Masters Program in Geospatial Technologies

MODELING MALARIA CASES ASSOCIATED WITH ENVIRONMENTAL RISK FACTORS IN ETHIOPIA USING GEOGRAPHICALLY WEIGHTED REGRESSION

Berhanu Berga Dadi

Dissertation submitted in partial fulfilment of the requirements for the Degree of *Master of Science in Geospatial Technologies*

NOVA IMS Information Management School
Universitat Jaume I
Westfälische Wilhelms-Universität Münster

MODELING MALARIA CASES ASSOCIATED WITH ENVIRONMENTAL RISK FACTORS IN ETHIOPIA USING THE GEOGRAPHICALLY WEIGHTED REGRESSION MODEL, 2015-2016

Dissertation supervised by

Dr.Jorge Mateu Mahiques,PhD

Professor, Department of Mathematics University of Jaume I
Castellon, Spain

Ana Cristina Costa, PhD

Professor, Nova Information Management School
University of Nova Lisbon, Portugal

Pablo Juan Verdoy, PhD

Professor, Department of Mathematics University of Jaume I
Castellon, Spain

March 2020

DECLARATION OF ORIGINALITY

I declare that the work described in this document is my own and not from someone else. All the assistance I have received from other people is duly acknowledged, and all the sources (published or not published) referenced.

This work has not been previously evaluated or submitted to the University of Jaume I Castellon, Spain, or elsewhere.

Castellon, 30th Feburaury 2020

Berhanu Berga Dadi

Acknowledgments

Before and above anything, I want to thank our Lord Jesus Christ, Son of GOD, for his blessing and protection to all of us to live. I want to thank also all consortium of Erasmus Mundus Master's program in Geospatial Technologies for their financial and material support during all period of my study. Grateful acknowledgment expressed to Supervisors: Prof.Dr.Jorge Mateu Mahiques, Universitat Jaume I(UJI), Prof.Dr.Ana Cristina Costa, Universidade NOVA de Lisboa, and Prof.Dr.Pablo Juan Verdoy, Universitat Jaume I(UJI) for their immense support, outstanding guidance, encouragement and helpful comments throughout my thesis work. Finally, but not least, I would like to thank my lovely wife, Workababa Bekele, and beloved daughter Loise Berhanu and son Nethan Berhanu for their patience, inspiration, and understanding during the entire period of my study.

MODELING MALARIA CASES ASSOCIATED WITH ENVIRONMENTAL RISK FACTORS IN ETHIOPIA USING THE GEOGRAPHICALLY WEIGHTED REGRESSION MODEL, 2015-2016

ABSTRACT

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. In this research, we explored spatial patterns and evaluated related potential environmental risk factors in the distribution of malaria cases 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 cases. Our results revealed a heterogeneous and highly clustered distribution of malaria cases in Ethiopia during the study period. The s-GWR model best explained the spatial correlation of potential risk factors with malaria cases and was used to produce predictive maps. The GWR model revealed that the relationship between malaria cases 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 cases. Hence, precipitation and relative humidity exhibited a varying relationship with malaria cases among the wards in both years. 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.

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

ACRONYMS

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

INDEX OF THE TEXT

INDEX OF TABLES

INDEX OF FIGURES

1 INTRODUCTION

1.1 Motivation, rationale and background

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 cases associated with environmental risk factors in of Ethiopia is very relevant. In Indonesia, (Hasyim et al. 2018) conducted a similar study. In their research, they didn't include climate factors data for their research data analysis, and they put that as a limitation of their research. Thus this research will try to solve that limitation by using the climate factors data and find out the association of climate factors with malaria cases.

1.2 Aim and objectives

This study aims to model malaria cases associated with environmental risk factors in Ethiopia, using geographically weighted regression in 2015 and 2016.
The specific objectives are:

- To map malaria risk areas (distinct) in the country.
- To map estimated malaria cases in the country.
- To investigate the impact of environmental risk factor in malaria cases distribution.
- To discuss and report spatial analysis results found.
- To model the association of environmental risk factors and malaria cases .

In addressing the problem, the following research questions were formulated for the study:

- Where are the annual malaria cases outbreak concentrated in 2015 and 2016?

- Which and where environmental risk factors are strongly associated with malaria cases.

1.3 Significance of the study

The finding of this research used to assist in planning, allocation of resource, drug distribution, and decision making concerning malaria control and monitoring in Ethiopia. In addition to this, the modeling of the spatial distribution of malaria cases associated with environmental risk factors is expected to help the country in preventing and controlling malaria. Moreover, the finding of this research can be used for other research as an input. This research work has an explicit significance for the researcher and used as a benchmark for interested researchers to explore the issues in the area for controlling and eradication of the epidemic. The outcome of the study also will provide information for government and nongovernment organizations to assist in malaria control and prevention in the country.

1.4 Structure of the report

This first chapter highlights the relevance of this research, lists the main objectives, and summarizes the methodological framework to address them. Section 2 dedicated to the literature review about malaria, particularly on modeling approaches to malaria cases. Chapter 3 dedicated to the methodological framework for modeling malaria cases associated with environmental risk factors using s-GWR. Chapter 4 highlights the result of all models and section 5 devoted to the discussion of all the results.

2 LITERATURE REVIEW

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 death 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 cases in 2017, which is about 92% or

2oo million malaria cases and followed by South East Asia, which is only 5% of malaria cases (WHO 2018). The annual report of malaria shows, over 80% of malaria cases 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 cases 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 cases 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 cases 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 et al. 2018).

Malaria is a leading cause of social and public health problems globally, including Ethiopia (WHO 2018). In Ethiopia, around 4-5 million malaria cases 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 et al. 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 Cases in Ethiopia (Ankamah et al. 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).

3 MATERIALS AND METHODS

In geospatial data analysis, nonstationarity is a condition in which a "global" model cannot clarify the relation or association between some sets of variables (Brunsdon, Fotheringham, and Charlton 2010). Therefore, besides global Ordinary Least Squares, local geographically weighted regression modeling used to analysis and model the association between environmental risk factors and malaria cases in Ethiopia at wards level. The local different of environmental risk factors to be studied potentially predict the response variable malaria cases (Y) are altitude (X1), Relative humidity (X2), Precipitation (X3), NDVI (X4) and rainfall (X5). Moreover, to generate a malaria risk map based on a statistically significant hotspot, this research work will use G* statistics (Yeshiwondim et al. 2009).

For this study, we used the R programming for data cleaning (de Jonge and van der Loo 2013), GWR4 (Acharya *et al.*, 2018;Manyangadze *et al.*, 2017;Hasyim *et al.*, 2018;Ge *et al.*, 2017) for modeling malaria cases, Arc Map (Acharya *et al.*, 2018;Manyangadze *et al.*, 2017;Hasyim *et al.*, 2018;Ge *et al.*, 2017) for interpolation independent variables, mapping malaria cluster, and mapping of the result of GWR and s-GWR. Malaria cases data was tested for spatial heterogeneity (non-stationarity) with Global Moran's I using GeoDa (Ge *et al.*, 2017; Fotheringham, Charlton and Brunsdon, 2002) as illustrated in (Table 1).

Table 1 Software used for analysis

Software used	Version	Comment
R programming	3.6.2	Package library(tidyverse) library(psych) library(skimr) library(broom)
ArcMap	10.6	
GWR	4.0	
GeoDa	1.14	

3.1 Study area and malaria dataset

Ethiopia is located in the eastern part of African content approximately 30 -150N latitude and 330 -480E, longitude. Its land and water coverage is 1,000,000 and 104,300 square kilometres, respectively (Hagose 2017) (Figure 1). 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 (Figure 1) uses the World Geodetic System (WGS84) map projection as its reference coordinate system for data analysis. As shown in Figure 3, in spatial data analysis, three stages of working with spatial data were eminent: data acquisition and processing, data analysis and data presentation (Hasyim et al. 2018). GWR 4.0 version 4.09, Geoda and Arc GIS 10.6 were used for data processing, analysis, and visualization. Malaria cases data were collected from Ethiopian Public Health Institute for all wards (i.e. administrative units similar to counties or districts).

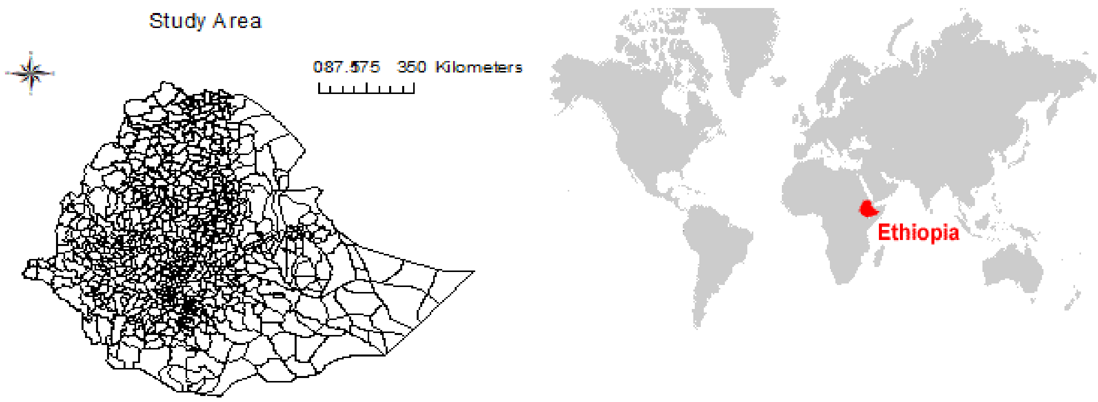


Figure 1: Map of the study area covering 679 wards (counties) of the entire country, Ethiopia.

The Ethiopia Public Health Institute summarized weekly malaria cases in each ward in the study area between 2015 and 2016 (96 weeks in all). The malaria data categorized into “clinical diagnosis,” and “confirmed malaria cases.” The total number of malaria cases were 39592.14 over 2015-2016 in the study area. Figure 2 shows the spatial distribution of annual malaria cases in each year.

Malaria cases data were collected from 679 counties (wards) from 2015 to 2016 in the study area. Malaria incidence was computed as malaria cases divided by population and multiplied by 1000.

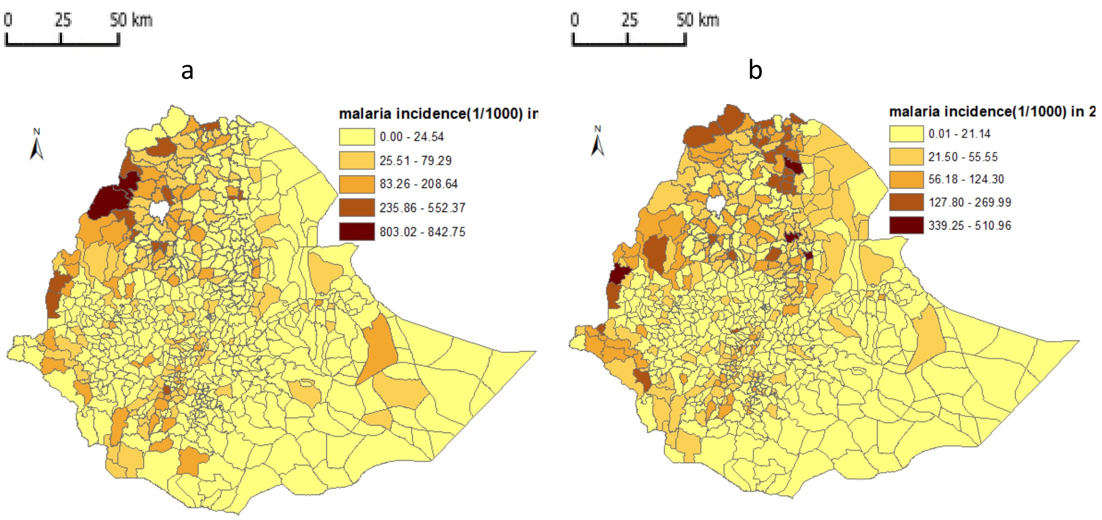


Figure 2: Annual average malaria incidences in each county from 2015 to 2016

3.2 Climate and environmental data

Research studies on climate variability do not show any consistent pattern or trends in the country (Mengistu, Bewket, and Lal 2014). In Ethiopia, different studies on key climate variability and trend indicators have also been conducted by (Osman and Sauerborn, 2002; Dereje Ayalew, 2012;Jury and Funk, 2013;Abtew, Melesse and Dessalegne, 2009;Taye, Zewdu and Ayalew, 2013;Viste, Korecha and Sorteberg, 2013; Mengistu et al., 2014).

Variables shown in Table 2 characterize the basic climate and environmental conditions of each county (ward). The predictors were selected based on their probable association among malaria cases, taking into consideration the literature review and data availability. Previous studies have demonstrated an association of socioeconomic variables with malaria, such as population density (persons/km2), persons (immigrant population), and gross domestic product (GDP) (Ge et al. 2017).

The climate and environmental predictors considered in this study, as well as their descriptions, are listed in Table 2. 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.

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 (Figure 3).

Population data (2012) were collected from the Ethiopian Central Statistical Agency and were used to compute the malaria cases.

Table 2: Variables used in the research and their sources

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

3.3 Data pre-processing and modeling

The Geographically Weighted Regression (GWR) modeling approach was applied for exploring the association among malaria cases and local spatial predictors across the study area. Figure 4 depicts the schematic overview of the methodology. The malaria cases data were tested for spatial autocorrelation, and the explanatory variables were obtained for the 679 wards by interpolation techniques as a first step. The subsections below detail each stage of the methodological framework.

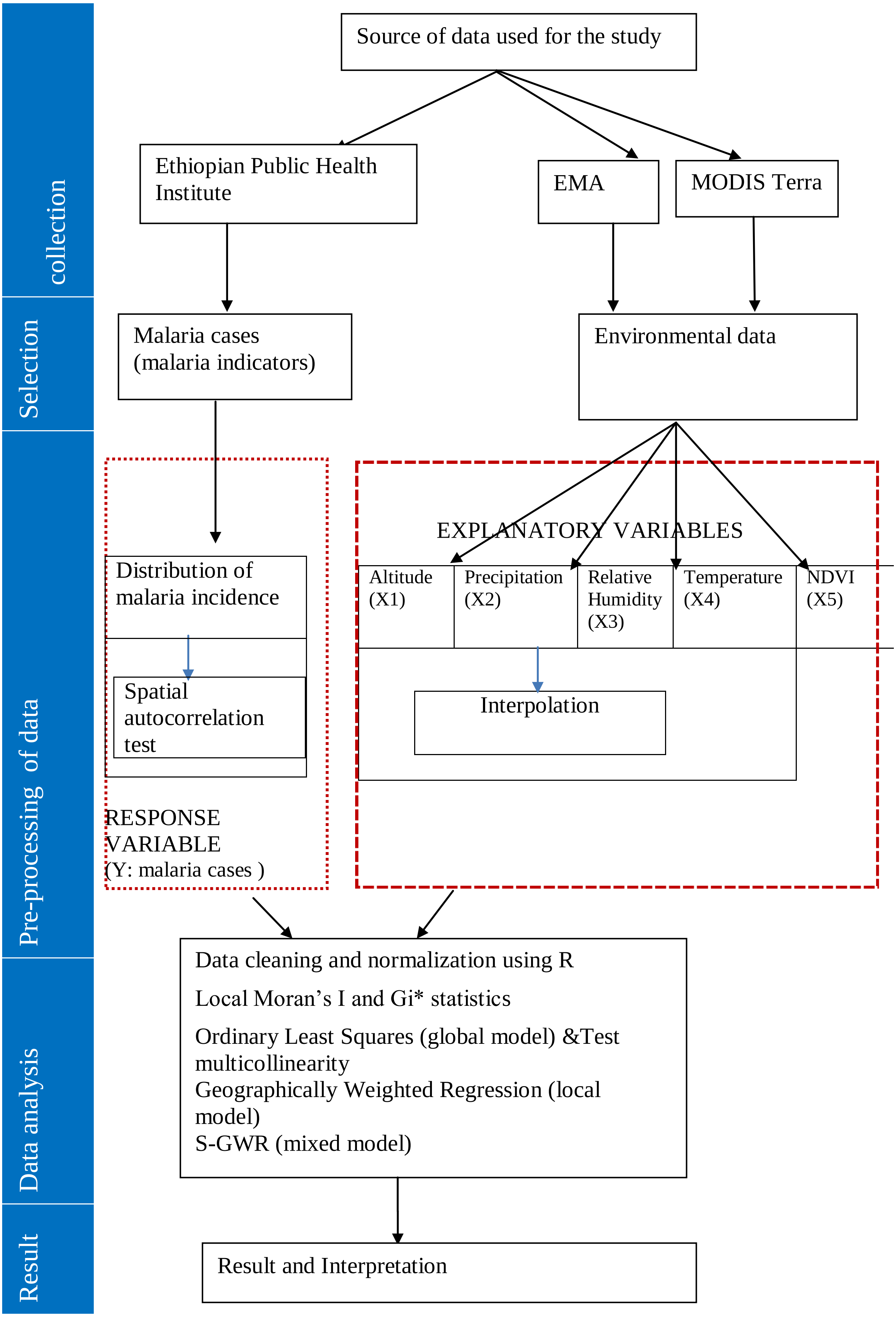
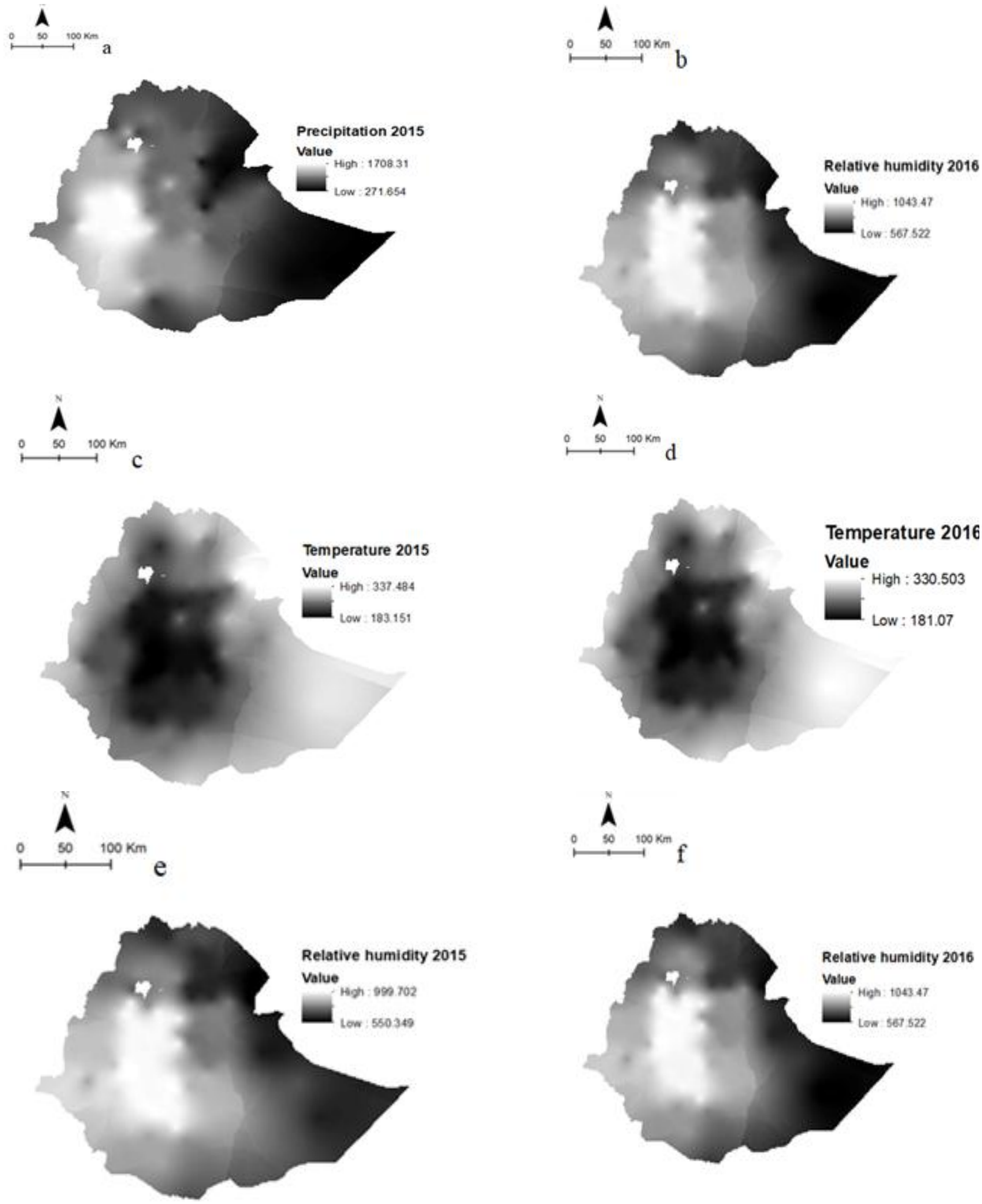


Figure 3: Flow chart of the research approach

3.3.1 Environmental data

Figure 4 depicts the kriging interpolation result of precipitation, temperature, relative humidity, and elevation for entire study area as shown bleow.



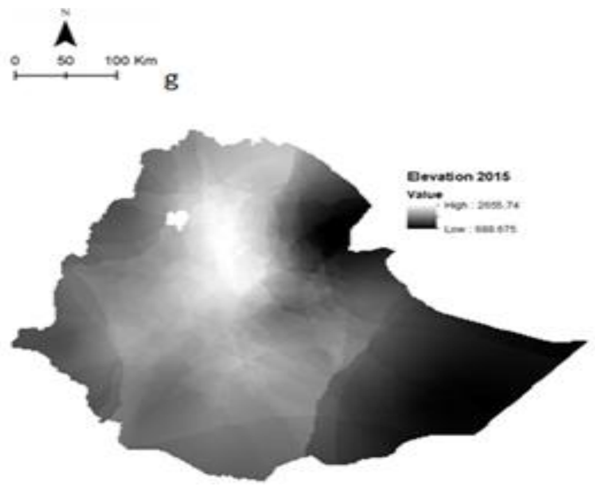


Figure 4: Each explanatory variable mapped 2015(left) and 2016(right) in the study area.

All environmental risk factors variable were interpolated by using Kriging Spherical models as is illustrated in Table 3. Mean yearly interpolated climate factors data was used for analysis.

Table 3: Variogram models used for interpolation of explanatory variables with ordinary kriging

Dataset	**Model and parameters**	**Model name and values**
Temperature	model	Spherical
	nugget	0.001
	range	4.2
	Partial sill	1.067
Precipitation	model	Spherical
	nugget	0.0009
	range	5.24
	Partial sill	0.998
Relative humidity	model	Spherical
	nugget	0.0010
	range	4.611
	Partial sill	1.025
Elevation	model	Spherical
	nugget	0.001049
	range	4.0217
	Partial sill	1.0495

3.3.2 Data cleaning and normalization

Data cleaning and exploring were done by using R software. Logit and log malaria cases were used to discover if bivariate relationship were linear or not using B-splines explanatory analysis (Figure 17) in Appendix 7.1.

3.3.3 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):

$$Gi*(d) = \frac{W_{ij}(d)_{xj}}{\sum_{i\ j} x} \qquad (1)$$

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 cases region. In contrast, a high negative Z-score value indicates that the location separate i is surrounded by relatively low (cold spot) malaria incidence in distinct areas. Otherwise, random distribution of malaria cases for high and negative value of Z ≥-1.96 and ≤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 map the local clusters of high malaria cases. 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

$$I_i = z_i \sum_j w_{ij} z_j \qquad (2)$$

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 weights matrix w_{ij}.

3.3.4 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 et al. 2002). GWR determines the spatial variabiliy 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 et al. 2002). In the first step, the average annual malaria cases 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 cases, 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 cases; 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 colinearity 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 cases and the selected envriomental risk factors. The method of least square expressed in the following equation (Acharya et al. 2018):

$$y_i = a + \sum_{j=1}^{k} a_j x_{ij} + \epsilon_i \tag{3}$$

Where y_i is the ith examination of the response variable, $a_j x_{ij}$ is the *i*th examination of the *K*th explanatory variable, and ε_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

$$y(u) = \beta_0(\boldsymbol{u}) + \sum_{t=1}^{n} \beta_t(\boldsymbol{u}) x_t + \epsilon \tag{4}$$

Where y is the value of malaria cases at the location $\mathbf{u}$, x_t is the value of explanatory variable t at the location $\mathbf{u}$, β_t (u) is the regression coefficient at the location $\mathbf{u}$, and ε is the random error with mean 0 and variance σ^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 et al. 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 et al. 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 decide the best bandwidth size for geographically weighting efficiently. The best bandwidth and the related weighting function were attained by selecting the smallest AICc score. The fixed Gaussian kernel for geographical weighting used in this study(Nakaya 2016) is as follows:

$$w_{ij} = \exp\left[-(d_{ij}/b)^2\right] \qquad (4)$$

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 $\beta_t(\mathbf{u})$ of the explanatory variable t at the location $\mathbf{u}$. If one explanatory variable X_t (i.e. environmental risk factor) has a positive [negative] coefficient at the location $\mathbf{u}$, it means that when X_t increases at the location $\mathbf{u}$, it is expected that the malaria cases(Y) increases [decreases] at the location $\mathbf{u}$, 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 et al. 2002):

$$y_i = \sum_{j=1}^{k_a} \alpha_j x_{ij} + \sum_{l=k+1}^{k_b} \beta_l (u_i v_i)\, x_{il} + \epsilon_i \qquad (5)$$

where for observation i, $(u_i v_i)$ is the geographical location, $\{\alpha_1 \dots \alpha_{k_a}\}$ are the k_a global coefficients associated with the set of global explanatory variables $\{x_1 \dots x_{k_a}\}$, $\{\beta_1(u,v) \dots \beta_{k_b}(u,v)\}$ are the k_b local coefficient functions associated with the set of local explanatory variables $\{x_1 \dots x_{k_b}\}$

The selected environmental variables (Elevation, precipitation, temperature, NDVI, and relative humidity) correspond to the explanatory variables, and the pre-processed annual average malaria cases 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 cases and residuals) were mapped using the ArcGIS10.6. Mapping local parameters make a straightforward interpretation based on recognized characteristics and spatial background of the research 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 pseudo-t values exceed ± 1.96 were considered as relevant (Ehlkes *et al.*, 2014; Acharya *et al.*, 2018; Wabiri *et al.*, 2016; Matthews and Yang, 2012).

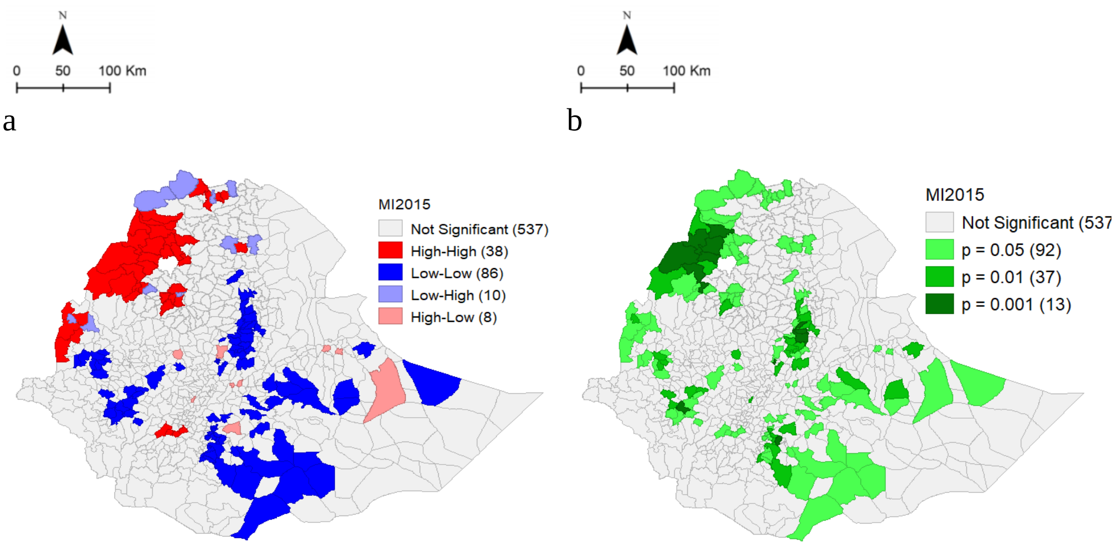
4 RESULTS

4.1 The spatial analysis and distribution of malaria Cases

A total of 19687.31 cases in 2015 and 456807.8 cases 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 cases of 0.197 per 1000 and 4.461 per 1000 inhabitants in 2015 and 2016, respectively.

The malaria cases were distributed in the wards, as shown in Figure 5. The results of the global spatial autocorrelation test for the 2015 and 2016 years data showed significant spatial dependence in several wards for all yearly cases: minimum Moran's I = 0.323, $p < 0.05$ (2015), and maximum Moran's I = 0.514, $p < 0.05$ (2016). The significant local spatial autocorrelation result for malaria cases ensured the suitability of the malaria cases data as the response variable in GWR models.

Local spatial autocorrelation of each year was assessed with local indicators of spatial association (LISA) (Anselin 1995), namely the Local Moran's I statistic Figure 5 and the Getis-Ord Gi* statistic (Figure 6). The Local Moran's I maps showed the hot-spot regions (high-high) and cold-spot regions (low-low), where hot-spot location means the malaria cases of a particular spatial unit is high, and malaria cases of its surrounding units are also high, and cold-spot region implies the opposite.



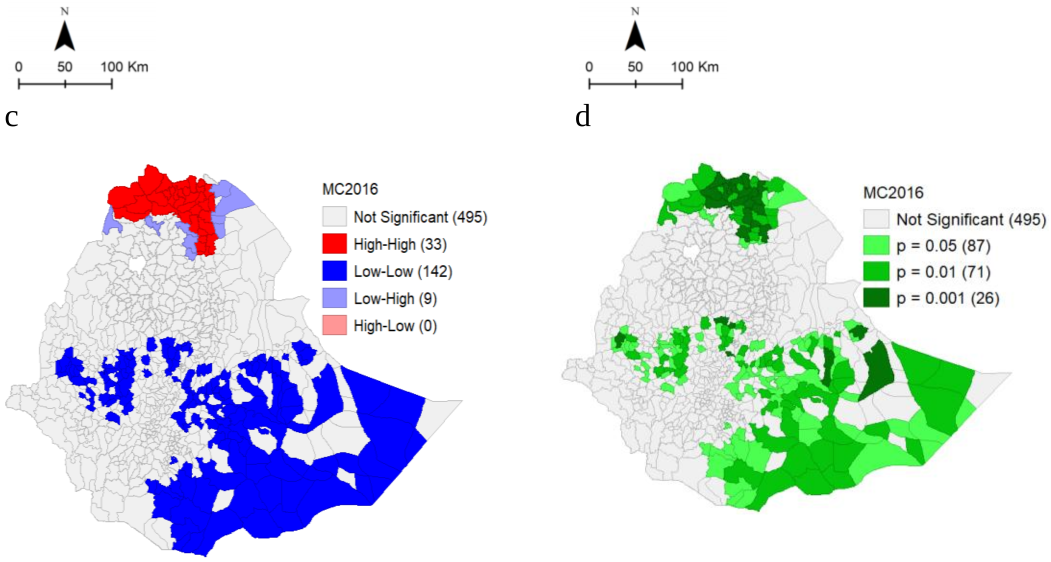


Figure 5 Local Moran's I test maps of malaria cases and corresponding significance for 2015 (top) and 2016 (bottom).

The hot-spot locations of malaria cases in 2015 Figure 5a were the wards along the west, the northwestern part of the country, whereas the cold-spot were along the west, the central and southeastern part of the country. The hot-spot locations of malaria cases in 2016 Figure 5c were very high malaria cases in the wards along the northwest, and the west northern part of the country, whereas the cold-spot were wards along the western, central, and southeastern part of the country. A few zones exhibit local negative spatial autocorrelation where wards with low values of malaria cases correlate with high neighbouring values (10 and 9 Low-High wards in 2015 and 2016 respectively), and wards with high values of malaria cases correlate with low neighbouring values (8 High-Low wards in 2015).

Figure 6 depicts the annual malaria cases 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 cases hot-spot distribution was along the north and northwestern region of the country. In contrast, in 2016, the annual malaria cases hot-spot distribution was along the northern part of the county (Figure 6b). These results also highlight the spatial nonstationarity of malaria cases.

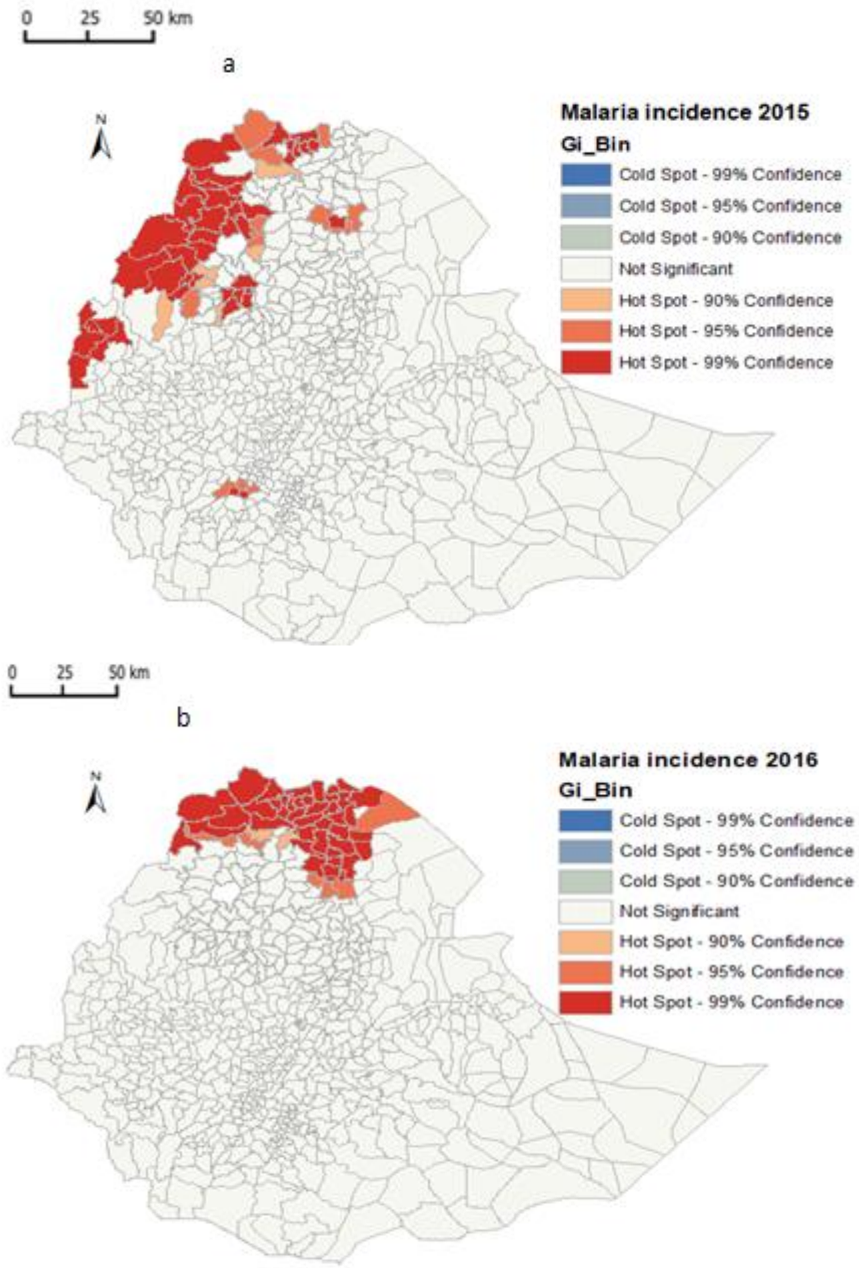
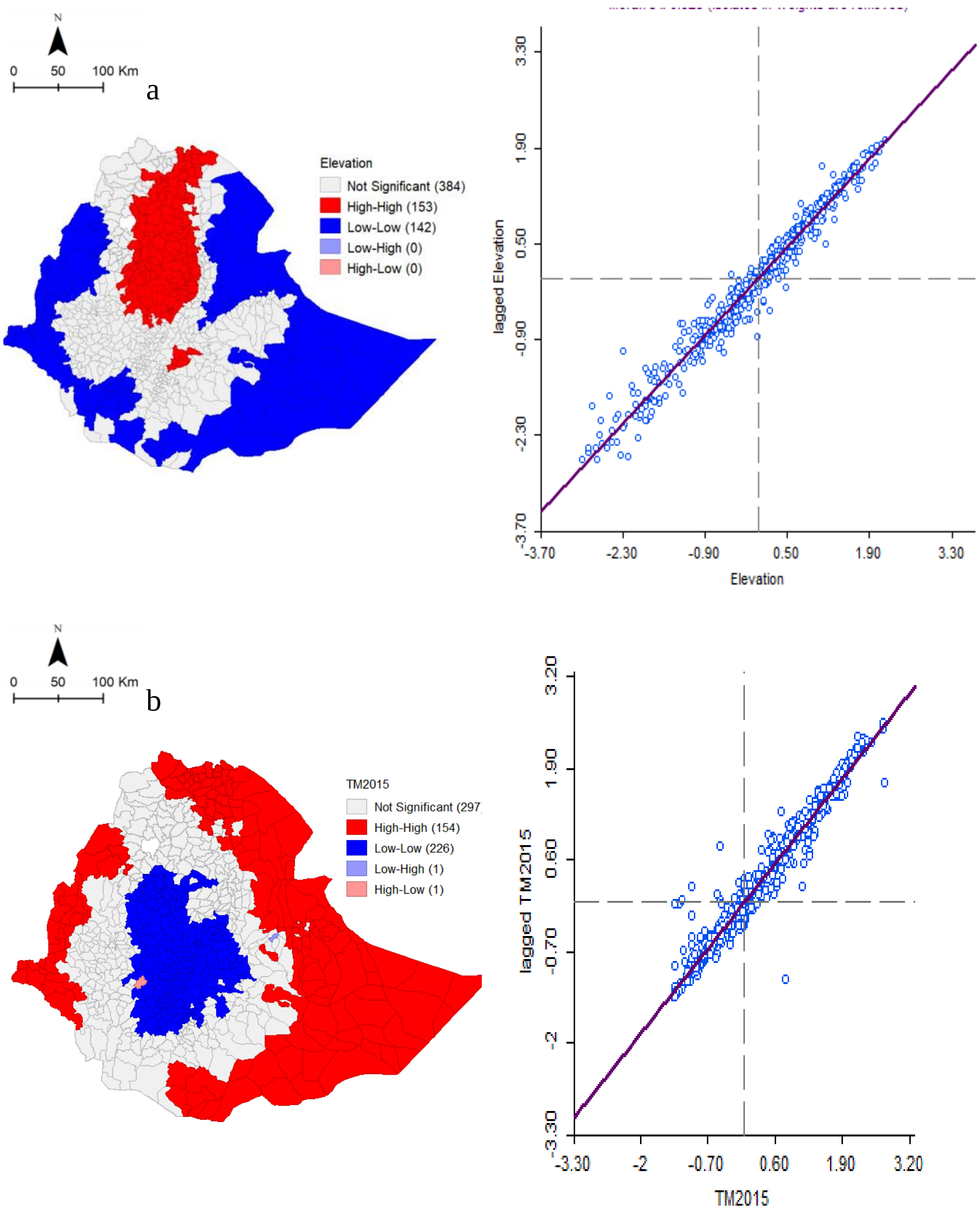


Figure 6. Hot-spot (Getis-Ord Gi* statistic) results of malaria cases in 2015 (a) and 2016 (b)

4.2 Spatial analysis of predictors

The explanatory variables were assessed using spatial autocorrelation, and it was found to be significant for five of the independent variables Figure **7**. A local Moran's I analysis result identified along with the hot-pot and cold-spot distributions of the five variables. The hot-spot locations of Elevation Figure 7a were the wards along the north-central part of the country. The temperature hot-spot areas were in the southern-eastern, northeastern lowlands and western regions of the country as depicted in Figure 7b. Relative humidity hot-spot distribution was along the central areas of the country as drew in Figure 7c. Precipitation hot-spot distribution was wards along the Western and Central regions of the country Figure 7d. Figure 7e showed hot-spots in the normalized difference vegetation index (NDVI) distribution along the southwestern part of the country.



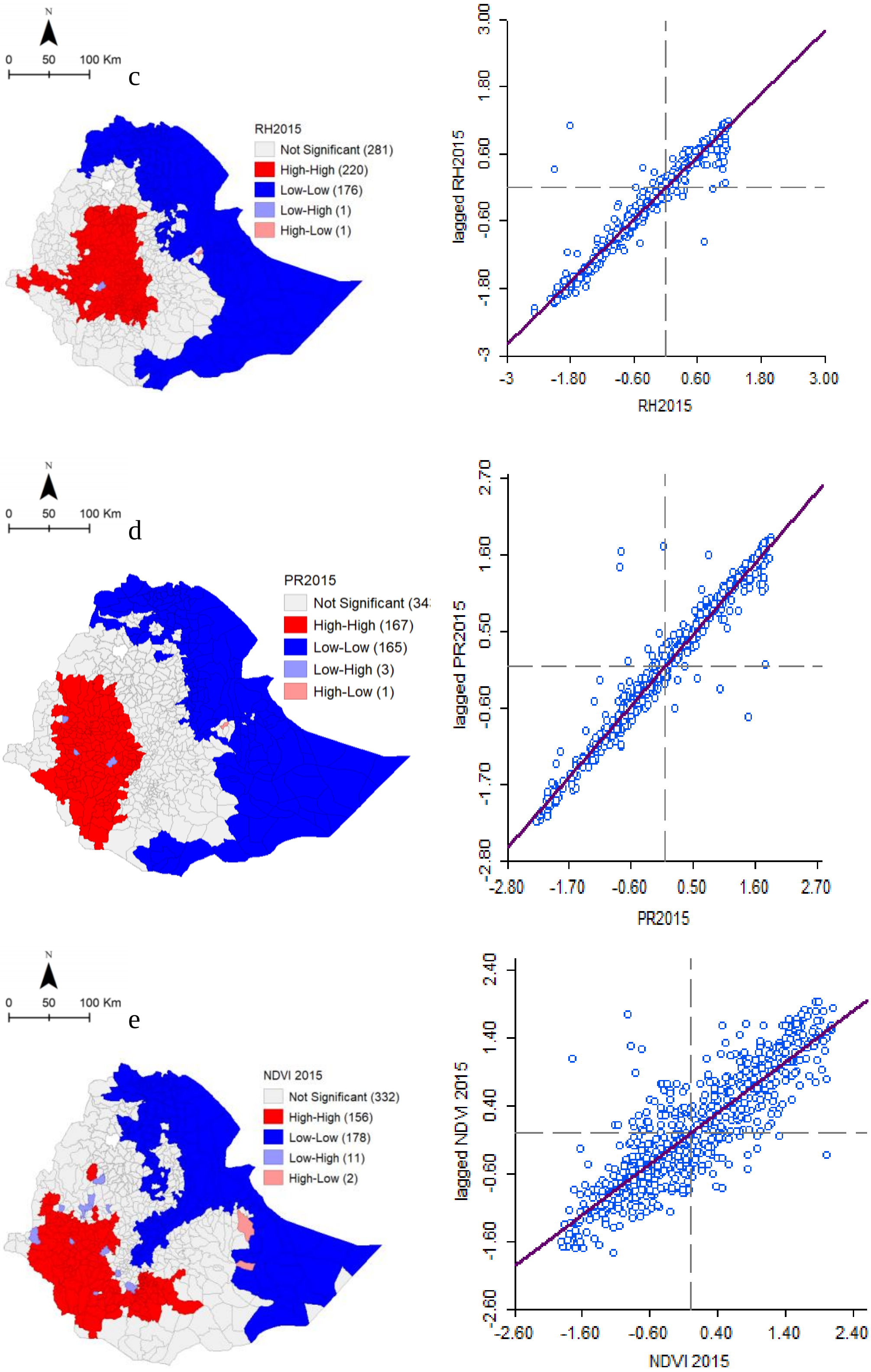


Figure 7: Distribution of selected explanatory variables with their corresponding local (1st column) and global Moran's I tests (2nd column)

4.3 Ordinary Least Squares model

The coefficients of the Ordinary Least Squares models have the same value for all points within the study area (Table 4) and (Table 5). 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.

In the two models (2015, 2016), all independent variables have VIF<4, so there is no evidence of multicollinearity among them as shown in Table 4 and Table 5. Therefore, it is appropriate to use them in the local models.

Table 4: Summary of OLS Results - Model Variables for 2015

Variable	Coefficient	StdError	t-Statistic	p-value	Robust StdError	Robust_t	Robust 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 5: Summary of OLS Results - Model Variables for 2016

Variable	Coefficient	StdError	t-Statistic	p-value	Robust StdError	Robust_t	Robust p-value	VIF
Intercept	-9252.87	2844.84	-3.252	0.001*	3677.108	-2.516	0.012*	……
Elevation	2.895	0.34	8.451	0.000*	0.551	5.251	0.000*	1.731
TM2016	411.115	69.41	5.922	0.000*	100.407	4.094	0.001*	3.807
PR2016	6.261	4.97	1.258	0.208	2.211	2.831	0.004*	2.096
RH2016	-51.043	21.51	-2.372	0.018*	25.473	-2.003	0.045*	3.951
NDVI2016	-0.027	0.08	-0.331	0.741	0.053	-0.508	0.611	1.655

4.4 Geographically Weighted Regression model

The GWR models were used to explore the local effects of variables on malaria cases 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 8 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 Figure 8, with a statistically significant positive association in red/orange and negative statistically significant relationship in green/light green. Figure 9(a-e) and Figure 10(a-e) 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 6 and Table 7 summarize the values of the maps of GWR local coefficients in Figure 9(a-e) and Figure 10(a-e), and also show global adjustment measures (R^2, Adjusted R^2 and AICc). Despite the higher Adjusted R^2 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.

Table 6. Summary of the locally varying coefficients of the variables on the GWR model in 2015.

Coefficients	Minimum	Lower quartile	Median	Upper quartile	maximum
Intercept	-113.919	9.149	29.276	63.560	431.059
Elevation	-237.415	-42.757	-13.933	-4.582	177.795
TM2015	-143.665	-9.475	2.745	17.829	294.543
RH2015	-362.737	-11.673	7.165	17.128	583.187
PR2015	-206.090	-15.310	-1.958	5.260	339.122
NDVI2015	-120.831	-8.291	-2.489	2.265	56.994

R^2 = 0.630, Adjusted R^2 =0.515, AICc= 7311.884

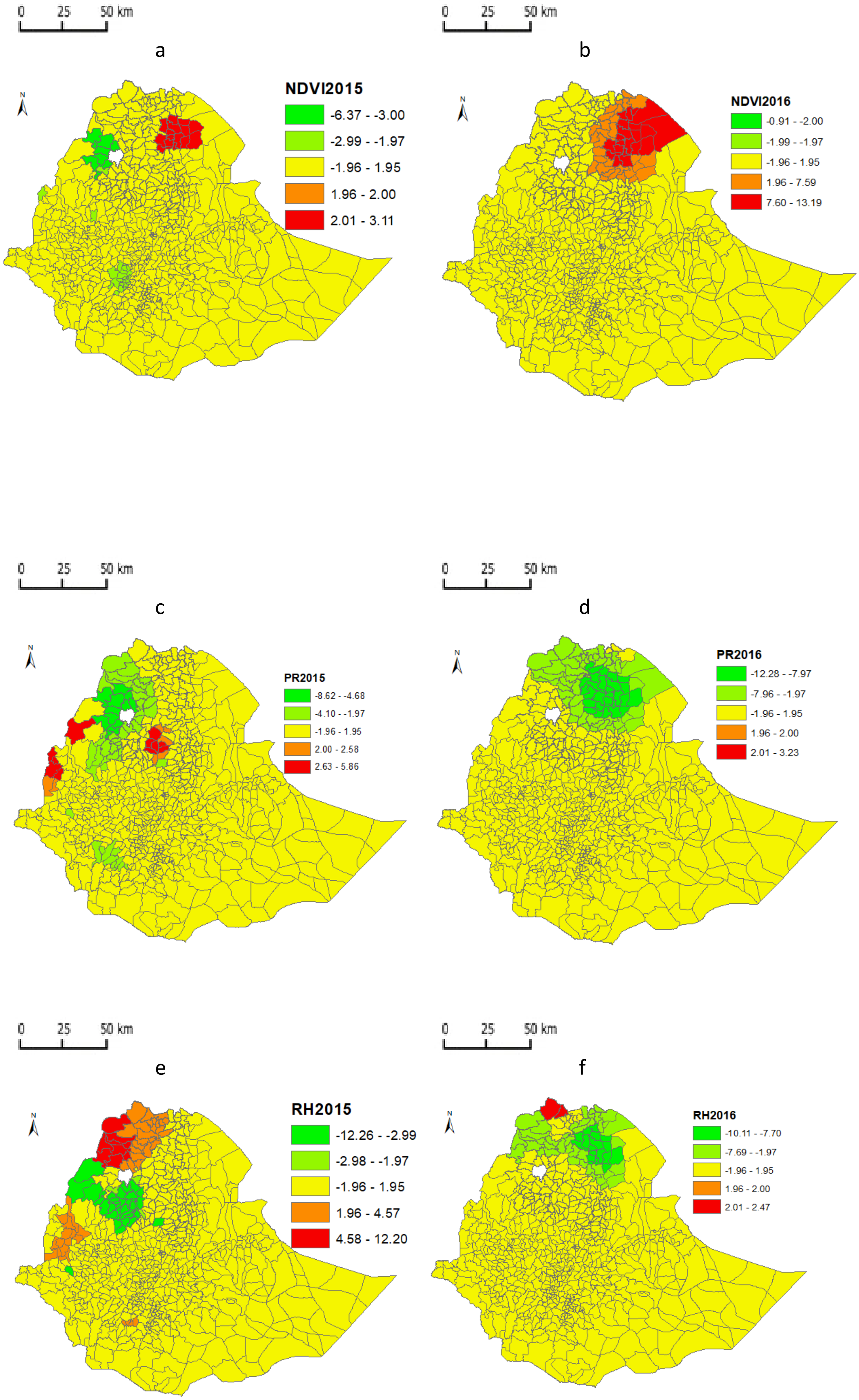
Table 7 Summary of the locally varying coefficients of the variables on the GWR model in 2016.

Coefficients	Minimum	Lower quartile	Median	Upper quartile	maximum
Intercept	-13284.062	5.503	15.931	30.822	9390.445
Elevation	-4007.037	-15.514	-2.653	11.094	4683.572
TM2016	-4299.357	-11.299	0.523	16.990	4500.589
PR2016	-19773.517	-26.501	-5.672	1.107	5874.614
RH2016	-13530.506	-26.221	0.522	9.315	4867.645
NDVI2016	-798.797	-4.291	-0.748	6.899	10148.366

R^2= 0.680, Adjusted R^2= 0.608, AICc=12349.729

In 2015, temperature coefficients showed a positive and negative correlation with malaria case per wards and were significant in some wards located to the northwestern, southwestern part of the study areas Figure 8g. Elevation 2015 estimated coefficients showed a positive and negative relationship with malaria cases per ward and was significant in some wards located to the northern, southwestern, and southern part of the study areas (Figure 8i). In 2015, the estimated NDVI coefficients showed a positive and negative relationship with malaria cases per wards and were significant in some wards located to the northwestern, northeastern and southern part of the country Figure 8a. In 2015 precipitation estimated coefficients showed a positive and negative correlation with malaria cases per wards and were significant in some wards located to the western, northwestern, southwestern, and south-central parts of the study area (Figure 8c). In 2015 relative humidity estimated coefficients showed a positive and negative association with malaria cases per wards and were significant in some wards located to the northern and western part of wards in the study area (Figure 8d). In 2016 NDVI estimated coefficients depicted only positive correlation with malaria cases per wards and were significant in some wards located to the northeastern wards of the country (Figure 8b).

In 2016 Precipitation estimated coefficients showed an only negative relationship with malaria cases per wards and were significant in some wards located to the northern part of the country (Figure 8d). In 2016 relative humidity, temperature, and elevation estimated coefficients depicted a positive and negative correlation with malaria cases per wards and were significant in some wards located to the northern part of the country (Figure 8h, f, and j).


0 25 50 km
a
NDVI2015
-6.37 - -3.00
-2.99 - -1.97
-1.96 - 1.95
1.96 - 2.00
2.01 - 3.11
0 25 50 km
b
NDVI2016
-0.91 - -2.00
-1.99 - -1.97
-1.96 - 1.95
1.96 - 7.59
7.60 - 13.19
0 25 50 km
c
PR2015
-8.62 - -4.68
-4.10 - -1.97
-1.96 - 1.95
2.00 - 2.58
2.63 - 5.86
0 25 50 km
d
PR2016
-12.28 - -7.97
-7.96 - -1.97
-1.96 - 1.95
1.96 - 2.00
2.01 - 3.23
0 25 50 km
e
RH2015
-12.26 - -2.99
-2.98 - -1.97
-1.96 - 1.95
1.96 - 4.57
4.58 - 12.20
0 25 50 km
f
RH2016
-10.11 - -7.70
-7.69 - -1.97
-1.96 - 1.95
1.96 - 2.00
2.01 - 2.47

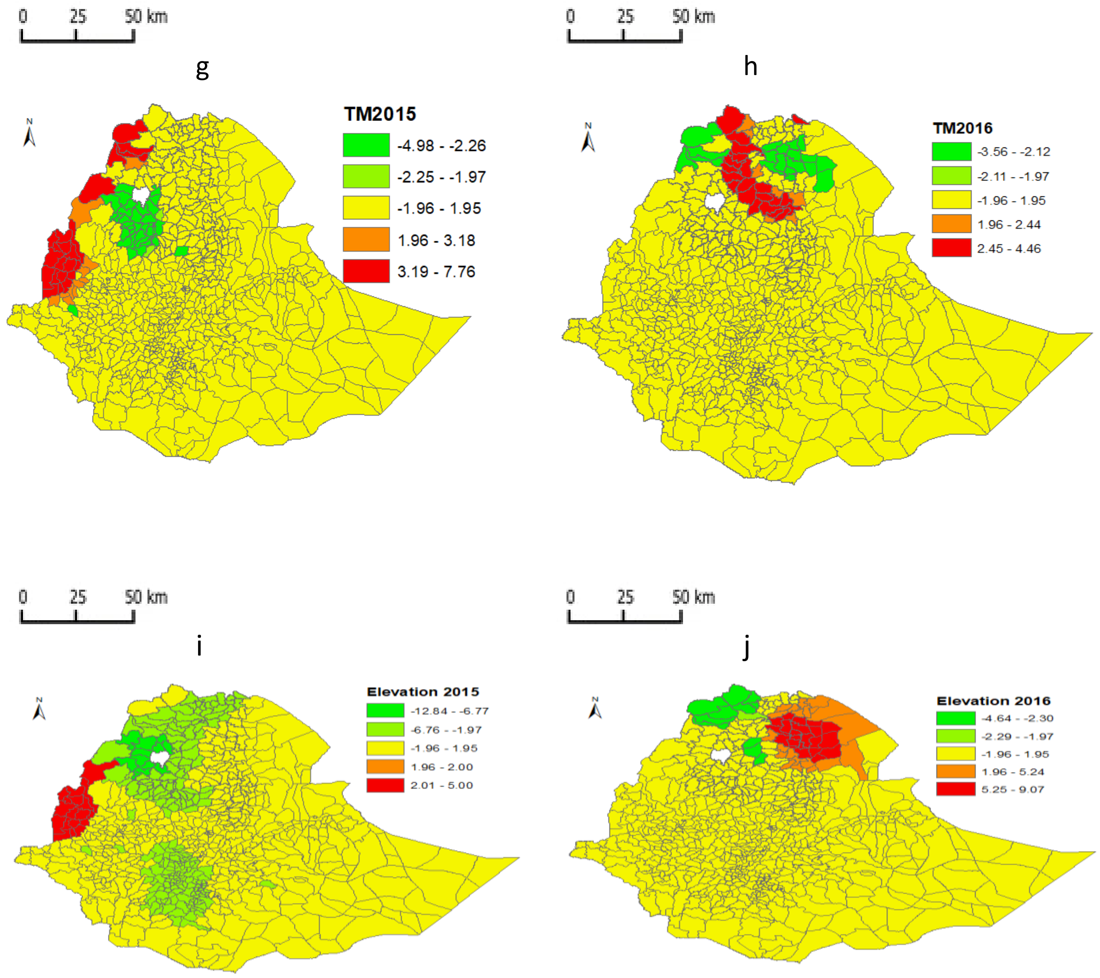
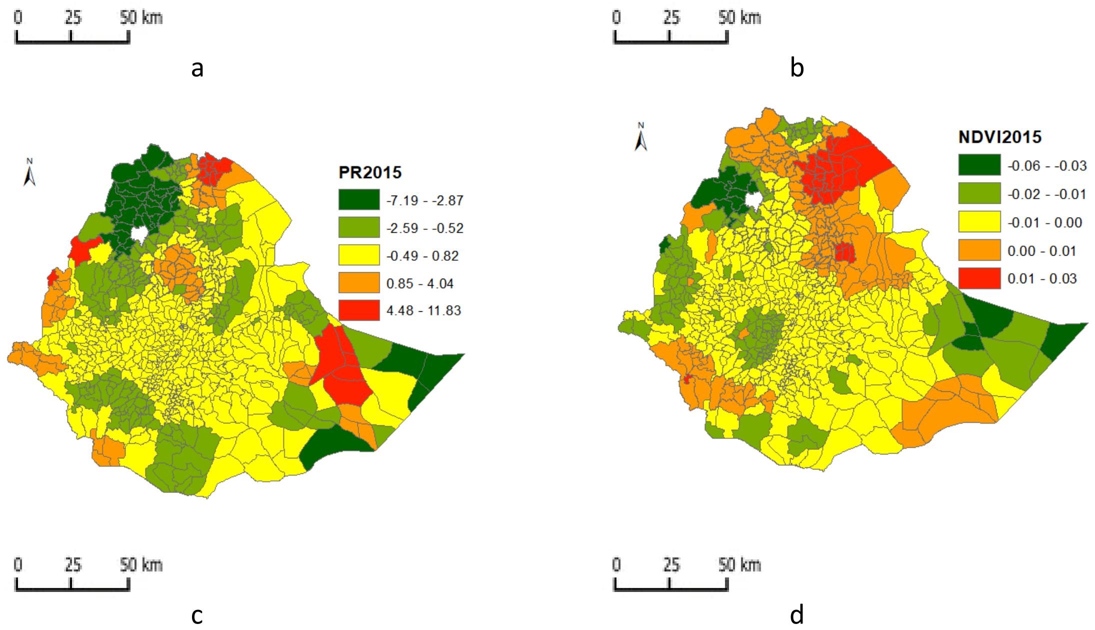
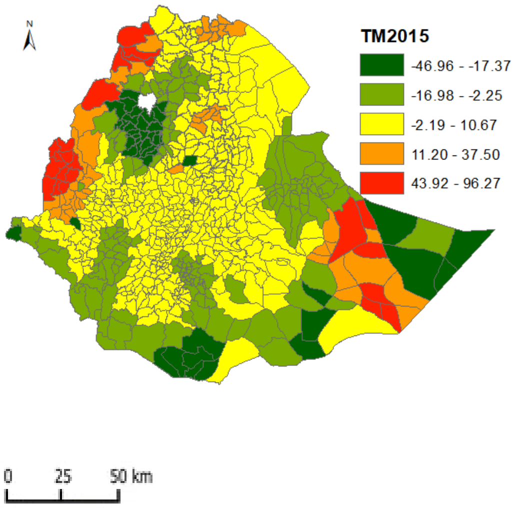
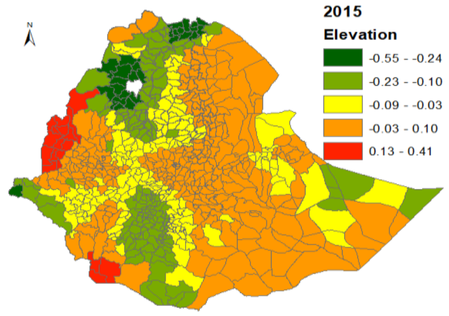


Figure 8. Pseudo t-values for independent variables in 2015 (left) and 2016 (right)







e

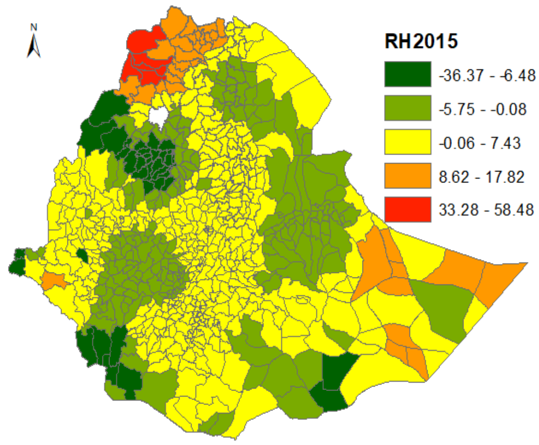


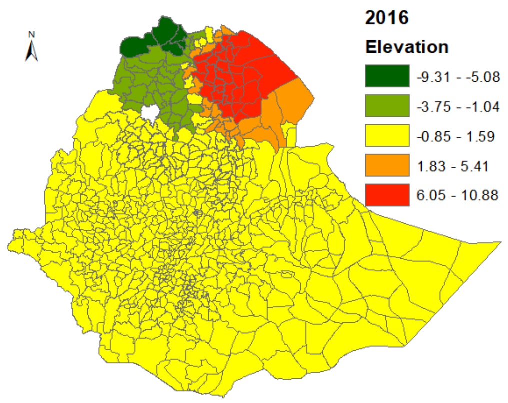
Figure 9. GWR local coefficients of the 2015 model (a-e)

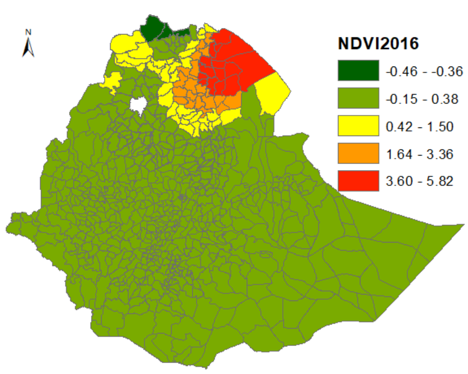
0 25 50 km

a

0 25 50 km

b





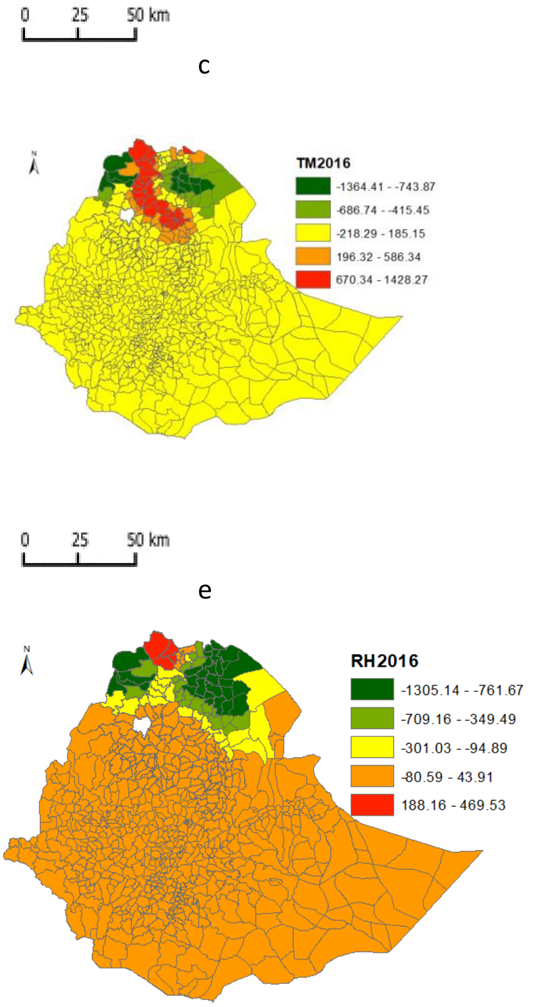


Figure 10.GWR local coefficients of the 2016 model (a-e)

In 2015 temperature is significantly and positively related to malaria cases in the following 28 Wards:

- Gidami,
- Gawo Kebe,
- Maok Omo,
- Mana Sibu,
- Menge,
- Sirba Abaya,
- Guba,
- Hawa Gelan,
- Boji Chekorsa,
- Jimma Horo,
- Babo,
- Begi,
- Bambasi,
- Homosha,
- Kumuruk,
- Anfilo,
- Dale sadi,
- Nejo,
- Dale Wabera,
- Gudetu Kondole,
- Kiltu kara,
- Assosa,
- Biligidillu,
- Sherkole,
- Yama Logi welel,
- Ayira Guliso,
- Agalmoeti,

- Qura,
- Kafta Humera,
- Mirab Armacho,
- Tach Armacho.
- Tsegede,

Temperature was also significant and negatively related to malaria cases in the following 31 wards (Figure 8g):

- Amuru,
- Michakel,
- Dembecha,
- Ankasha,
- Bibugn,Banja,
- Fagta lakoma,
- Pawe,
- Bahirdar Zuria,
- Fogera,
- Dera,
- Debra Elias,
- Bure,
- Senan,
- Guagusa,
- Sekela, Quarit,
- Yilmana Densa,
- Anchefer,
- Dera ,
- Libo Kemekm,
- Bure,
- Guzamn,
- Wemberama,
- Jabi Tehnan,
- Dega Damot,
- Gonje,
- Dangila,
- Mecha,
- Jawi,
- Takusa ,
- and Alfa.

Elevation 2015 estimated coefficient was significant and positively related to malaria cases in the following 18 wards in the country:

- Gidami,
- Gawo Kebe,
- Maok Omo,
- Mana Sibu,
- Homosha,
- Kumuruk,
- Jimma Horo,
- Babo,
- Begi,
- Bambasi,
- Biligidillu,
- Sherkole
- Dale Wabera,
- Gudetu Kondole,
- Kiltu kara,
- Assosa, Menge,
- Sirba Abaya,
- and Guba.

Moreover, Elevation 2015 was also significant and negatively related to malaria cases in the following 160 wards in the country (Figure 8i):

- Asgede Tsimbila,
- Kola Temben,
- Tahtay koraro,
- Afeshum,
- Ahferom,
- Mirab Armacho,
- Addi Arekay,
- Welkait,
- Farta,
- Ebenat,
- Takusa,
- GonderZuria,
- Dabat,
- Pawe,
- yilmana Densa,
- Bahirdar Zuria,
- Limu,
- Awabel,
- Dera,
- Shebel Bereta,
- Aneded,
- Medebay,
- Degua Temben,
- laelay Maychew,
- Erop,
- Mereb Leka,
- Tsegede,
- Beyeda,
- Alfa,
- Lay Gayint,
- west Belesa,
- Chilgam,
- Lay Armachewo,
- Metema,
- Dangila,
- West Esite,
- Debub Anchefer,
- Ababo,
- Dejen,
- Wegde,
- Dejen,
- Baso liben,
- Naeder Adet,
- Hawzen,
- Adwa,
- Gulomekeda,
- laelay Adiyabo,
- Debark,
- Tselemt,
- Fogera,
- Libo Kemkem,
- East Belesa,
- Dembia,
- Wegera,
- Danguara,
- Mecha,
- Dera,
- Jawi,
- Baso liben,
- Wara Jarso,
- Debresina,
- Awabel,
- Aneded,

- Guzamn,
- Dembecha,
- Debay,
- Enar Enawa,
- Wemberma,
- Dega Damot,
- Sekela,
- Ameya,
- Tocha,
- Boloso Bombe,
- Omo Nada,
- Endiguagn,
- Ezha,
- limu kosa,
- Ameya,
- Koochere,
- Bule Hora,
- Borke,
- zala,
- Chencha,
- Wenagol,
- Afele Kola,
- Hulla,
- Humbo,
- Denibu Gofa,
- Sodo Zuria,
- Boloso Sore,
- Awasa Zuria,
- limu,
- Ofa,
- Tibaro,
- Dedo,
- Shashemene zuria
- Debre Elias,
- Michakel,
- Telatgen,
- Jabi Tehnan,
- Guangua,
- Sekela,
- Mandura,
- Nono,
- Mareka,
- Kacha Bira,
- Sekoru,
- Admi Tulu,
- Cheha,
- Sekoru,
- Nono,chora,
- Gedeb,
- Hambela,
- Arba Minch Zuria,
- Daramalo,
- Abaya,
- Bule,
- Bore,
- Bursa,
- Loka Abaya,
- Kucha,
- Damot,
- Bomb,
- Goro,
- Kacha biraa,
- Kindo Dida,
- Dune,
- Omo Nada,
- And Bure
- Bure,
- Senan,
- Enemay,
- Bure Dembecha,
- Ankasha,
- Quarit,
- Fagta Lakoma,
- Goro,
- Gena Bosa,
- Tibaro,
- Yem, Gibe,
- Selti,
- Chora,
- Tiro Afeta,
- Amaro SP,
- Kercha,
- Wamena,
- kemba,
- Dita,
- Dila Zuria,
- Yirgachefe,
- Dara,
- Dale,
- Zala,
- Boreda,
- Boricha,
- Siraro,
- Arsi Negeli,
- loma Bosa,
- Mareka,
- Daniboya,
- Goro,

In 2015 the estimated NDVI local coefficient was significant and positively related to malaria cases in the following 15 wards in the country:

- Gaz gibla,
- Sekota,
- Yalo,
- Megale,
- Saharti Samre,
- Alamata,
- Endamehoni,
- Teru,
- Erebti,
- Ab Ala,
- Olfa,
- Raya Azebo,
- Alaje,
- Hintalo Wejirat,
- and Enderta.

NDVI also significant and negatively related to malaria cases in the following 54 wards in the country (Figure 8a):

- Mereb leke,
- Chilga,
- Kindo dida,
- Kindo koysha,
- Damot pulasa,
- Gena Bosa,
- Badawacho,
- Hadero Tubito,
- Gena,
- Dedo,
- Gembora,
- Sekoru,
- sekoru,
- wilbareg,
- Gawo kebe,
- Babo,
- Pawe,
- and Metema
- Quara,
- Dembia,
- Ofa,
- Damot sore,
- Mareka,
- Boloso Bombe,
- Tibaro,
- Dune,Soro,
- Daniboya,
- Omo Nada,
- LImu,
- Gibe,
- Alfa,
- Yama logi welel,
- Dale wabera,
- Gumurk,
- Achefer,
- Enemorina Eaner
- Takusa,
- Loma bosa,
- Sodo zuria,
- Damot gale,
- Tocha,
- Boloso sore,
- Kacha Bira,
- Doya,
- Shashogo,
- Yem,
- Analemmo,
- Tiro Afeta,
- Geta,
- Jimma Horo,
- Gudetu condole,
- Homosha,
- Jawi,

In 2015 precipitation estimated local coefficient was significant and positively to malaria cases in the following 16 wards in the country:

- Kurmuk,
- Bambasi,
- Guba,shebel berta,
- Debresina,
- Sayit,
- and Dawunt
- Homosha,
- Begi,
- Wegede,
- Enbise sar midir,
- Simada,
- Assosa,
- Maok omo,
- Enarj enawaga,
- Mehal sayint,
- Tach gayint

Precipitation was also significant and negatively related to malaria cases in the following 49 wards in the country (Figure 8c):

- Cheta,
- Melekoza,
- Zala,
- Esira,Yaso,
- Bulen,
- Bure,
- Banja,
- Dangila,
- Bahirdar Zuria,
- Jawi,
- Ebenat,
- Dembia,
- Wegera,
- Dabat,
- Debark,
- Addi arekay,
- and Bure.
- Decha,
- Geze gofa,
- Darmalo,
- Ibantu,
- Guangua,
- Wemberma,
- Fagta lakoma,
- Pawe,
- Yilana densa,
- Takusa,
- Metema,
- Gonder zuria,
- Lay armacho,
- Armarcho,
- Addi Arekay,
- Kefta humera,
- Ela,
- Ayida,
- Denibu gofa,
- Dibat,
- Ankasha,
- Madura,
- Mecha,
- Anchefer,
- Dera, Alfa,
- libo kemekem,
- Chilga,
- Belesa,
- Janamora,
- Tsegede,
- Welkait,
- Dera

In 2015 relative humidity was significant and positively related to malaria in the following 43 wards in the country.

- Metema,
- Mirab Armacho,
- Dembia,
- Wegera,
- Debark,
- Beyeda,
- Asgede Tsimbila,
- Tshtay koraro,
- Laelay Adiyabo,
- Gidami,
- Yama logi Wele,
- Gudetu Kondole,
- Mana Sibu,
- Daramalo,
- and Mirab Abaya.
- Chilga,
- Tach Armacho,
- Gonder Zuria,
- Dabat,
- Addi Arekay,
- Welkait,
- Kola Temben,
- Maychew,
- Mereb Leke,
- Jimma Horo,
- Gawo Kebe,
- Babo,
- Kiltu Kara,
- Dita,
- Lay Armacho,
- Tsegede,
- West Belesa,
- Janamora,
- Tselemt,
- Kafta Humera,
- Medebay,
- Adwa,
- Tahtay Adiyabo,
- Dale Wabera,
- Begi,
- Nejo,
- Agalometi,
- Chenech

Relative humidity was also significant and negatively related to malaria Cases in the following 32 wards (Figure 8e):

- Quara,
- Pawe,
- Sekela,
- Yilmana dense,
- Bahirdar Zuria,
- Amuru,
- Michakel,
- Michakel,
- Wembera,
- Hulet Ej Enese,
- Guangua,
- Guba,
- Dangila,
- Quarit,
- Mecha,
- Dera,
- Debre Elias,
- Senan,
- Dembecha,
- Daga damot,
- Quarit,
- Guagusa
- Dangura,
- Fafata lakoma,
- Gonje,
- Anchefer,
- Fogera,
- Guzamn,
- Debay Telatgen,
- Bure,
- Bibugn,
- Ankasha,
- And shiludad.

In year 2016 NDVI was significant and positive related to malaria cases in the following 50 wards in the country:

- Sahla,
- Alaje,
- Erebti,
- Tanqua Abergele,
- Degua Temben,
- Atsbi Wenberta,
- Werei leke,
- Ganta Afeshum,
- Dawunt,
- Ambasel,
- Chifra,
- Meket,
- Gidan,
- Ziquala,
- Hintalo Wejirat,
- Afdera,
- Abala,
- Kola Temben,
- Koneba,
- Hawazen,
- Dalul, Berahle,
- Wadla,
- Habru,
- Ewa,
- Bugna,
- Kobo,
- Saharti Samre,
- Megale,
- Beyeda,
- Enderta,
- Kelete Awelallo,
- Berahle,
- Sekota,
- Lay Gayint,
- Delanta,
- Worebabu,
- Guba Lafto,
- Lasta,
- Gulina,

- Awra,
- Dehana,
- Teru,
- SaesieTs aedaemba
- Aalmata,
- Ofla,
- Raya Azebo,
- and Erob.
- Gaz gibla,
- Yalo,
- Endamehoni,

NDVI also significant and negatively related to malaria cases in the Tahtay Adiyabo, Laelay Adiyabo and Mereb Leke wards (Figure 8b).

Precipitation estimated local coefficient showed only negative relationship with malaria cases per wards and was significant in the following 58 wards (Figure 8d):

- Ambasel,
- Meket,
- Worebabu,
- Guba lafto,
- Meket,
- Lasta,
- Gulina,
- Gaz Gibla,
- Wegera,
- Enamehoni,
- Sahla,
- Tsegede,
- Welkait,
- laelay Adiyabo,
- Adwa,
- Hawzen,
- Beyeda,
- Abala ,
- Berahle,
- and Berahle.
- Tach Gayint,
- Wadla,
- Chifa,
- Ewa,
- Ebenat,
- Gidan,
- Alamata,
- Dehana,
- Ziquala,
- Yalo,
- Janamora,
- Debark,
- keftay Humera,
- Mereb leke,
- Kola Temben,
- Ganta Afeshum,
- Tanqua Abergele,
- Erebti,
- Koneba,
- Farta, Lay Gayint,
- Delanta,
- Habru,
- Awra,
- Gugna,
- Kobo,
- Ofla,
- Belesa,
- Sekota,
- Teru,
- Debat,
- Addi arekay,
- Tahtay Adiyabo,
- Maychew,
- Wereri leke ,
- Tselmti,
- Enderta ,
- Afdera,
- kelete Awalalo

Relative humidity was significant and positively related to malaria cases in the Tahtay Adiyabo and, Laelay Adiyabo Wards Moreover; it was significant and negatively related to malaria cases in the following 52 wards (Figure 8f).

- Metema,
- Tsegede,
- Tanqua Abergele,
- Dabat,
- Beyeda,
- Endamehoni,
- Erebit,
- Habru,
- Guba lafto,
- Gulina,
- Alamata,
- Sekota,
- Degua Temben,
- Tach Armacho,
- Mirab Armacho,
- Enderta,
- Debark,
- Sahla,
- Hintalo Wejirat,
- Worebabu,
- Ewa,
- Gidan,
- Teru,
- Gaz gibila,
- Ofla,
- Kelete Awelallo,
- Lay armacho,
- Kaftay Humura,
- Wegera,
- Janamora,
- Ziquala,
- Abala,
- Chira,
- Awra,
- kobo,
- Yalo,
- Dehana,
- Tselemt,
- Atsbi Wenberta,

- Koneba,
- Hawazen,
- Ganta afsshum,

.

- Berahle,
- Werei leke,
- Erob,
- Gulomekede,
- Saesis Ts sedaemba,
- Daul,
- Naeder Adet,
- Dalul,
- Ahferom
- and Welkait

Temperature was significant and positively related to malaria cases in the in the following 26 wards in the country.

- Dawunt,
- Wadla,
- Laygayint,
- Gidan,Kobo,
- Gibla,
- Wegera,
- Sahla,
- Tselemti,
- Tahtay Koraro
- Ambasel,
- Guba Lafto,
- Bugna,
- Lasta,
- Dehana,
- Dabat,
- Debark,
- Asgede tsimbila,
- and Erop.
- Delanta,
- Maket,
- Ebenat,
- Gaz,
- Belesa,
- Janamora,
- Addi Arekay,
- Tahtay Adityabo,

Temperature also significant and negatively associated with malaria cases in the following 17 wards (Figure 8h).

- Mirab Armacho,
- Endameoni,
- Megale,
- Hintalo Wejirat,
- Tanqua Abergele,
- kelele Awelalo
- Tsegede,
- Alaje,
- Erebti,
- Saharti Samre,
- Degua Temben,
- ,andAtsbi Wenberta.
- Kafta Humera,
- Teru,
- Abala,
- Abergele,
- Kola Tamben,

Elevation was significant and positively related to malaria cases in the Delanta, Ambesal, Habru, Guba lafto, Chifa, Dubti, Ewa, Awara, Gulina, Kobo, Lasta, Gas Gibla, Almata, Ofla, Yalo, Teru, Afdera, Raya azebo, Sekota, Ziquala, Abergele, Alaje, Hintalo Wejrat, Megale, Erebti, Saharti Samre, Tanqua Abergele, Enderta, Degua Temben, Kola Temben, Werei leke, Hawzen, Atsbi Wenberta, Saesie Ts aedaemba, koneba, Berahle, and Dalul Wards. Moreover, Elevation is significant and negatively related with malaria cases in the Kafta Humera, Tsegede, Welkait, Asgede Tsimbila, Tselemti, Tahtay koraro, Tahtay Adiyabo, laelay Adiyabo, East Belesa, Ebenat, and Lay Gayint Wards as it depicted in Figure 8j.

The observed malaria cases map in 2015 (Figure 11a) should be compared with caution with the estimated map (Figure 11b), as well as the observed cases in 2016 (Figure 11c) with the estimated cases in 2016 (Figure 11d). According to the above discussion, the models' coefficients are not relevant in a large number of wards, thus the predictive power of the models is low in most of the country. However, it is important to point out that the models were not developed for prediction purposes. The usefulness and aim of the models is to identify relevant varying relationships between malaria cases and environmental variables.

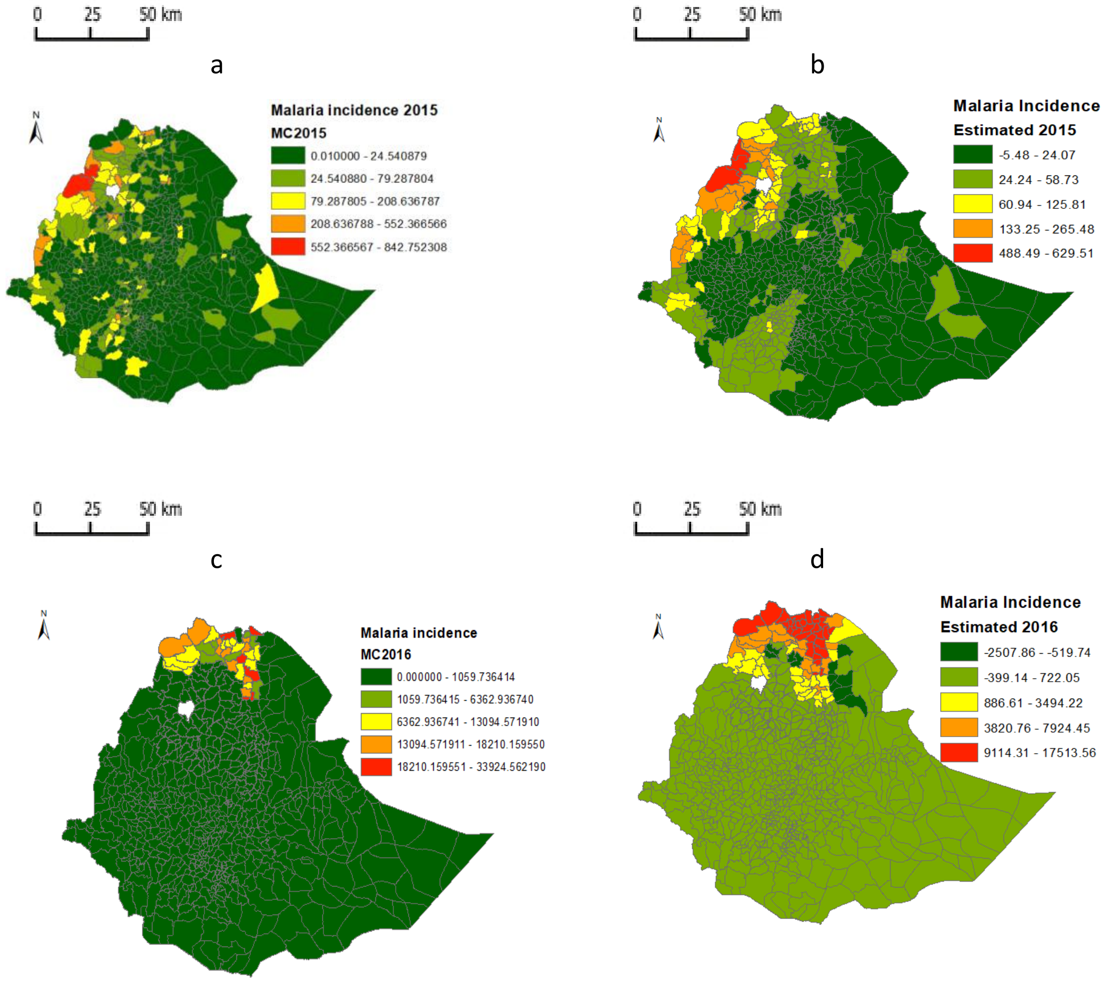


Figure 11. Observed (a) and GWR estimated (b) malaria incidence in 2015; observed (c) and GWR estimated (d) malaria incidence in 2016

Table 8 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 2016 GWR model, but the Global Moran's I statistic was closer to zero in GWR than in the OLS models.

Table 8 Comparison of goodness-of-fit results and residual analysis of the GWR and OLS models

Year	AICc		Adjusted R^2		RSS		Global Moran's I of residual (significance-score)	
	GWR	OLS	GWR	OLS	GWR	OLS	GWR	OLS
2015	7311.88	7683.95	0.515	0.012	715.43	3199364.47	0.007 (p=0.143, Z=1.463)	0.395 (p=0.000, Z=10.662)
2016	12349.73	12772.77	0.608	0.182	2250710287	5753288780	-0.104 (p=0.020, Z=-2.319)	0.316 (p=0.000, Z=23.683)

Global Moran's I results showed (Table 8) there is significant autocorrelation in the residuals of the GWR model in 2016, and authenticates the variables we considered in this study were unable to appropriately predict the malaria cases distribution spatially 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 results of spatial autocorrelation of residuals of the 2015 model was not statistically significant so that the model was well specified.

4.5 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 9), 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 9 Comparison of GWR and s-GWR models performances based on AICc

Year	Model	AICc	Improvement
2015	GWR model (before L -> G selection)	7311.884	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	

Table 10 reviews the result of GWR 4.0 (Nakaya *et al.*, 2005;Nakaya, 2016), where contrast of OLS, GWR, and s-GWR in terms of AICc, R^2, and adjusted R^2. The OLS model explained only 1.2% in 2015 and 18.2% in 2016 of the variability of malaria cases, 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 to12305 in 2016. In summary, both s-GWR models performed better than the other competing models, thus they are considered the final models for malaria cases in this study.

Table 10 Comparison of OLS, GWR and s-GWR models performances based on goodness-of-fit measures

Year	Fitness measures	OLS (global model)	GWR (local model)	s-GWR (mixed model)
2015	AICc	7683.95	7311.88	7273.69
	R^2	0.021	0.630	0.642
	Adjusted R^2	0.012	0.515	0.538
2016	AICc	12772.77	12349.73	12304.72
	R^2	0.189	0.683	0.685
	Adjusted R^2	0.182	0.609	0.624

The outcome of geographic variability test and local to global variable selection approach were based on DIFF of Criterion (Table 11 and Table 12) 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.

Table 11 Summary of s-GWR model coefficients in 2015

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	-132.546712
Elevation	-254.355	-41.107	-14.286	-4.255	209.723	-26.453838
TM2015	-171.932	-12.620	1.321	16.662	505.169	-16.845405
RH2015	-436.018	-13.593	6.545	18.316	745.869	-14.245095
PR2015	-286.025	-15.893	-1.410	8.455	464.719	-5.695182

R^2 = 0.642, Adjusted R^2 =0.578, AICc= 7273.690

Table 12 Summary of s-GWR model coefficients in 2016

Global coefficients						
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	-49.857796
RH2016	-11305.206	86.091	226.052	292.885	7926.942	-16.984620
NDVI2016	-1485.792	-119.889	-32.930	21.994	13194.320	-55.513555

R^2 = 0.685, Adjusted R^2 =0.624, AICc= 12304.719

The local estimated coefficients variation and associated t statistics are shown in (Figure 12, Figure 13, Figure 14, and Figure 15) below.

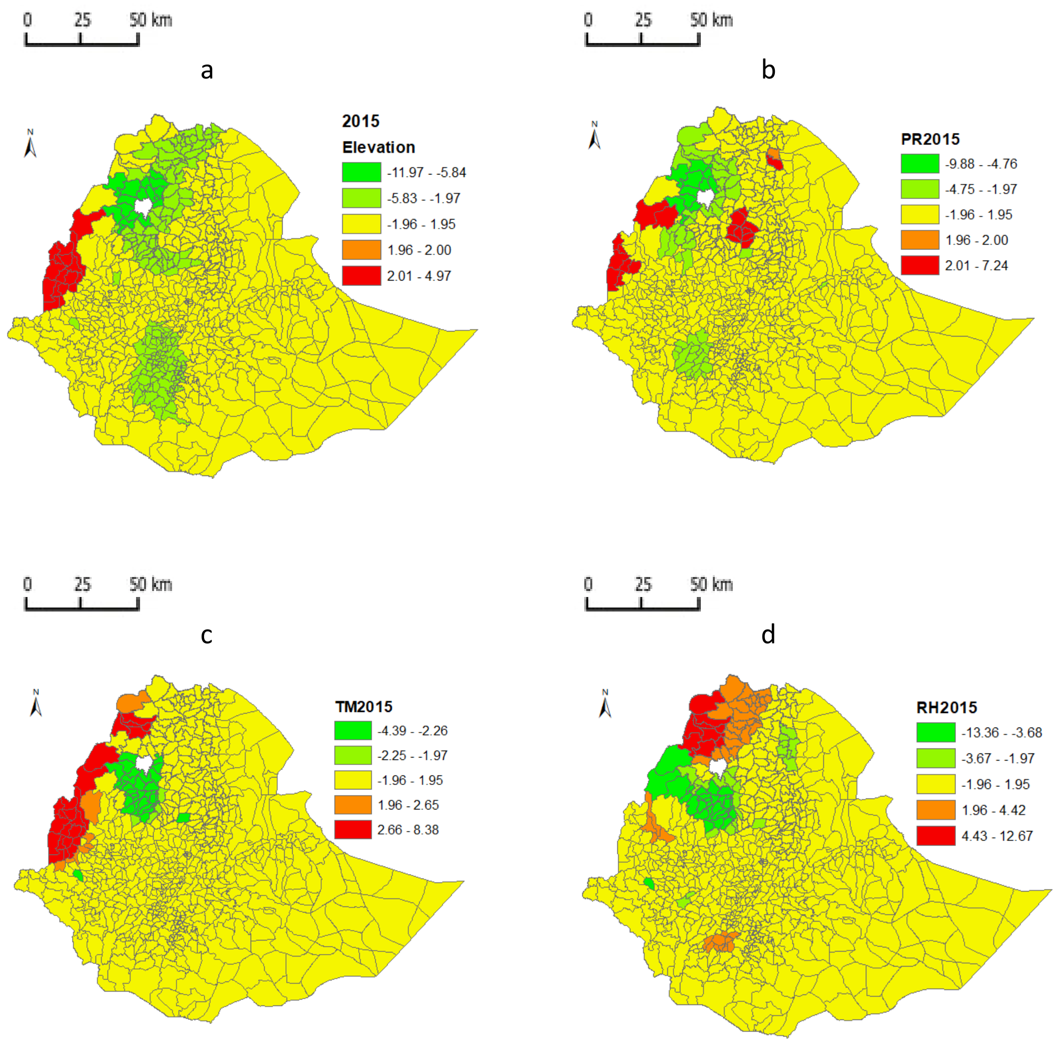
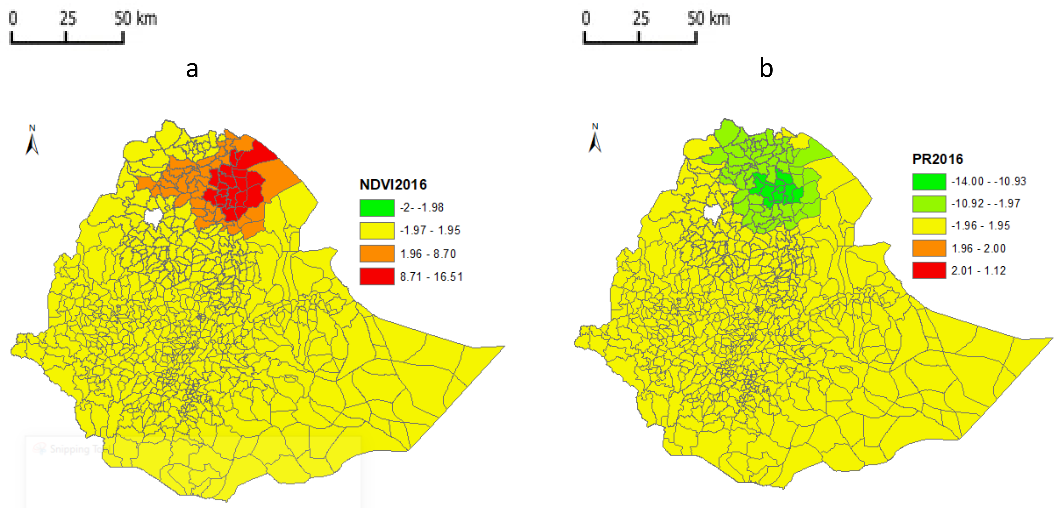


Figure 12.S-GWR Pseudo t-values for independent variables in 2015 with significance levels



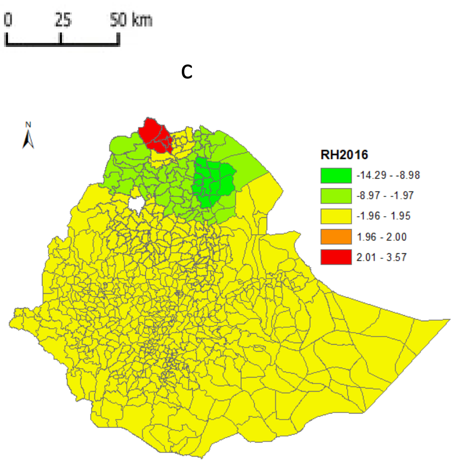


Figure 13.s-GWR Pseudo t-values for independent variables in 2016 with significance levels

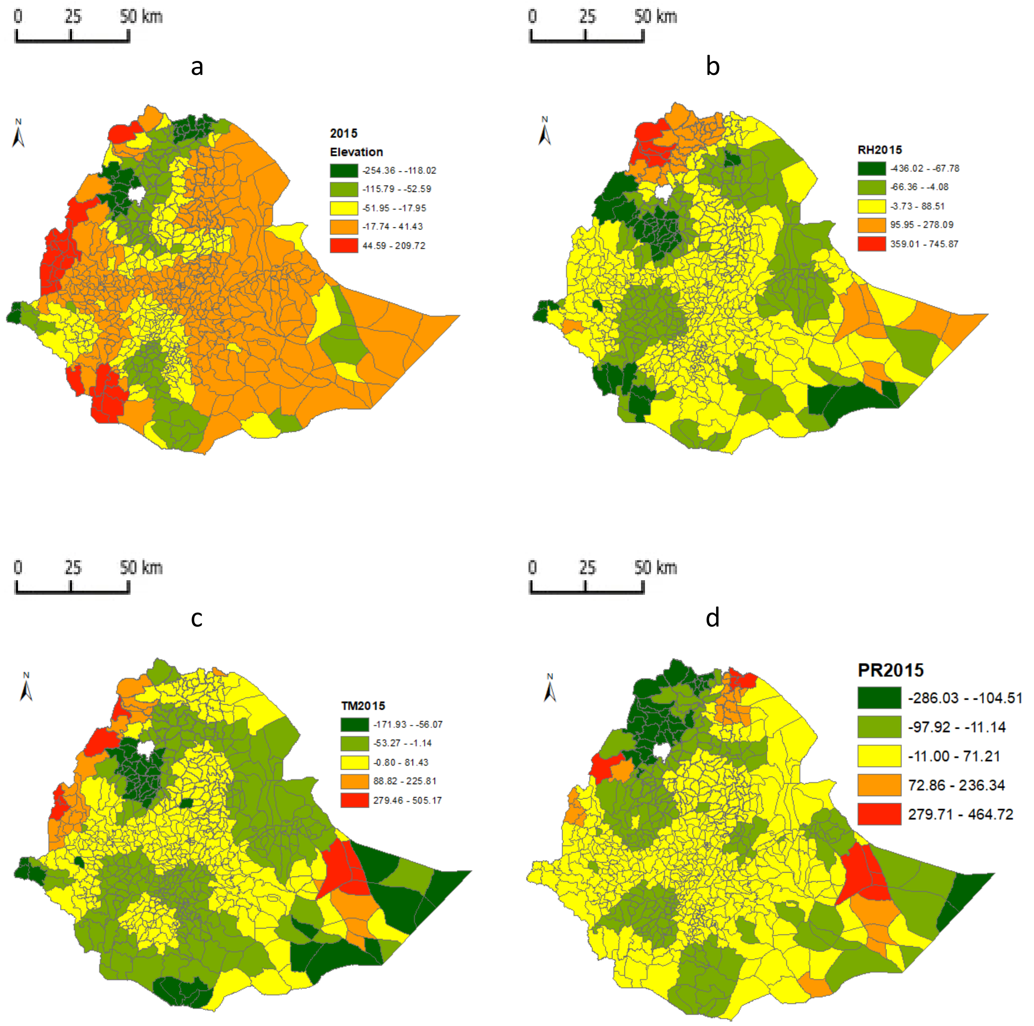


Figure 14. s- GWR local coefficients of the 2015 model (a-d)

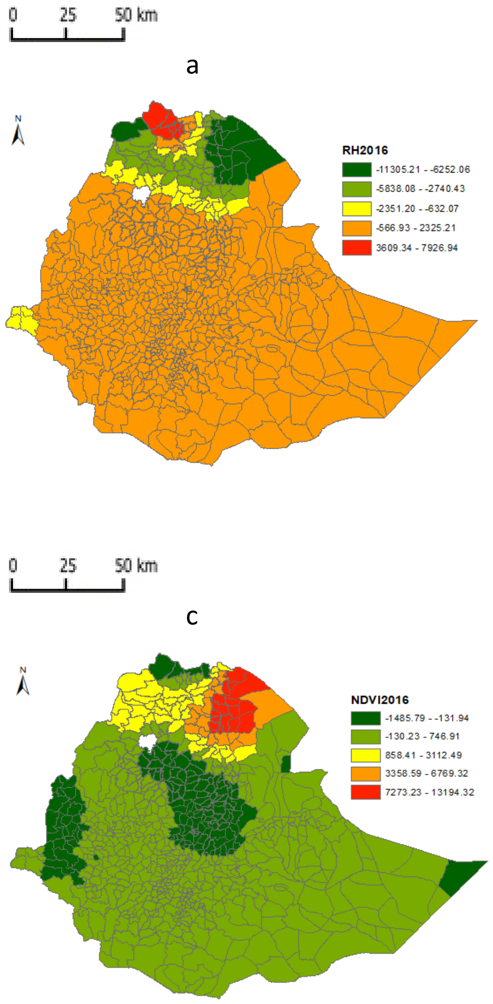


Figure 15. s- GWR local coefficients of the 2016 model (a-c)

In 2015 Precipitation was significant and positively associated with malaria cases in the following 19 wards:

- Kurmuk,
- Bambasi,
- Guba,
- Enarj enawaga,
- Mehal sayint,
- Tach gayaint,
- AndHintalo Wejrat
- Homosha,
- Begi,
- shebel berta,
- Debresina,
- Sayit,
- Dawunt,
- Assosa,
- Maok omo,
- Wegede,
- Enbise sar midir,
- Simada,
- Hindarta and

Precipitation also significant and negatively related to malaria cases in the following 50 wards in the country (Figure 12b):

- Cheta,
- Melekoza,
- Zala,
- Esira,
- Decha,
- Geze gofa,
- Darmalo,
- Yaso,
- Ela,
- Ayida,
- Denibu gofa,
- Ibantu,

- Dibat,
- Ankasha,
- Madura,
- Mecha,
- Anchefer,
- Dera,
- Takusa,
- Metema,
- Gonder zuria,
- Lay armacho,
- Armarcho,
- Addi Arekay,
- Kefta humera,
- Bulen,
- Bure,
- Banja,
- Dangila,
- Bahirdar Zuria,
- Alfa,
- libo kemekem,
- Chilga,
- Belesa,
- Janamora,
- Tsegede,
- Welkait,
- Dera and Bure.
- Guangua,
- Wemberma,
- Fagta lakoma,
- Pawe,
- Yilana densa,
- Jawi,
- Ebenat,
- Dembia,
- Wegera,
- Dabat,
- Debark,
- Addi arekay,
-

Elevation was significant and positively associated with malaria cases in the following 19 wards:

- Gidami,
- Gawo Kebe,
- Maok Omo,
- Mana Sibu,
- Menge,
- Sirba Abaya,
- Guba.
- Jimma Horo,
- Babo,
- Begi,
- Bambasi,
- Homosha,
- Kumuruk,
- Dale Wabera,
- Gudetu Kondole,
- Kiltu kara,
- Assosa,
- Biligidillu,
- Sherkole, and

Elevation also significant and negatively related to malaria cases in the following 155 wards (Figure 12a):

- Asgede Tsimbila,
- Kola Temben,
- Tahtay koraro,
- Erob, Gulomekeda,
- laelay Adiyabo,
- Beyeda, Tselemt,
- Fogera,
- Libo Kemkem,
- East Belesa,
- Dembia,
- Wegera,
- Danguara,
- Mecha,
- Dera,
- Jawi,Limu,
- Awabel,
- Dera,
- Shebel Bereta,
- Aneded,
- Guzamn,
- Dembecha,
- Debay,
- Medebay,
- Degua Temben,
- laelay Maychew,
- Ahferom,
- Debark,
- Welkait,
- Farta,
- Ebenat,
- Takusa,
- GonderZuria,
- Dabat,
- Pawe,
- yilmana Densa,
- Bahirdar Zuria,
- Ababo,
- Dejen,
- Wegde,
- Dejen,
- Baso liben,
- Debre Elias,
- Michakel,
- Telatgen,
- Naeder Adet,
- Hawzen,
- Ganta Afeshum,
- Mereb Leka,
- Addi Arekay,
- Alfa,
- Lay Gayint,
- west Belesa,
- Chilgam,
- Lay Armachewo,
- Metema,
- Dangila,
- West Esite,
- Debub Anchefer,
- Baso liben,
- Wara Jarso,
- Debresina,
- Awabel,
- Aneded,
- Bure,
- Senan,
- Enemay,

- EnarJ Enawa,
- Wemberma,
- Dega Damot,
- Sekela,
- Ameya,
- Tocha,
- Boloso Bombe,
- Omo Nada,
- Gibe,
- Selti, Ezha,
- limu kosa,
- Ameya,
- Amaro SP,
- Kercha,
- Wamena,
- kemba,
- Dita,
- Dila Zuria,
- Yirgachefe,
- Dara,
- Dale,
- Zala,
- Boreda,
- Boricha,
- Siraro,
- Arsi Negeli,
- loma Bosa,
- Mareka,
- Daniboya,
- Goro,
- Jabi Tehnan,
- Guangua,
- Sekela,
- Mandura,
- Nono,
- Mareka,
- Kacha Bira,
- Sekoru,
- Endiguagn,
- Cheha,
- Sekoru,
- Nono,
- Koochere,
- Bule Hora,
- Borke,
- zala,
- Chencha,
- Wenagol,
- Afele Kola,
- Hulla,
- Humbo,
- Denibu Gofa,
- Sodo Zuria,
- Boloso Sore,
- Awasa Zuria,
- limu,
- Ofa,
- Tibaro,
- Dedo,
- Bure,
- Bure Dembecha,
- Ankasha,
- Quarit,
- Fagta Lakoma,
- Goro,
- Gena Bosa,
- Tibaro,
- Yem,
- Admi Tulu,
- Chora,
- Tiro Afeta,
- chora,
- Gedeb,
- Hambela,
- Arba Minch Zuria,
- Daramalo,
- Abaya,
- Bule,
- Bore,
- Bursa,
- Loka Abaya,
- Kucha,
- Damot,
- Bomb,
- Shashemene zuria,
- Kacha biraa,
- Kindo Dida,
- Dune,
- Omo Nada,
- And Adwa.

Relative humidity was significant and positively related with malaria cases in the following 40 wards:

- Zala,
- Chencha,
- Kucha,
- Dembia,
- Wegera,
- Debark,
- Tselemt,
- Tach Armacho,
- Kafta Humera,
- Medebay Zana,
- Tahtay maychew,
- Mereb leke,
- Bilidigilu,
- and Agalmoeti.
- Daramalo,
- Mirab Abaya,
- Denibu Gofa,
- Gonder zuria,
- Dabat,
- Beyeda,
- Chilga,
- Tsegede,
- Welkait,
- Naeder Adet,
- Laelay Mayechew,
- Laelay Adiyabo,
- Sirba Abay,
- Dita,
- Boreda,
- Ofa, Humbo,
- West Belesa,
- Janamora,
- Addi Arekay,
- Metema,
- Mirab Armacho,
- Asgede Tsimbila,
- Tahtay koraro,
- Adwa,
- Tahtay Adiyabo,
- Kiltu kara,

Relative humidity also significant and negatively related to malaria cases in the following 47 wards (Figure 12d):

- Kobo,
- Raya Azebo,
- Hintalo Wejirat,
- Amuru,
- Guzamn,
- Senan,
- Wemberma,
- Jabi Tehnan,
- Ankasha,
- Banja,
- Gonje,
- Fagta lakoma,
- Pawe,
- Achefer,
- Yilmana densa,
- Guba,
- Alamata,
- Endamehoni,
- Dera,
- Baso Liben,
- Debre Elias,
- Dembecha,
- Debay Telatgen,
- Bibugn,
- Guangua,
- Sekela,
- Hulet EJ enense,
- Manduara,
- Jawi,
- Mecha,
- Dera ,
- and Fogera.
- Ofla,
- Alaje,
- Jarte jardega,
- Aneded,
- Michakel,
- Bure,
- Bure,
- Dega Damot,
- Guagusa shikudad,
- Quarit,
- Sekela,
- Dangura,
- Dangila,
- Bahirdar Zuria,
- West esite,
-

Moreover, Temperature was significant and positively associated with malaria cases in the following 28 wards:

- Quara,
- Tach Armacho,
- Mana Sibu,
- Assosa,
- Sirba Abay,
- Sherkole,
- Gidami,
- Ayira Guliso,
- Babo,
- and Kafta Humara.
- Mirab Armacho,
- Maok omo,
- Babo, Nejo,
- Bilidigilu,
- Wenbera,
- Guba,Anfilo,
- Jimma Horo,
- Boji Chekorsa,
- Begi,
- Tsegede,
- Bambasi,
- Kiltu Kara,
- Menge,
- Kurmuk,
- YamaLogi welel,
- Dale Wabera,
- Gawo kebe,
- Gudetu Kondole,

Temperature also significant and negatively related to malaria cases in the wards in the following 33 wards (Figure 12c):

- Dera,
- Senan,
- Bure,
- Dega damot,
- Banja,
- Gonje,
- Fagata Lakoma,
- Dangila,
- Bahirdar Zuria,
- Pawe,
- Fogera,
- Debre Elias,
- Michakel,
- Wemberma,
- Bibugn, Ankasha,
- Sekela,
- Hulet Ej enese,
- Dangila,
- Mecha,
- Dera,
- Jawi,
- Libo Kemkem
- Guzamn,
- Dembecha,
- Jabi Tehnan,
- Guagusa,
- Quarit,
- Sekela,
- Takusa,
- Yilmana densa,
- Anchefer,
- Alfa,
- and Bure.

In 2016 Relative humidity was significant and positively associated with malaria cases in the Medebay Zena, Asgede Tsimbila, Tahtay koraro, and Laelay Adiyabo Wards. Relative humidity also significant and negatively related to malaria cases in the following 53 wards (Figure 13c):

- Chilga,
- Tsegede,
- Wegera,
- Ebenat,
- Degua Temben,
- Koneba,
- Ganta Afeshum,
- Ambasel,
- Ewa, Awra,
- Wadla,
- Gidan,
- Yalo,
- Gaz Gibla,
- Ziquala,
- Raya Azebo,
- Alaje,
- Afdera,
- Enderta,
- Metema,
- Debark,
- West Belesa,
- East Belesa,
- Kelete Awelallo,
- Berahle, Dalul,
- Erob,
- Habru,
- Teru, Gulina,
- Meket,
- Kobo,
- Alamata,
- Bugna,
- Sekota,
- Yalo,
- Megale,
- Abala,

andTanqua Abergele.

- Armacho,
- Dabat,
- Gonder Zuria,
- Beyeda,
- Atsbi Wenberta,
- Saesieaedaemba,
- Worebabu,
- Chifra,
- Guba lafto,Delanta,
- Lasta,
- Gulina,
- Gidan,
- Dehana,
- Ofla,
- Endamehoni,
- Erebti,
- Hintalo Wejirat,

Precipitation was significant and only negatively related with malaria cases in the following 60 wards in the country (Figure 13b):

- Afdera,
- Wegera,
- Tach Gayint,
- Meket,
- Worebabu,
- Guba lafto,
- Meket,
- Lasta,
- Gulina,
- Gaz Gibla,
- Wegera,
- Enamehoni,
- Sahla,
- Tsegede,
- Welkait,
- Mereb leke,
- Kola Temben,
- Ganta Afeshum,
- Enderta,
- Koneba,
- Kafta Humera,
- Farta,
- Farta,
- Wadla,
- Chifa,
- Ewa,
- Ebenat,
- Gidan,
- Alamata,
- Dehana,
- Ziquala,
- Yalo,
- Janamora,
- Debark,
- Tahtay Adiyabo,
- Maychew,
- Wereri leke ,
- Tselmti, Beyeda,
- Erebti,
- kelete Awalalo
- Tseged,
- Ambasel,
- Lay Gayint,
- Delanta,
- Habru,
- Awra,
- Gugna,
- Kobo,
- Ofla,
- Belesa,
- Sekota,
- Teru,
- Debat,
- Addi arekay,
- laelay Adiyabo,
- Adwa,
- Hawzen,
- Tanqua Abergele,
- Berahle,
- and Berahle.

Finally, NDVI was significant and only positively related with malaria cases in the following 52 wards in the country as depicted in Figure 13a):

- Tach Armacho,
- East Belesa,
- Lasta,
- Gaz Gibla,
- Dabat,
- Addi Arekay,
- Abergele,
- Ewa,
- Afdera,
- Yalo,
- Olfa,
- Raya Azebo,
- Megale,
- Saharti Samre,
- Kola Temben,
- Atsbi wenberta,
- Hawzen,
- Saesies aedaemba,
- Lay Armacho,
- Dehana,
- Meket,
- Ziquala,
- Janamora,
- Beyeda,
- Guba lafto,
- Chifra,
- Kobo,
- Teru,
- Sekota,
- Alaje,
- Erebti,
- Enderta,
- Degua Temben,
- Koneba,
- Dalul,
- And Erob.
- Wegera,
- Bugna,
- Gidan,
- Sahla,
- Debark,
- Tanqua Abergele,
- Habru,
- Awra,
- Gulina,
- Alamata,
- Endamehoni,
- Hintalo Weejirat,
- Abala,
- Tselemt,
- Kelete Awelallo,
- Berahle,
- Ganta Afeshum,
-

5 DISCUSSION

In this study, the effects environmental variables on malaria cases 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 cases 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 outcome of this research depicted that malaria incidence in Ethiopia heterogeneously distributed and spatially clustered at the ward level in the country during the study period. The finding of theis research are consistent with research results 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 cases associated with environmental risk factors in the country. The result of this research could be helpful for ward-level planning, policy making, and implementation of malaria control.

Our research demonstrated the relevance of the semiparametric geographical modeling approach of local-level risk factors analysis by contrasting global (OLS), local (GWR), and mixed (s-GWR) model. Our study exhibited the drawbacks of the OLS method to explain the variation of malaria cases in terms of estimation of model performance, model correctness and complexities evaluated to the GWR model. We showed that model goodness-of-fit could be enhanced through the execution of the s-GWR model. The finding of this research work 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 suggested to evade the model complexity as an alternative 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 final s-GWR model of the Moran’s I of the residuals was -0.079349 (z score = -3.622420 and p-value = 0.053), so the spatial autocorrelation in the residual are not statistical significant thus the model was properly well 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). The variation in local R^2 over the wards revealed significant location differences in the malaria incidence transmission process in the study area (Figure 16a-b).The local R^2 depicted that the s-GWR model had higher performance in malaria cluster areas when it compared to the other parts of the study area identical with earlier similar studies from Nepal (Acharya et al. 2018), Colombia(Delmelle et al. 2016) and South Africa (Manyangadze et al. 2016).

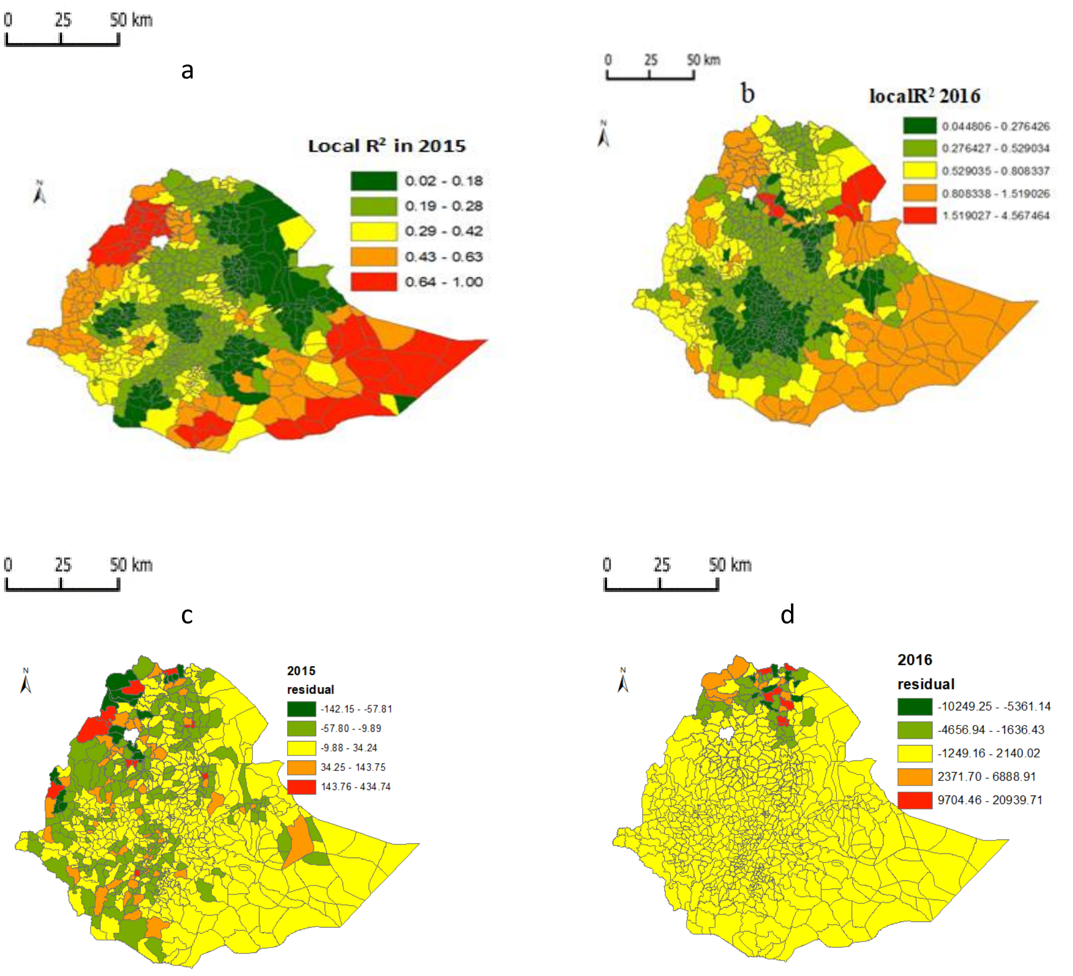


Figure 16.Local R2 (a, b), residual distribution (c, d) in the s-GWR-based prediction model of the malaria case in 2015 and 2016

Our final mixed s-GWR models show that the distribution of annual malaria cases is heterogeneous (Figure 12 and Figure 13) 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 cases that vary strongly at the village level. Identifying the malaria hot-spot wards (wards with a high number of malaria incidences) is important in implementing 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 hot-spot may point to the wards that require prompt notice in terms of planning and execution of the disease control strategies.

The heterogeneity of malaria incidence determined by ecological, biological, and sociological factors (Pinchoff et al. 2015). As distinguished by researcher (Ehlkes et al. 2014), nearly all research presume homogeneous effect 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 research, the analysis result demonstrated assuming some variables vary at the local level, while others have a global effect, significantly make better the model performance. Permitting spatial non-stationary in the regression model lets clear interpretation regarding the true nature of the possible correlation (Ehlkes et al. 2014). That could be due to the Long-lasting insecticidal nets (LLINs) distributed to some of the wards that have malaria cases 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 evaluating the relationship between environmental risk factors and malaria cases, one should think about the pathways in which these variables under research lie (Ehlkes et al. 2014). For instance, the environmental variables: temperature, NDVI, elevation, relative humidity, and precipitation, which influence the malaria cases considered in this research as they determine by the plenty of mosquitos.

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 vector control factors and malaria cases may bring out unpredicted results.

In this research, there was an association between elevation and malaria cases. 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 (Hanandita 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 transference possible decreases as the elevation raises (Chikodzi, 2013; Meyrowitsch *et al.*, 2011). In our study, also we noted elevation was significant in 2015 and depicted its expected negative correlation with malaria cases in some of the wards in the northern and southern wards, but also depicted a positive correlation in some wards to the western part of the country (Figure 12a).

In Ethiopia, precipitation was significantly correlated with malaria cases in tropical areas (Midekisa et al. 2015). Moreover, in Botswana, precipitation showed association with the incidence of clinical malaria cases (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 cases 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 cases 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 Figure 12b. In Ethiopia, minimum temperatures significantly correlated with malaria cases in cold areas (Midekisa et al. 2015). In this study also local coefficients of temperature in 2015 showed positive and negative relationship with malaria cases and were significant in some wards located in the northwestern and western part of the country, as depicted in Figure 12c.

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 transference of malaria by given that water to create aquatic habitats. In this study also local coefficients of relative humidity in 2015 depicted the expected positive and negative relationship with malaria cases, and they were significant in some wards located in the northwestern and western part of the country, as depicted in Figure 12d. 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 cases in some wards in the country. NDVI was significant in some wards located in the northern part of the country, as it showed in Figure 13a. In 2016 precipitation local coefficients showed an only negative relationship with malaria cases and were significant in some wards located in the northern part of the country, as it showed in Figure 13b. In 2016 relative humidity local coefficients showed a positive and negative relationship with malaria cases in some wards in the country. Relative humidity was significant in some wards located in the northern part of the country (Figure 13c). 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. Researchers (Gwitira et al. 2015) distinguished that in malaria incidence where there is effective malaria control wards, there would be weak association among enviromentl risk factors and malaria cases. This was observed in this study in 2015, NDVI was a weak association with malaria cases in the country. In 2016 elevation and temperature were also weak correlations with malaria cases 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 cases 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 cases 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 cases 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 cases in some wards in Ethiopia. However, s-GWR model allowing for spatial heterogeneity explains better the relationship of malaria cases 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).

6 CONCLUSION

This research analyzed the modeling of malaria cases and its association with environmental risk factors in Ethiopia. The finding of this research showed that malaria cases 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 relevant risk factors accountable for the spatial variation of the malaria 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 cases that varies over space. In our study, the variability of malaria cases 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 result of this research has a direct suggestion for health policy planning and decision making. Moreover, this research shows the relevance of a mixed geographical regression modeling approach in geostatistics analysis of malaria cases 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 cases study in Ethiopia to map and explore environmental risk factors in the entire country at the ward-level. The methodological framework implemented in this research is convertible in other county depending up on spatial data availability. Moreover, this study demonstrates the importance of a mixed s-GWR modeling approach in the spatial analysis of malaria cases affected by complex environmental risk factors at the ward-level.

This research also revealed the relevance of s-GWR approach to make better the knowledge about malaria cases and its determinants, so that this research can be used for the malaria control 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 cases.

REFERENCES

Abtew, Wossenu, Assefa M. Melesse, and Tibebe Dessalegne. 2009. “Spatial, Inter and Intra-Annual Variability of the Upper Blue Nile Basin Rainfall.” *Hydrological Processes*.

Acharya, Bipin Kumar, Chun Xiang Cao, Tobia Lakes, Wei Chen, Shahid Naeem, and Shreejana Pandit. 2018. “Modeling the Spatially Varying Risk Factors of Dengue Fever in Jhapa District, Nepal, Using the Semi-Parametric Geographically Weighted Regression Model.” *International Journal of Biometeorology*.

Alemu, Abebe, Gemeda Abebe, Wondewossen Tsegaye, and Lemu Golassa. 2011. “Climatic Variables and Malaria Transmission Dynamics in Jimma Town, South West Ethiopia.” *Parasites and Vectors* 4(1):1–11.

Alimi, Temitope O., Douglas O. Fuller, Whitney A. Qualls, Socrates V. Herrera, Myriam Arevalo-Herrera, Martha L. Quinones, Marcus V. G. Lacerda, and John C. Beier. 2015. “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 and Vectors*.

Ankamah, Sylvia, Kaku S. Nokoe, and Wahab A. Iddrisu. 2018. “Modeling Trends of Climatic Variability and Malaria in Ghana Using Vector Autoregression.” *Malaria Research and Treatment*.

Anselin, Luc. 1995. “Local Indicators of Spatial Association—LISA.” *Geographical Analysis*.

Berke, Olaf. 2004. “Exploratory Disease Mapping: Kriging the Spatial Risk Function from Regional Count Data.” *International Journal of Health Geographics*.

Brunsdon, Chris, A. Stewart Fotheringham, and Martin E. Charlton. 2010. “Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity.” *Geographical Analysis*.

Buck, Kyle D. 2016. “Modeling of Geographic Cancer Risk Factor Disparities in US Counties.” *Applied Geography*.

De Castro, Marcia Caldas, Roberto L. Monte-Mór, Diana O. Sawyer, and Burton H. Singer. 2006. “Malaria Risk on the Amazon Frontier.” *Proceedings of the National Academy of Sciences of the United States of America*.

Chasco, Coro, Isabel García, and José Vicéns. 2007. “Modeling Spatial Variations in

Household Disposable Income with Geographically Weighted Regression." *Munich Personal RePEc Archive*.

Chi, Sang Hyun, Diana S. Grigsby-Toussaint, Natalie Bradford, and Jinmu Choi. 2013. "Can Geographically Weighted Regression Improve Our Contextual Understanding of Obesity in the US? Findings from the USDA Food Atlas." *Applied Geography*.

Chikodzi, David. 2013. "Spatial Modeling of Malaria Risk Zones Using Environmental, Anthropogenic Variables and GeograPhical Information Systems Techniques." *Journal of Geosciences and Geomatics*.

Chirebvu, Elijah, Moses John Chimbari, Barbara Ntombi Ngwenya, and Benn Sartorius. 2016. "Clinical Malaria Transmission Trends and Its Association with Climatic Variables in Tubu Village, Botswana: A Retrospective Analysis." *PloS One*.

Craig, Marlies, David Le Sueur, and Bob Snow. 1999. "A Climate-Based Distribution Model of Malaria Transmission in Sub-Saharan Africa." *Parasitology Today*.

Dambach, Peter, Vanessa Machault, Jean Pierre Lacaux, Cécile Vignolles, Ali Sié, and Rainer Sauerborn. 2012. "Utilization of Combined Remote Sensing Techniques to Detect Environmental Variables Influencing Malaria Vector Densities in Rural West Africa." *International Journal of Health Geographics*.

Day, Jonathan F. 2016. "Mosquito Oviposition Behavior and Vector Control." *Insects*.

Delmelle, Eric, Michael Hagenlocher, Stefan Kienberger, and Irene Casas. 2016. "A Spatial Model of Socioeconomic and Environmental Determinants of Dengue Fever in Cali, Colombia." *Acta Tropica*.

Dereje Ayalew. 2012. "Variability of Rainfall and Its Current Trend in Amhara Region, Ethiopia." *AFRICAN JOURNAL OF AGRICULTURAL RESEEARCH*.

Dysoley, Lek, Akira Kaneko, Hideaki Eto, Toshihiro Mita, Doung Socheat, Anders Börkman, and Takatoshi Kobayakawa. 2008. "Changing Patterns of Forest Malaria among the Mobile Adult Male Population in Chumkiri District, Cambodia." *Acta Tropica*.

Ehlkes, Lutz, Anne C. Krefis, Benno Kreuels, Ralf Krumkamp, Ohene Adjei, Matilda Ayim-Akonor, Robin Kobbe, Andreas Hahn, Christof Vinnemeier, Wibke Loag, Udo Schickhoff, and Jürgen May. 2014. "Geographically

Weighted Regression of Land Cover Determinants of Plasmodium Falciparum Transmission in the Ashanti Region of Ghana.” *International Journal of Health Geographics*.

Elyazar, Iqbal R. F., Marianne E. Sinka, Peter W. Gething, Siti N. Tarmidzi, Asik Surya, Rita Kusriastuti, Winarno, J. Kevin Baird, Simon I. Hay, and Michael J. Bangs. 2013. “The Distribution and Bionomics of Anopheles Malaria Vector Mosquitoes in Indonesia.” in *Advances in Parasitology*.

Ethiopian, Projecting, Demographics From, Spectrum Suite, and O. F. Models. 2014. “Projecting Ethiopian Demographics from 2012–2050 Using the Spectrum Suite of Models.” (July).

Fazzini, Massimiliano, Carlo Bisci, and Paolo Billi. 2015. “The Climate of Ethiopia Massimiliano.” *Choice Reviews Online*.

Fotheringham, A. S., M. E. Charlton, and C. Brunsdon. 1998. “Geographically Weighted Regression: A Natural Evolution of the Expansion Method for Spatial Data Analysis.” *Environment and Planning A*.

Fotheringham, Stuart, Martin Charlton, and Chris Brunsdon. 2002. “Chp 2: Geographically Weighted Regression : The Basics.” in *Geographically Weighted Regression: The analysis of spatially varying relationships*.

Gao, Jiangbo and Shuangcheng Li. 2011. “Detecting Spatially Non-Stationary and Scale-Dependent Relationships between Urban Landscape Fragmentation and Related Factors Using Geographically Weighted Regression.” *Applied Geography*.

Ge, Yong, Yongze Song, Jinfeng Wang, Wei Liu, Zhoupeng Ren, Junhuan Peng, and Binbin Lu. 2017. “Geographically Weighted Regression-Based Determinants of Malaria Incidences in Northern China.” *Transactions in GIS*.

Goodchild, Michael F. and Donald G. Janelle. 2004. *Spatially Integrated Social Science Thinking Spatially in the Social Sciences*.

Guthmann, Jean Paul, A. Llanos-Cuentas, A. Palacios, and A. J. Hall. 2002. “Environmental Factors as Determinants of Malaria Risk. A Descriptive Study on the Northern Coast of Peru.” *Tropical Medicine and International Health*.

Gwitira, Isaiah, Amon Murwira, Fadzai M. Zengeya, Mhosisi Masocha, and Susan Mutambu. 2015. “Modelled Habitat Suitability of a Malaria Causing Vector (Anopheles Arabiensis) Relates Well with Human Malaria Incidences in Zimbabwe.” *Applied Geography*.

Hagose, Lemlem. 2017. “Strategic Analysis of Sesame (Sesamum Indicum L.) Market Chain in Ethiopia a Case of Humera District.” *International Journal of Plant & Soil Science*.

Halimi, Mansour, Manuchehr Farajzadeh, Mahdi Delavari, Ashraf Takhtardeshir, and Abbas Moradi. 2014. “Modeling Spatial Relationship between Climatic Conditions and Annual Parasite Incidence of Malaria in Southern Part of Sistan&Balouchistan Province of Iran Using Spatial Statistic Models.” *Asian Pacific Journal of Tropical Disease*.

Hanandita, Wulung and Gindo Tampubolon. 2016. “Geography and Social Distribution of Malaria in Indonesian Papua: A Cross-Sectional Study.” *International Journal of Health Geographics*.

Hasyim, Hamzah, Afi Nursafingi, Ubydul Haque, Doreen Montag, David A. Groneberg, Meghnath Dhimal, Ulrich Kuch, and Ruth Müller. 2018. “Spatial Modeling of Malaria Cases Associated with Environmental Factors in South Sumatra, Indonesia.” *Malaria Journal* 17(1):1–15.

Hoshen, Moshe B. and Andrew P. Morse. 2004. “A Weather-Driven Model of Malaria Transmission.” *Malaria Journal*.

Jephcote, Calvin and Haibo Chen. 2012. “Environmental Injustices of Children’s Exposure to Air Pollution from Road-Transport within the Model British Multicultural City of Leicester: 2000-09.” *Science of the Total Environment*.

de Jonge, Edwin and Mark van der Loo. 2013. “An Introduction to Data Cleaning with R.” *Statistics Netherlands*.

Jury, Mark R. and Chris Funk. 2013. “Climatic Trends over Ethiopia: Regional Signals and Drivers.” *International Journal of Climatology*.

Kleinschmidt, I., B. L. Sharp, G. P. Y. Clarke, B. Curtis, and C. Fraser. 2001. “Use of Generalized Linear Mixed Models in the Spatial Analysis of Small-Area Malaria Incidence Rates in KwaZulu Natal, South Africa.” *American Journal of Epidemiology*.

Kuo, Chi Chien, Nicola Wardrop, Chung Te Chang, Hsi Chieh Wang, and Peter M. Atkinson. 2017. “Significance of Major International Seaports in the Distribution of Murine Typhus in Taiwan.” *PLoS Neglected Tropical Diseases*.

Lin, Chia Hsien and Tzai Hung Wen. 2011. “Using Geographically Weighted Regression (GWR) to Explorespatial Varying Relationships of Immature Mosquitoes Andhuman Densities with the Incidence of Dengue.” *International*

Journal of Environmental Research and Public Health.

Loha, Eskindir, Wakgari Deressa, Taye Gari, Meshesha Balkew, Oljira Kenea, Tarekegn Solomon, Alemayehu Hailu, Bjarne Robberstad, Meselech Assegid, Hans J. Overgaard, and Bernt Lindtjørn. 2019. "Long-Lasting Insecticidal Nets and Indoor Residual Spraying May Not Be Sufficient to Eliminate Malaria in a Low Malaria Incidence Area: Results from a Cluster Randomized Controlled Trial in Ethiopia." *Malaria Journal*.

Loha, Eskindir and Bernt Lindtjørn. 2010. "Model Variations in Predicting Incidence of Plasmodium Falciparum Malaria Using 1998-2007 Morbidity and Meteorological Data from South Ethiopia." *Malaria Journal*.

Lu, Binbin, Paul Harris, Martin Charlton, and Chris Brunsdon. 2015. "Calibrating a Geographically Weighted Regression Model with Parameter-Specific Distance Metrics." *Procedia Environmental Sciences*.

Manyangadze, Tawanda, Moses J. Chimbari, Margaret Macherera, and Samson Mukaratirwa. 2017. "Micro-Spatial Distribution of Malaria Cases and Control Strategies at Ward Level in Gwanda District, Matabeleland South, Zimbabwe." *Malaria Journal*.

Manyangadze, Tawanda, Moses John Chimbari, Michael Gebreslasie, and Samson Mukaratirwa. 2016. "Risk Factors and Micro-Geographical Heterogeneity of Schistosoma Haematobium in Ndumo Area, UMkhanyakude District, KwaZulu-Natal, South Africa." *Acta Tropica*.

Mar'ah, Z., A. Djuraidah, and A. H. Wigena. 2017. "Semi-Parametric Geographically Weighted Regression Modeling Using Linear Model of Coregionalization." *Researchgate.Net* (July).

Masaninga, Freddie, Nawa Mukumbuta, Ketty Ndhlovu, Busiku Hamainza, Pauline Wamulume, Emmanuel Chanda, John Banda, Mercy Mwanza-Ingwe, John M. Miller, Birkinesh Ameneshewa, Abraham Mnzava, and Elizabeth Kawesha-Chizema. 2018. "Insecticide-Treated Nets Mass Distribution Campaign: Benefits and Lessons in Zambia." *Malaria Journal*.

Matthews, Stephen A. and Tse-Chuan Yang. 2012. "Mapping the Results of Local Statistics." *Demographic Research*.

McKinley, Jennifer M., Ulrich Ofterdinger, Michael Young, Amy Barsby, and Anna Gavin. 2013. "Investigating Local Relationships between Trace Elements in Soils and Cancer Data." *Spatial Statistics*.

Mei, Chang Lin, Ning Wang, and Wen Xiu Zhang. 2006. “Testing the Importance of the Explanatory Variables in a Mixed Geographically Weighted Regression Model.” *Environment and Planning A*.

Mengistu, Daniel, Woldeamlak Bewket, and Rattan Lal. 2014. “Recent Spatiotemporal Temperature and Rainfall Variability and Trends over the Upper Blue Nile River Basin, Ethiopia.” *International Journal of Climatology*.

Meyrowitsch, Dan W., Erling M. Pedersen, Michael Alifrangis, Thomas H. Scheike, Mwelecele N. Malecela, Stephen M. Magesa, Yahya A. Derua, Rwehumbiza T. Rwegoshora, Edwin Michael, and Paul E. Simonsen. 2011. “Is the Current Decline in Malaria Burden in Sub-Saharan Africa Due to a Decrease in Vector Population?” *Malaria Journal*.

Midekisa, Alemayehu, Belay Beyene, Abere Mihretie, Estifanos Bayabil, and Michael C. Wimberly. 2015. “Seasonal Associations of Climatic Drivers and Malaria in the Highlands of Ethiopia.” *Parasites and Vectors*.

Mordecai, Erin A., Krijn P. Paaijmans, Leah R. Johnson, Christian Balzer, Tal Ben-Horin, Emily de Moor, Amy Mcnally, Samraat Pawar, Sadie J. Ryan, Thomas C. Smith, and Kevin D. Lafferty. 2013. “Optimal Temperature for Malaria Transmission Is Dramatically Lower than Previously Predicted.” *Ecology Letters*.

Mushinzimana, Emmanuel, Stephen Munga, Noboru Minakawa, Li Li, Chen Chieh Feng, Ling Bian, Uriel Kitron, Cindy Schmidt, Louisa Beck, Guofa Zhou, Andrew K. Githeko, and Guiyun Yan. 2006. “Landscape Determinants and Remote Sensing of Anopheline Mosquito Larval Habitats in the Western Kenya Highlands.” *Malaria Journal*.

Nakaya, Tomoki. 2016. “GWR4.09 User Manual.” *GWR 4 Development Team* 40.

Nakaya, Tomoki, A. S. Fotheringham, C. Brunsdon, and M. Charlton. 2005. “Geographically Weighted Poisson Regression for Disease Association Mapping.” *Statistics in Medicine*.

Osman, Mahdi and Petra Sauerborn. 2002. “A Preliminary Assessment of Characteristics and Long Term Variability of Rainfall in Ethiopia–Basis for Sustainable Land Use and Resource Management.” *Challenges to Organic Farming and Sustainable Land Use in the Tropics and Subtropics, Deutscher Tropentag*.

Páez, Antonio, Takashi Uchida, and Kazuaki Miyamoto. 2002. “A General

Framework for Estimation and Inference of Geographically Weighted Regression Models: 1. Location-Specific Kernel Bandwidths and a Test for Locational Heterogeneity.” *Environment and Planning A*.

Parker, Daniel M., Stephen A. Matthews, Guiyun Yan, Guofa Zhou, Ming Chieh Lee, Jeeraphat Sirichaisinthop, Kirakorn Kiattibutr, Qi Fan, Peipei Li, Jetsumon Sattabongkot, and Liwang Cui. 2015. “Microgeography and Molecular Epidemiology of Malaria at the Thailand-Myanmar Border in the Malaria Pre-Elimination Phase.” *Malaria Journal*.

Pasculli, Antonio, Sergio Palermi, Annalina Sarra, Tommaso Piacentini, and Enrico Miccadei. 2014. “A Modeling Methodology for the Analysis of Radon Potential Based on Environmental Geology and Geographically Weighted Regression.” *Environmental Modeling and Software*.

Pinault, Lauren L. and Fiona F. Hunter. 2011. “New Highland Distribution Records of Multiple Anopheles Species in the Ecuadorian Andes.” *Malaria Journal*.

Pinchoff, Jessie, German Henostroza, Bryan S. Carter, Sarah T. Roberts, Sisa Hatwiinda, Busiku Hamainza, Moonga Hawela, and Frank C. Curriero. 2015. “Spatial Patterns of Incident Malaria Cases and Their Household Contacts in a Single Clinic Catchment Area of Chongwe District, Zambia.” *Malaria Journal*.

Ramezankhani, Roghieh, Arezoo Hosseini, Nooshin Sajjadi, Mostafa Khoshabi, and Azra Ramezankhani. 2017. “Environmental Risk Factors for the Incidence of Cutaneous Leishmaniasis in an Endemic Area of Iran: A GIS-Based Approach.” *Spatial and Spatio-Temporal Epidemiology*.

Rulisa, Stephen, Fredrick Kateera, Jean Pierre Bizimana, Steven Agaba, Javier Dukuzumuremyi, Lisette Baas, Jean de Dieu Harelimana, Petra F. Mens, Kimberly R. Boer, and Peter J. de Vries. 2013. “Malaria Prevalence, Spatial Clustering and Risk Factors in a Low Endemic Area of Eastern Rwanda: A Cross Sectional Study.” *PLoS ONE*.

Shapiro, Lillian L. M., Shelley A. Whitehead, and Matthew B. Thomas. 2017. “Quantifying the Effects of Temperature on Mosquito and Parasite Traits That Determine the Transmission Potential of Human Malaria.” *PLoS Biology*.

Siraj, A. S., M. Santos-Vega, M. J. Bouma, D. Yadeta, D. Ruiz Carrascal, and M. Pascual. 2014. “Altitudinal Changes in Malaria Incidence in Highlands of Ethiopia and Colombia.” *Science*.

Stefani, Aurélia, Emmanuel Roux, Jean Marie Fotsing, and Bernard Carme. 2011.

"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*.

Su, Shiliang, Rui Xiao, and Yuan Zhang. 2012. "Multi-Scale Analysis of Spatially Varying Relationships between Agricultural Landscape Patterns and Urbanization Using Geographically Weighted Regression." *Applied Geography*.

Taye, Mulugojjam, Ferede Zewdu, and Dereje Ayalew. 2013. "Characterizing the Climate System of Western Amhara , Ethiopia : A GIS Approach." *American Journal of Research Communication*.

Thomson, Madeleine C., Israel Ukawuba, Christine L. Hershey, Adam Bennett, Pietro Ceccato, Bradfield Lyon, and Tufa Dinku. 2017. "Using Rainfall and Temperature Data in the Evaluation of National Malaria Control Programs in Africa." *American Journal of Tropical Medicine and Hygiene*.

Videras, Julio. 2014. "Exploring Spatial Patterns of Carbon Emissions in the USA: A Geographically Weighted Regression Approach." *Population and Environment*.

Viste, Ellen, Diriba Korecha, and Asgeir Sorteberg. 2013. "Recent Drought and Precipitation Tendencies in Ethiopia." *Theoretical and Applied Climatology*.

Wabiri, Njeri, Olive Shisana, Khangelani Zuma, and Jeffrey Freeman. 2016. "Assessing the Spatial Nonstationarity in Relationship between Local Patterns of HIV Infections and the Covariates in South Africa: A Geographically Weighted Regression Analysis." *Spatial and Spatio-Temporal Epidemiology*.

Wang, Yiyi, Kara M. Kockelman, and Xiaokun Wang. 2011. "Anticipation of Land Use Change through Use of Geographically Weighted Regression Models for Discrete Response." *Transportation Research Record*.

Weli, Vincent Ezikornwor and Sunday Ighovie Efe. 2015. "Climate and Epidemiology of Malaria in Port Harcourt Region, Nigeria." *American Journal of Climate Change*.

WHO. 2016. "Fact Sheet on Malaria." *Geneva: World Health Organization (WHO)*.

WHO. 2018. *World Malaria Report 2018. Geneva:World Health Organization; 2018*.

Wijayanti, Siwi P. M., Thibaud Porphyre, Margo Chase-Topping, Stephanie M. Rainey, Melanie McFarlane, Esther Schnettler, Roman Biek, and Alain Kohl. 2016. "The Importance of Socio-Economic Versus Environmental Risk Factors for Reported Dengue Cases in Java, Indonesia." *PLoS Neglected Tropical*

Diseases.

Yao, Yonghui and Baiping Zhang. 2013. "MODIS-Based Estimation of Air Temperature of the Tibetan Plateau." *Journal of Geographical Sciences*.

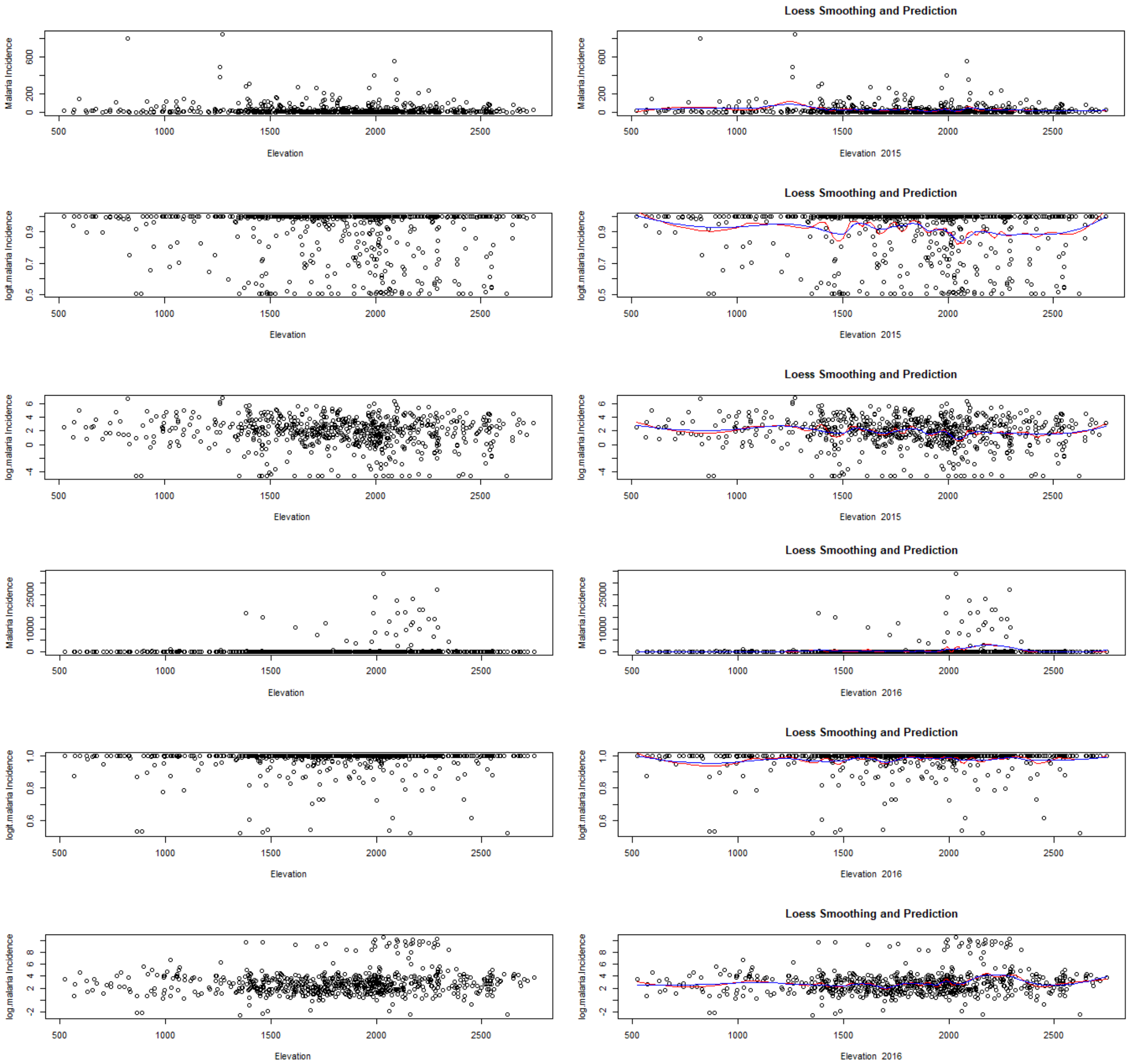
Yeshiwondim, Asnakew K., Sucharita Gopal, Afework T. Hailemariam, Dereje O. Dengela, and Hrishikesh P. Patel. 2009. "Spatial Analysis of Malaria Incidence at the Village Level in Areas with Unstable Transmission in Ethiopia." *International Journal of Health Geographics*.

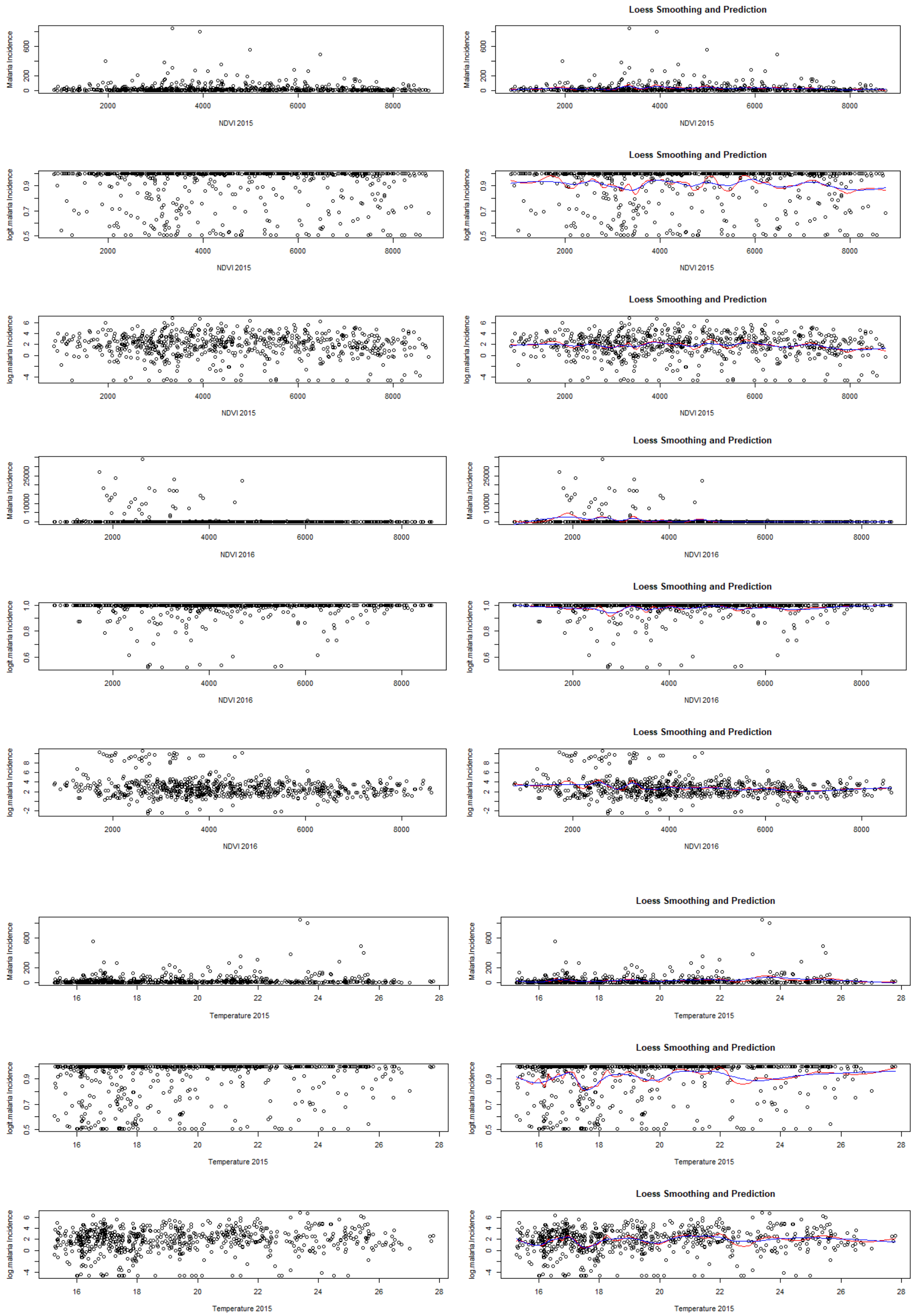
7 APPENDICES

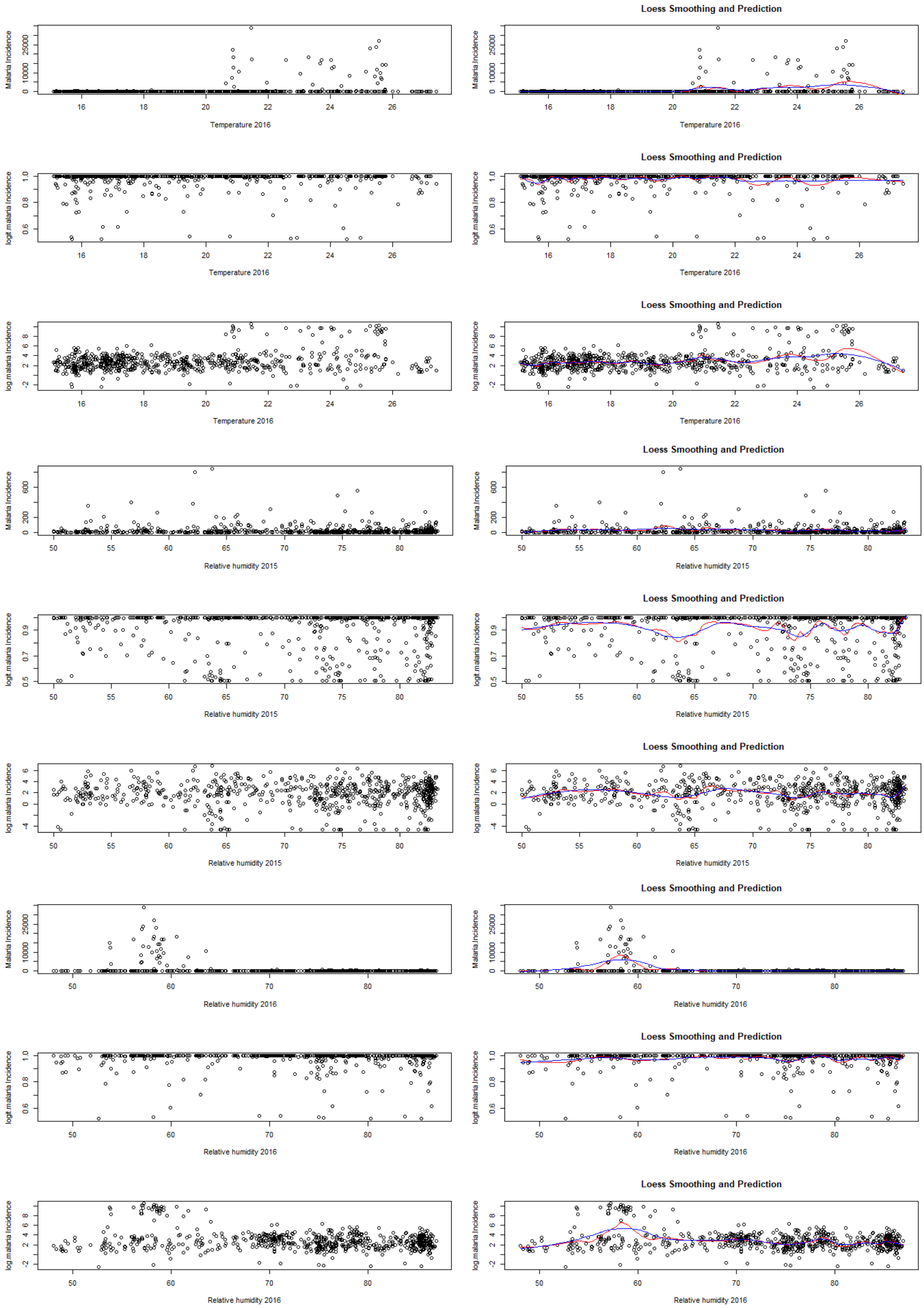
7.1 Data cleaning R code

```
setwd("C:/Users/Gamu/Desktop")

s<-read.csv("2016datacleaning.csv", TRUE , ",")

library(tidyverse)
library(psych)
library(skimr)
library(broom)
glimpse(s)
summary(s)
describe(s)
skim(s)
corr.test(s$MC2016 , s$TM2016)
cor.test(s$MC2016 , s$TM2016)
plot(s$MC2016 , s$TM2016)
reOutlier <- filter(s, WOREDANAME !=  'Diga')
ct1<-cor.test(s$MC2016 , s$TM2016)

plot(reOutlier$MC2016 , reOutlier$TM2016)
cor.test(reOutlier$MC2016 , reOutlier$TM2016)
ct2<-cor.test(reOutlier$MC2016 , reOutlier$TM2016)

ct1_t<-tidy(ct1)
ct1_tp<-ct1_t$p.value
```

7.2 Scatter plot of log and logit of malaria incidence for all explanatory variables in 2015 and 2016




Loess Smoothing and Prediction
Malaria Incidence
NDVI 2015
logit malaria Incidence
log malaria Incidence
NDVI 2016
Temperature 2015


Loess Smoothing and Prediction
Malaria Incidence
Temperature 2016
logit malaria incidence
Temperature 2016
log malaria incidence
Temperature 2016
Malaria Incidence
Relative humidity 2015
logit malaria incidence
Relative humidity 2015
log malaria incidence
Relative humidity 2015
Malaria Incidence
Relative humidity 2016
logit malaria incidence
Relative humidity 2016
log malaria incidence
Relative humidity 2016

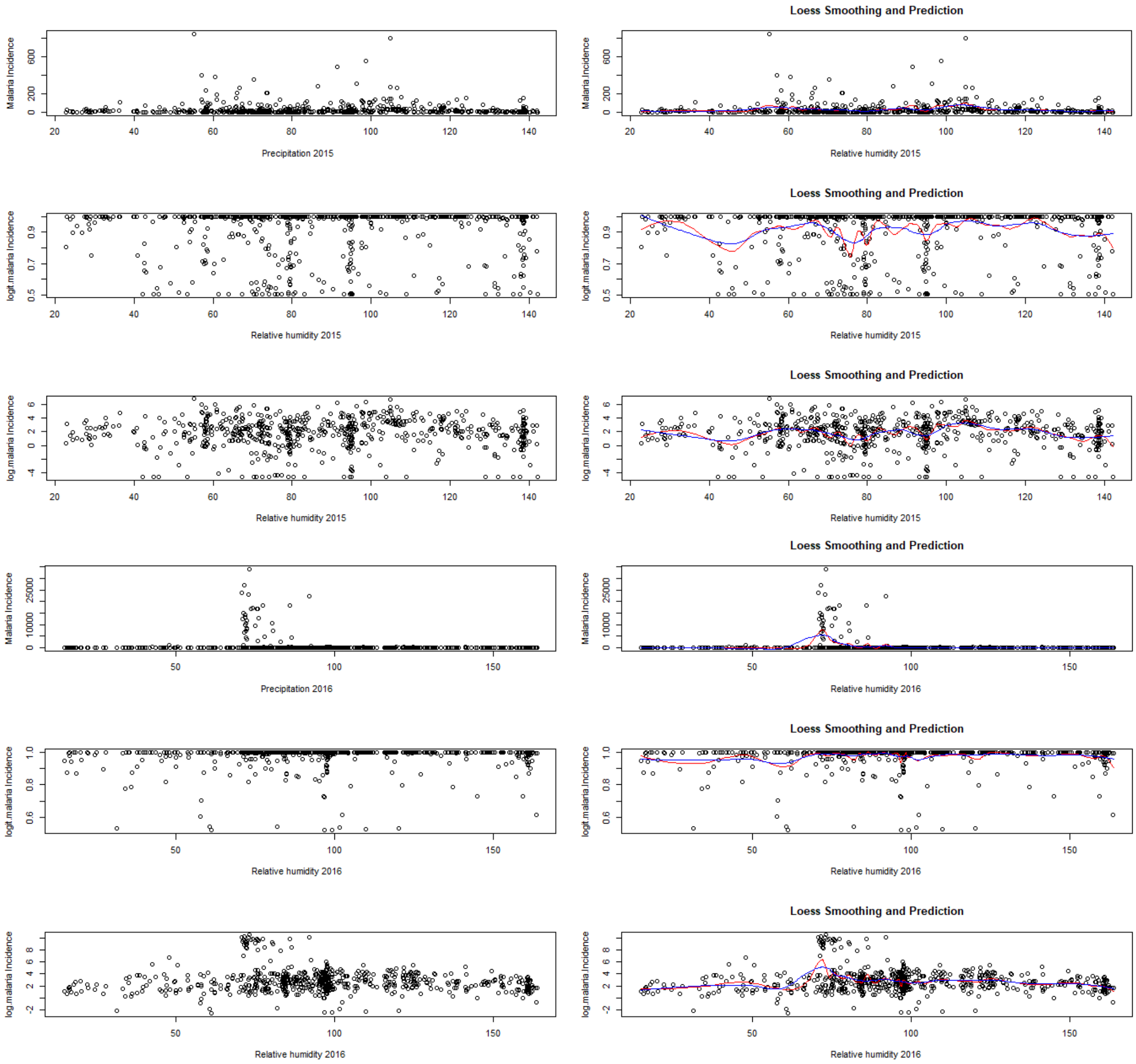


Figure 17 scatter plot of log and logit of malaria incidence for all explanatory variables in 2015 and 2016

7.3 R code for exploratory analysis to discover if bivariate relationships were linear or not using B-splines

```
setwd("C:/Users/Gamu/Desktop/GWR2020/Jorge Mateu/Rcode")
s<-read.csv("dec_20.csv", TRUE , ",")
library(tidyverse)
library(psych)
library(skimr)

malaria=s[,1]
january.rh=s[,2]
january.tm=s[,3]
january.pr=s[,4]
january.el=s[,5]
january.en=s[,6]
```

```
glimpse(s)
summary(s)
describe(s)
skim(s)

corr.test(s$MC2016,s$TM2016)
cor.test(s$MC2016,s$TM2016)
plot(s$MC2016,s$TM2016)
names(s)
sc<-filter(s,names != TM2016) # TO Removing the outliers
```

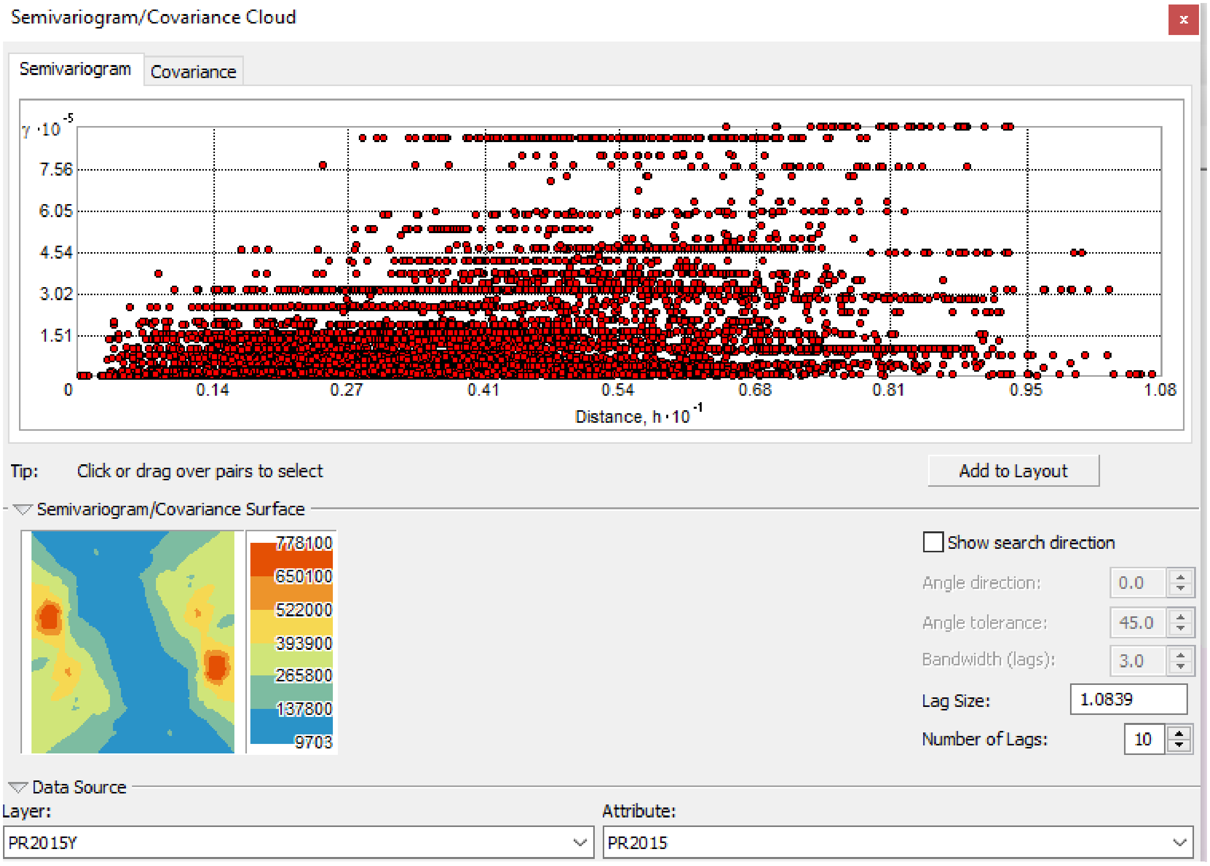
Explanatory data analysis

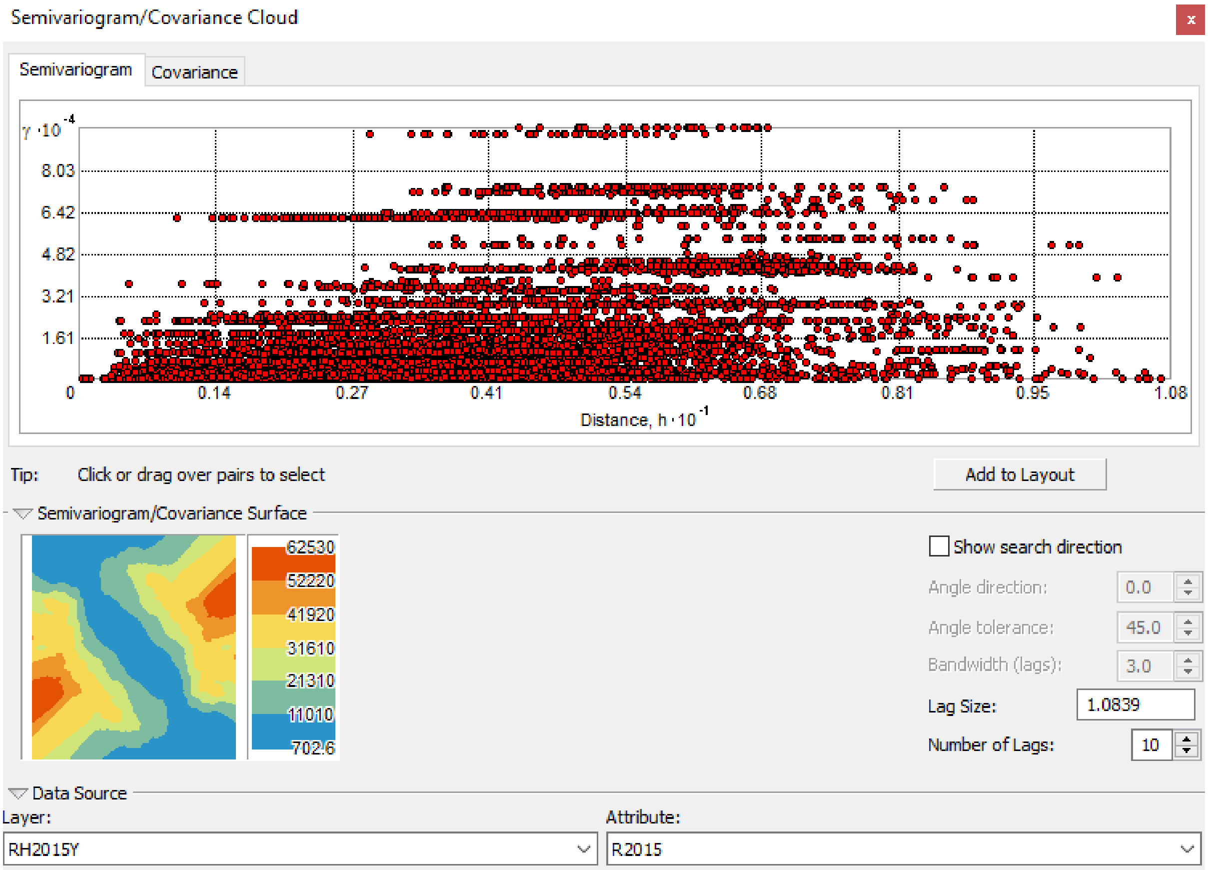
```
# STEP 1: defining the data set

malaria=s[,1]
january.pr=s[,4]
january.tm=s[,5]
january.rh=s[,6]
# checking if there is negative values or errore in the data
summary(data1[,5])
sum(data1[,5]<=0)




logit.malaria=logistic(malaria,d=0, a=1,c=0, z=1)
log.malaria=log(malaria)
data=cbind(malaria,logit.malaria,log.malaria,january.pr,january.tm,january.rh)
data1=data[-c(678,679),] #remove two rows that have zeros in the covs

data.ordered.temp=data1[order(data1[,5]),] #order in increasing order from Temp data
data.ordered.Rh=data1[order(data1[,6]),] #order in increasing order from Rh data
data.ordered.Pr=data1[order(data1[,4]),] #order in increasing order from Pr data

# STEP 2: some plots for temperature

par(mfrow=c(2,2))
plot(data1[,5],data1[,1],xlab="Temperatures",ylab="Malaria cases")
plot(data1[,5],data1[,2],xlab="Temperatures",ylab="logit.malaria",ylim=c(0,1))
plot(data1[,5],data1[,3],xlab="Temperatures",ylab="log.malaria",ylim=c(0,1))

# STEP 3: loess of "logit of malaria" prevalence against temperature
# & loess of "malaria cases" prevalence against temperature

# Loess Regression...Loess Regression is the most common method used to smoothen a
volatile time series. # It is a non-parametric methods where least squares regression is
performed in localized subsets, which # makes it a suitable candidate for smoothing any
numerical vector.

# data.ordered.temp[,2] this is "logit of malaria"
```

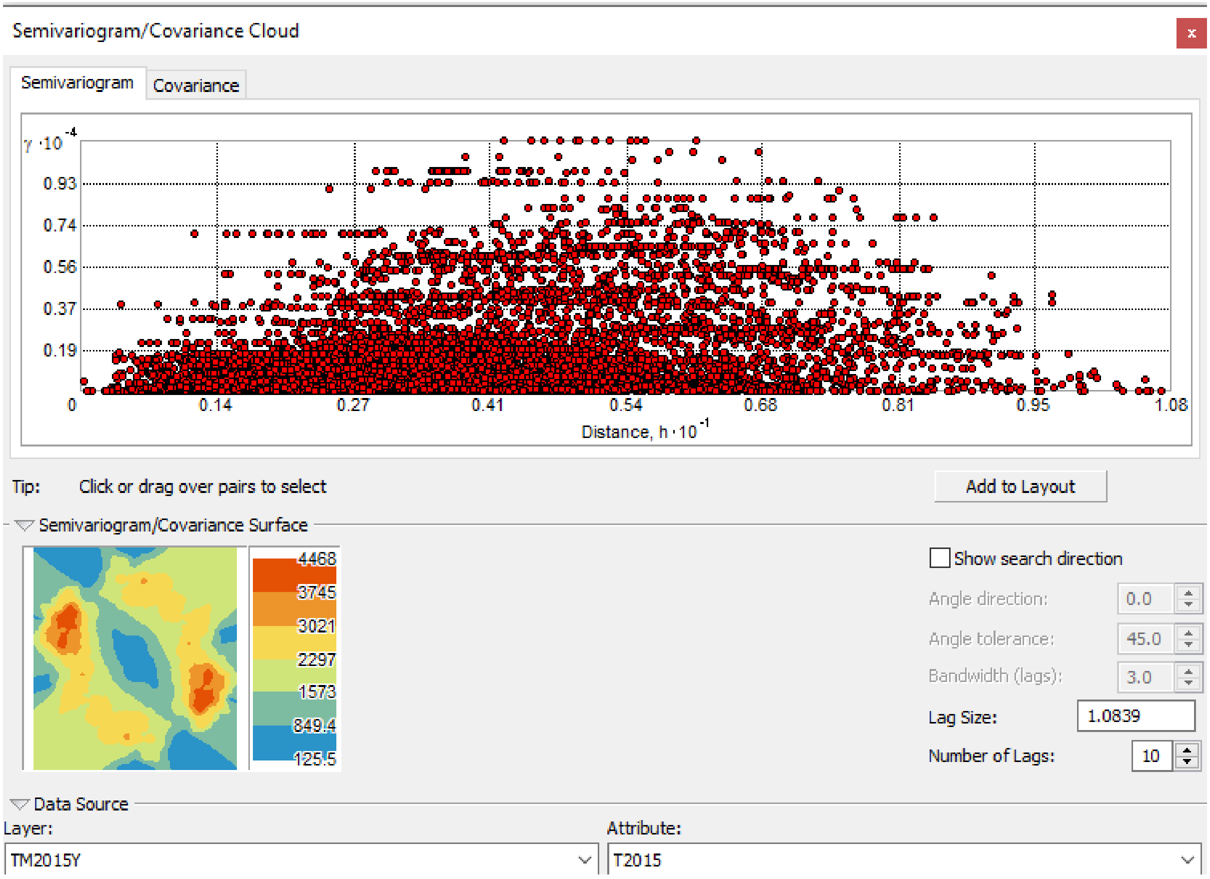
```
# data.ordered.temp[,4] this is "temp"
# data.ordered.temp[,3] this is "Pr"
# data.ordered.temp[,5] this is "RH"
# if you are using Rh then use data.ordered.Rh[,5]
# if you are using Pr then use data.ordered.Pr[,3]

loess.logit.malaria.Temp10=loess(data.ordered.temp[,2] ~ data.ordered.temp[,5],
span=0.10) # 10% span
loess.logit.malaria.Temp20=loess(data.ordered.temp[,2] ~ data.ordered.temp[,5],
span=0.20) # 20% span
loess.log.malaria.Temp11=loess(data.ordered.temp[,3] ~ data.ordered.temp[,5],
span=0.10) # 10% span
loess.log.malaria.Temp21=loess(data.ordered.temp[,3] ~ data.ordered.temp[,5],
span=0.20) # 20% span

loess.malaria.Temp10=loess(data.ordered.temp[,1] ~ data.ordered.temp[,5], span=0.10)
loess.malaria.Temp20=loess(data.ordered.temp[,1] ~ data.ordered.temp[,5], span=0.20)

loess.malaria.Temp11=loess(data.ordered.temp[,1] ~ data.ordered.temp[,5], span=0.10)
loess.malaria.Temp21=loess(data.ordered.temp[,1] ~ data.ordered.temp[,5], span=0.20)
# STEP 4: get smoothed output

smoothed.logit.malaria.10=predict(loess.logit.malaria.Temp10)
smoothed.logit.malaria.20=predict(loess.logit.malaria.Temp20)
smoothed.log.malaria.11=predict(loess.log.malaria.Temp11)
smoothed.log.malaria.21=predict(loess.log.malaria.Temp21)

smoothed.malaria.10=predict(loess.malaria.Temp10)
smoothed.malaria.20=predict(loess.malaria.Temp20)
smoothed.malaria.11=predict(loess.malaria.Temp11)
smoothed.malaria.21=predict(loess.malaria.Temp21)

# # STEP 5: some plots

par(mfrow=c(1,2))
plot(data1[,5],data1[,1],xlab="Temperatures",ylab="Malaria cases")
plot(data1[,5],data1[,2],xlab="Temperatures",ylab="logit.malaria",ylim=c(0,1))

par(mfrow=c(1,2))

plot(x=data.ordered.temp[,5], y=data.ordered.temp[,2], main="Loess Smoothing and
Prediction", xlab="january.Tmp", ylab="logit.malaria")
lines(data.ordered.temp[,5], smoothed.logit.malaria.10, col="red")
lines(data.ordered.temp[,5], smoothed.logit.malaria.20, col="blue")

plot(x=data.ordered.temp[,5], y=data.ordered.temp[,1], main="Loess Smoothing and
Prediction", xlab="january.Tmp", ylab="Malaria cases")
lines(data.ordered.temp[,5], smoothed.malaria.10, col="red")
lines(data.ordered.temp[,5], smoothed.malaria.20, col="blue")
```

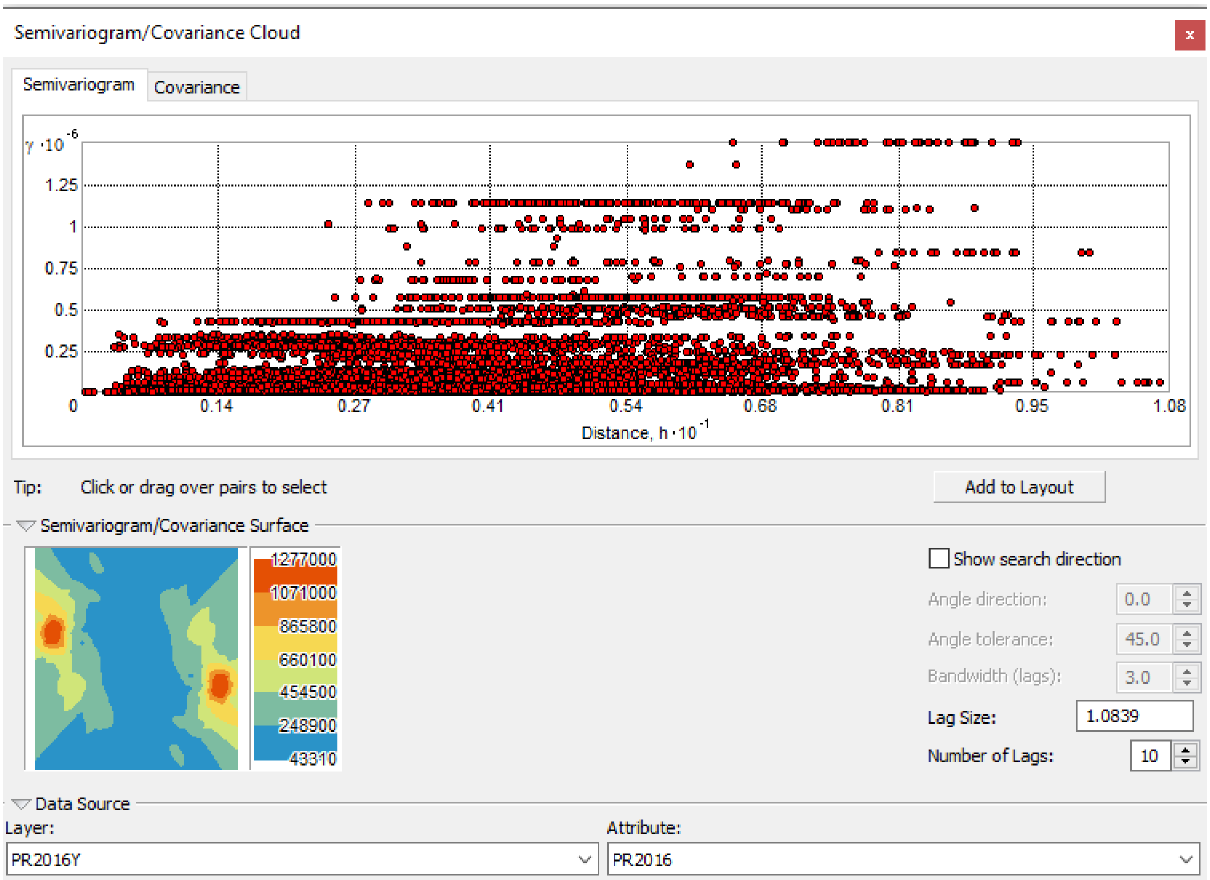
```
par(mfrow=c(3,2))

plot(data1[,5],data1[,1],xlab="Elevation       ",ylab="Malaria Incidence")

plot(x=data.ordered.temp[,5], y=data.ordered.temp[,1], main="Loess Smoothing and Prediction", xlab=" Elevation 2016", ylab="Malaria.Incidence")
lines(data.ordered.temp[,5], smoothed.malaria.10, col="red")
lines(data.ordered.temp[,5], smoothed.malaria.20, col="blue")

plot(data1[,5],data1[,2],xlab="Elevation      ",ylab="logit.malaria Incidence",ylim=c(0,1))

plot(x=data.ordered.temp[,5], y=data.ordered.temp[,2], main="Loess Smoothing and Prediction", xlab= "Dec Elevation 2016", ylab="logit.malaria.Incidence")
lines(data.ordered.temp[,5], smoothed.logit.malaria.10, col="red")
lines(data.ordered.temp[,5], smoothed.logit.malaria.20, col="blue")

aa=cbind(data.ordered.temp[,5],data1[,5])



plot(x=data.ordered.temp[,5], y=data.ordered.temp[,3] ,xlab="Elevation ",ylab="log.malaria Incidence")
plot(x=data.ordered.temp[,5], y=data.ordered.temp[,3], main="Loess Smoothing and Prediction", xlab= "2016", ylab="log.malaria.Incidence")
lines(data.ordered.temp[,5], smoothed.log.malaria.11, col="blue")
lines(data.ordered.temp[,5], smoothed.log.malaria.21, col="green")
```

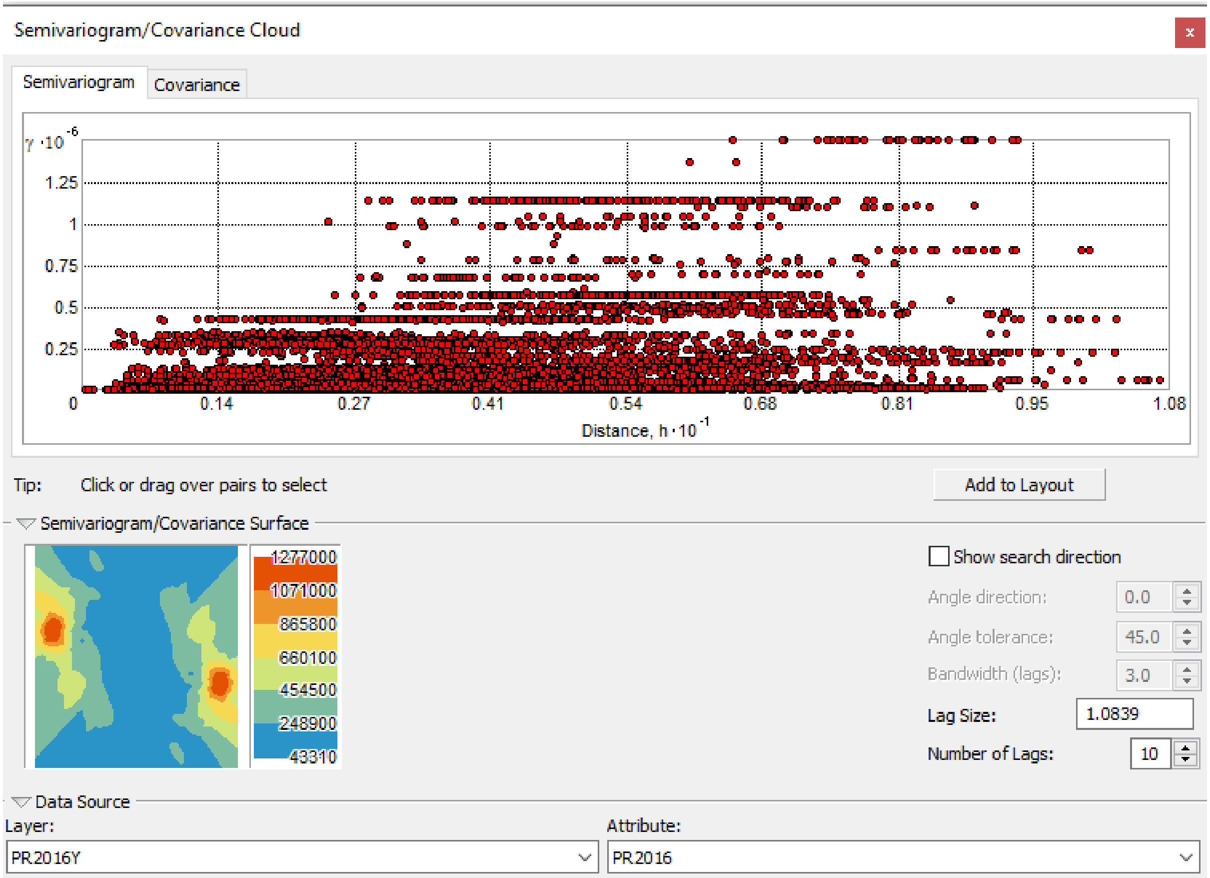
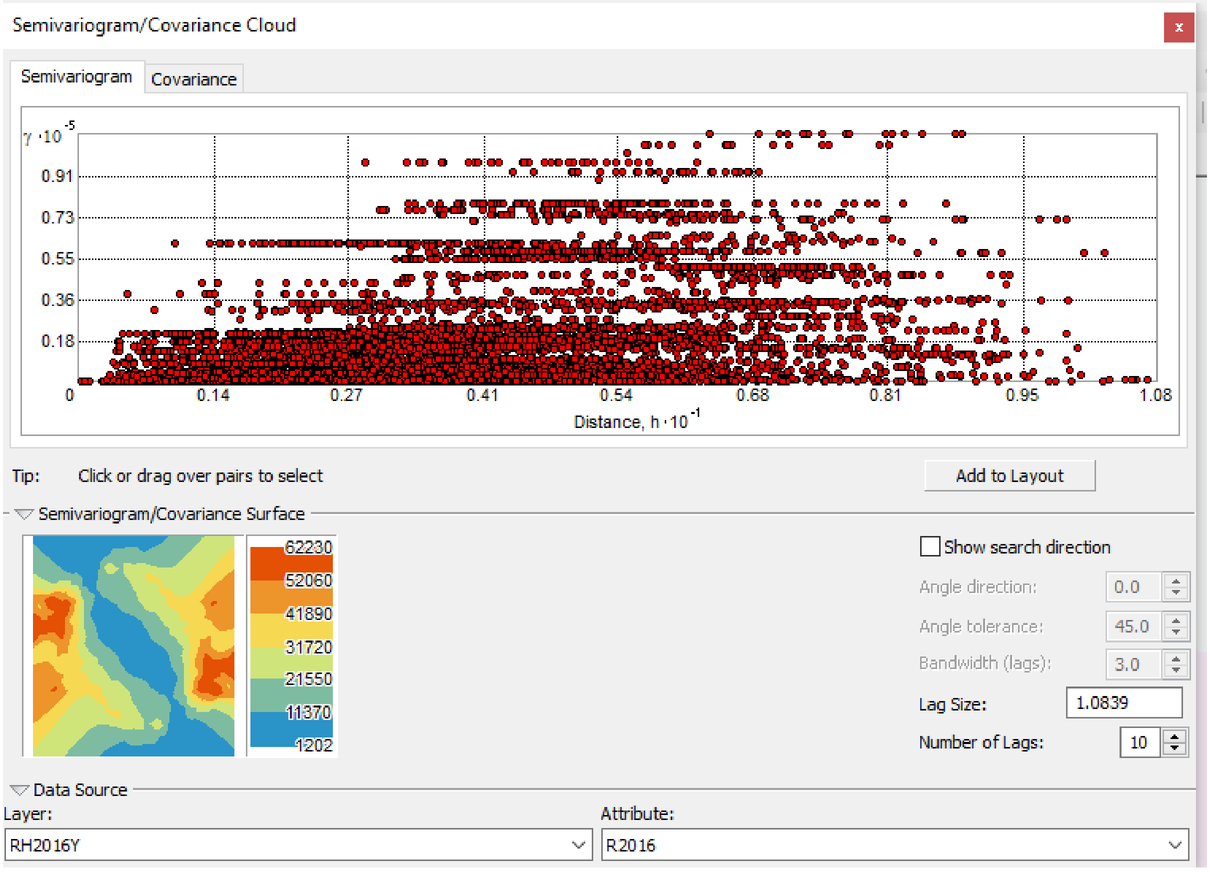
7.4 Results of (temperature, Precipitation, elevation and Relative humidity) Variogram models used for interpolation of explanatory variables with ordinary kriging

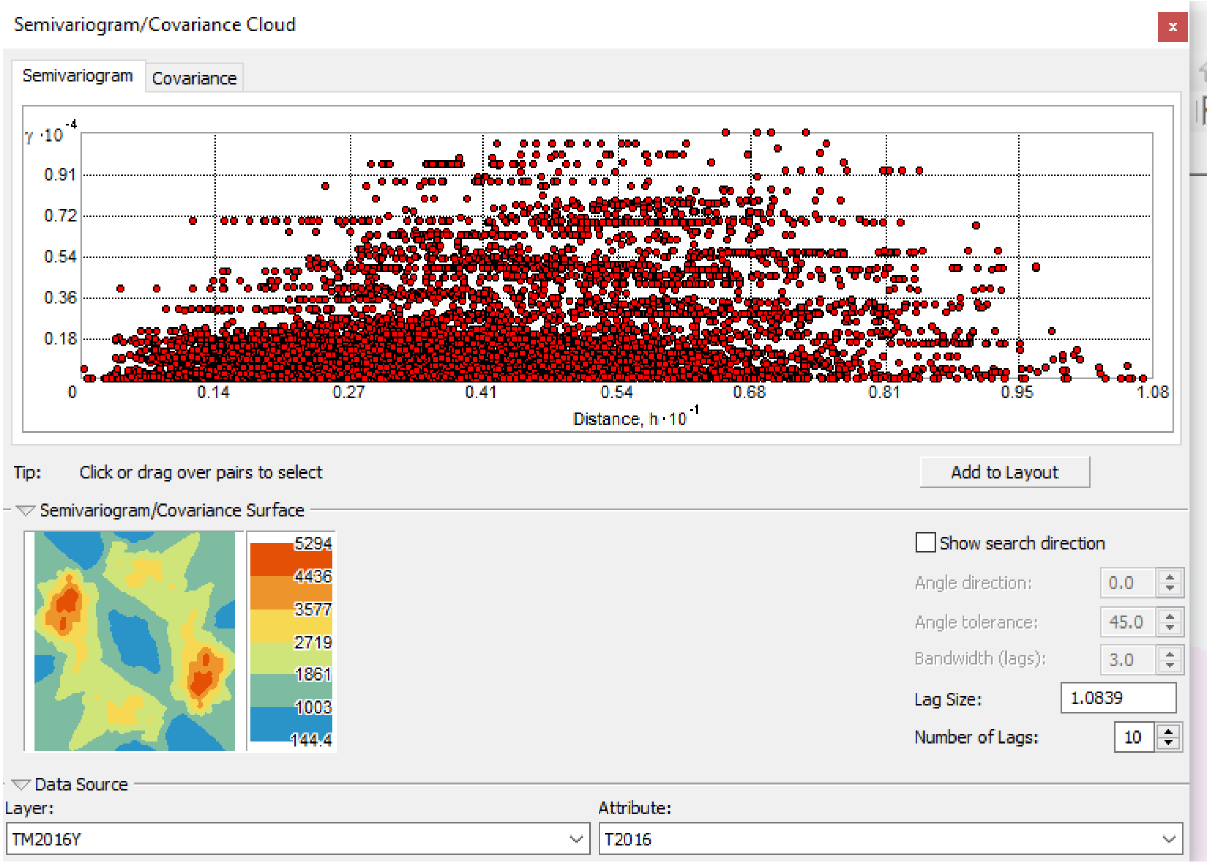

Semivariogram/Covariance Cloud
Semivariogram
Covariance
γ ·10^-5
7.56
6.05
4.54
3.02
1.51
0
0.14
0.27
0.41
0.54
0.68
0.81
0.95
1.08
Distance, h·10^-1
Tip: Click or drag over pairs to select
Add to Layout
Semivariogram/Covariance Surface
778100
650100
522000
393900
265800
137800
9703
Show search direction
Angle direction: 0.0
Angle tolerance: 45.0
Bandwidth (lags): 3.0
Lag Size: 1.0839
Number of Lags: 10
Data Source
Layer:
PR2015Y
Attribute:
PR2015


Semivariogram/Covariance Cloud
Semivariogram
Covariance
γ ·10^-4
8.03
6.42
4.82
3.21
1.61
0
0.14
0.27
0.41
0.54
0.68
0.81
0.95
1.08
Distance, h·10^-1
Tip: Click or drag over pairs to select
Add to Layout
Semivariogram/Covariance Surface
62530
52220
41920
31610
21310
11010
702.6
Show search direction
Angle direction: 0.0
Angle tolerance: 45.0
Bandwidth (lags): 3.0
Lag Size: 1.0839
Number of Lags: 10
Data Source
Layer:
RH2015Y
Attribute:
R2015


Semivariogram/Covariance Cloud
Semivariogram
Covariance
γ ·10^-4
0.93
0.74
0.56
0.37
0.19
0
0.14
0.27
0.41
0.54
0.68
0.81
0.95
1.08
Distance, h · 10^-1
Tip: Click or drag over pairs to select
Add to Layout
Semivariogram/Covariance Surface
4468
3745
3021
2297
1573
849.4
125.5
Show search direction
Angle direction: 0.0
Angle tolerance: 45.0
Bandwidth (lags): 3.0
Lag Size: 1.0839
Number of Lags: 10
Data Source
Layer:
TM2015Y
Attribute:
T2015


Semivariogram/Covariance Cloud
Semivariogram
Covariance
γ ·10^-6
1.25
1
0.75
0.5
0.25
0
0.14
0.27
0.41
0.54
0.68
0.81
0.95
1.08
Distance, h · 10^-1
Tip: Click or drag over pairs to select
Add to Layout
Semivariogram/Covariance Surface
1277000
1071000
865800
660100
454500
248900
43310
Show search direction
Angle direction: 0.0
Angle tolerance: 45.0
Bandwidth (lags): 3.0
Lag Size: 1.0839
Number of Lags: 10
Data Source
Layer:
PR2016Y
Attribute:
PR2016


Semivariogram/Covariance Cloud
Semivariogram
Covariance
γ ·10^-6
1.25
1
0.75
0.5
0.25
0
0.14
0.27
0.41
0.54
0.68
0.81
0.95
1.08
Distance, h ·10^-1
Tip: Click or drag over pairs to select
Add to Layout
Semivariogram/Covariance Surface
1277000
1071000
865800
660100
454500
248900
43310
Show search direction
Angle direction: 0.0
Angle tolerance: 45.0
Bandwidth (lags): 3.0
Lag Size: 1.0839
Number of Lags: 10
Data Source
Layer:
PR2016Y
Attribute:
PR2016

Semivariogram/Covariance Cloud
Semivariogram
Covariance
γ ·10^-5
0.91
0.73
0.55
0.36
0.18
0
0.14
0.27
0.41
0.54
0.68
0.81
0.95
1.08
Distance, h ·10^-1
Tip: Click or drag over pairs to select
Add to Layout
Semivariogram/Covariance Surface
62230
52060
41890
31720
21550
11370
1202
Show search direction
Angle direction: 0.0
Angle tolerance: 45.0
Bandwidth (lags): 3.0
Lag Size: 1.0839
Number of Lags: 10
Data Source
Layer:
RH2016Y
Attribute:
R2016



Semivariogram/Covariance Cloud
Semivariogram
Covariance
γ ·10^-4
0.91
0.72
0.54
0.36
0.18
0
0.14
0.27
0.41
0.54
0.68
0.81
0.95
1.08
Distance, h · 10^-1
Tip: Click or drag over pairs to select
Add to Layout
Semivariogram/Covariance Surface
5294
4436
3577
2719
1861
1003
144.4
Show search direction
Angle direction: 0.0
Angle tolerance: 45.0
Bandwidth (lags): 3.0
Lag Size: 1.0839
Number of Lags: 10
Data Source
Layer: TM2016Y
Attribute: T2016