Geographically Weighted Regression Modelling of the Spatial Association Between Malaria Cases and Environmental Factors in Cameroon

Tewara AM[1, 2, 3*], Yunxia L[1, 2], Ling W[1, 2], Binang HB[4], Nji CP[3], Mbah-Fongkimeh PN[5], Zhaolei Z[1], Xinhui L[1], Miao Z[1], Liu X[1] and Xue F[1, 2]

[1]Department of Biostatistics, School of Public Health, Shandong University Cheeloo College of Medicine, Jinan 250012, PR China

[2]Institute for Medical Dataology, Shandong University, Jinan, 250002, P.R. China

[3]Center for Health Promotion and Research (Former Tuberculosis Reference Laboratory Bamenda) P.O. Box 586, Bamenda Cameroon

[4]Department of Clinical Laboratory Diagnostics, Shandong University Cheeloo College of Medicine, Jinan 250012, PR China

[5]School of Medicine and Public Health, Faculty of Health and Medicine, The University of New Castle, Callaghan, Australia

*Corresponding author:

Marlvin Anemey Tewara,
Institute for Medical Dataology,
Shandong University, Jinan, 250002,
P.R. China,
Email: tewaramarlvin@yahoo.com

1. Abstract

1.1. Background: Studies have illustrated the association of malaria cases with environmental factors in Cameroon but limited in addressing how these factors vary in space for timely public health interventions. Thus, we want to find the spatial variability between malaria hotspot cases and environmental predictors using Geographically weighted regression (GWR) spatial modelling technique.

1.2. Materials and Methods: The global Ordinary least squares (OLS) in the modelling spatial relationships tool in ArcGIS 10.3. was used to select candidate explanatory environmental variables for a properly specified GWR model. The local GWR model used the global OLS candidate variables to examine, predict and explore the spatial variability between environmental factors and malaria hotspot cases generated from Getis-Ord Gi* statistical analysis.

1.3. Results: The OLS candidate environmental variable coefficients were statistically significant (adjusted R2 = 22.3% and p < 0.01) for a properly specified GWR model. The GWR model identified a strong spatial association between malaria cases and rainfall, vegetation index, population density, and drought episodes in most hotspot areas and a weak correlation with aridity and proximity to water with an overall model performance of 0.243 (adjusted R2= 24.3%).

1.4. Conclusion: The generated GWR maps suggest that for policymakers to eliminate malaria in Cameroon, there should be the creation of malaria outreach programs and further investigations in areas where the environmental variables showed strong spatial associations with malaria hotspot cases.

2. Abbreviations

SSA: Sub-Saharan Africa; WHO: World health organization; GTS: Global Technical Strategy; GIS: Geographic Information Systems; OLS: Ordinary Least Squares; GWR: Geographically Weighted Regression; ITN: Insecticide Treated Net; DHS: Demographic and Health Survey; USAIDS: United States Agency for International Development; EVI: Enhanced Vegetation index; AICc: Akaike's Information Criterion; VIF: Variance Inflation Factor.

3. Introduction

About 70% of the world's malaria burden is focused in just 11 countries; 10 in Sub-Saharan Africa (SSA) (Burkina Faso, Cameroon, the Democratic Republic of Congo, Ghana, Mali, Mozambique, Niger, Nigeria, Uganda and the United Republic of Tanzania) and India. These high-burden nations housed an estimated 151 million cases of malaria and 275 000 deaths. In 2017, all the high burden endemic countries in Africa saw an increase in malaria cases with Cameroon witnessing an increase of 131,000 malaria cases [1], from a total of

1,618,738 confirmed cases reported in health facilities and communities in 2016 [2]. The epidemiological transmission of malaria in Cameroon is high (> 1 case per 1000 population) in about 71% (16.6 million people) and low (0–1 case per 1000 population) in about 29% (6.8 million) in people of all sexes and age groups with children less than five at greater risk of the disease [2]. In 2014, the morbidity of malaria in Cameroon was 30% in children and 18% in Adults [3]. The government of Cameroon and partners have been combating malaria through the creation of national intervention programs including the distribution of free insecticide-treated nets (ITN) that was established in 2011 to populations at high risk, provision of sulfadoxine-pyrimethamine drugs to pregnant woman, parasitological screening of suspected malaria cases, and the application of other WHO standard treatments [3,4]. The socio-economic and environmental challenges posed by the malaria disease to Africa countries is a global concern. The WHO's Global Technical Strategy (GTS) for Malaria 2016- 2030 has been developed with the aim of helping countries reduce the human suffering caused by the disease. Adopted by the World Health Assembly in May 2015, the strategy provides comprehensive technical guidance to countries and development partners for 15 years, emphasizing the importance of scaling up malaria responses and moving towards elimination. It also highlights the urgent need to increase investments across all interventions – including preventive measures, diagnostic testing, treatment and disease surveillance – as well as in harnessing innovation and expanding research [5]. Intensifying investments in malaria research by endemic SSA is a key to attaining the GTS targets and eradicating the disease from the SSA geolocations.

The application of spatial statistical methods to geolocational health data research has enabled complex scenarios of the malaria disease to be visualized through the creation of spatial maps within the Geographical information systems (GIS) technology [6-11]. The study of the spatial variation between disease outcomes and associative socioeconomic or environmental factors using the GIS has greatly improved our understanding of these factors with the health outcome in question. Malaria has been reported to be associated with environmental and climatic factors such as rainfall, humidity, temperature [12,13] and understanding the behavior of these factors in space with the application of spatial regression statistics [14] will further improve on timely control measures and resource allocations.

Regression analyses are statistical techniques that allow for the modelling, examining, and exploring of spatial relationships, to better understand the factors behind observed spatial patterns and hotspots, and to predict outcomes based on that understanding [14]. Ordinary Least Squares regression (OLS) is a global regression method that provides a global model of the variable or process to be predicted or studied. It creates a single regression equation to represent that process. Geographically Weighted Regression (GWR) is a local spatial regression method that allows the relationships to be modelled to vary across the study area by fitting a regression equation to every

feature in the dataset using candidate explanatory variables from the OLS. It is a local form of linear regression used to model spatially varying relationships. GWR statistical modelling technique has been applied to a range of malaria studies: Hasyim [15], used the GWR to find the spatial association between malaria cases and environmental factors in South Sumatra, Indonesia where altitude, distance from forest and rainfall were associated with malaria, Moise [16], in the seasonal and geographic variation of pediatric malaria in Burundi, identified the spatial variation between monthly rainfall and malaria prevalence. The GWR spatial modelling technique has been a powerful tool in the understanding of malaria prevention and the spatial variability of malaria cases and environmental factors [15-18]. Its application has been valuable in the understanding of other infectious diseases such as the spatial association between dengue fever, and socioeconomic and environmental determinants [19,20]. Also, GWR has been applied in other health outcomes and social science studies including cancer events [21], mental depression [22], fire events [23], hospital accessibility study [24], alcohol and violence [25] and real estate housing crisis [26].

Massoda [27], compared malaria survey programs in different ecological zones in Cameroon and recommended on the needs of intervention programs during high transmission rainy seasons. Furthermore, Tewara [28], in a recent study on small area spatial statistical analysis of malaria clusters and hotspots in Cameroon, illustrated the linear association between malaria cases and environmental factors using the Pearson correlation statistics [29] but didn't demonstrate any spatial variability that would become the main aim of this study. The specific objective of this study is to find the spatial variability between malaria hotspot cases and environmental predictors using the GWR spatial modelling technique.

4. Materials and Methods

4.1. Data Source

Data for this study was obtained from the demographic and health survey (DHS) program website (https://dhsprogram.com) funded by the United States Agency for International Developments (US-AID) following a written administrative clearance [30]. The DHS program has been described elsewhere [28]. Households are grouped as either an urban (city block or apartment building) or rural (village or group of villages) cluster points (figure 1) and displaced a distance up to 2 km for urban-city clusters and 5 km for rural clusters due to confidentiality. For this study, points data for the Cameroon malaria survey was obtained from the DHS spatial data repository site [31]. The malaria survey data for 2015 was linked with environmental covariates data; Enhanced Vegetation Index (EVI), rainfall, drought episodes, population density, aridity, proximity to water, and analyzed using the ArcGIS 10.3 (ESRI, Redlands, California, USA) software.

4.2. Malaria and Environmental Data Description

The WHO recommends that all cases of suspected malaria should be confirmed using parasite-based diagnostic testing (either micros-

copy or rapid diagnostic test) before administering treatment. Thus, malaria cases were confirmed based on both rapid diagnostic tests and on laboratory analysis. A clinical case was defined as a malaria-attributable febrile episode (body temperature in excess of 37.5 °C), accompanied by headaches, nausea, excess sweating and/or fatigue censored by a 30-day window [6,27]. Since the households are the variables to be analyzed, a malaria year as described in this study, is the average number of people per year who show clinical symptoms of Plasmodium falciparum malaria within the cells whose centroid falls within a radius of 10 km (for rural points) or 2 km (for urban points) [28]. The environmental covariates data set used for this study are described in table 1.
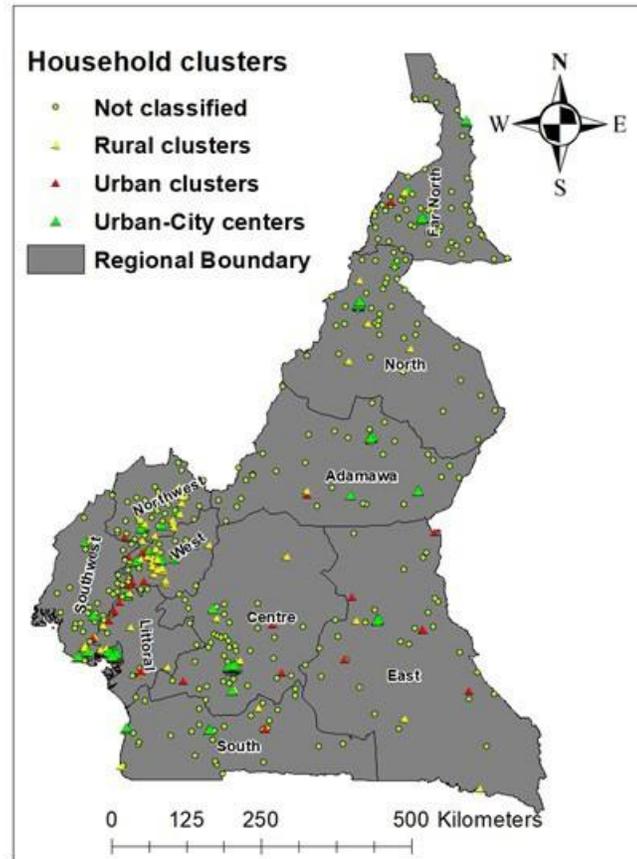

Figure 1: DHS household survey cluster points.

Table 1: Covariates data used for this study

Variable	Units	Description	Resolution	Data source
Population density	Number of people	The average number of people in the cells whose centroid falls within a radius of 10 km[#] or 2 km*.	2.5 minutes	[32]
Aridity	Aridity Index between 0.01 (Hyper-Arid) and 0.99 (Humid)	The average aridity index of the cells whose centroid falls within a radius of 10 km[#] or 2 km*.	Unspecified	[33]
Drought episodes	Individual classes between 1 (Low Drought) and 10 (High Drought	The average of the drought episodes indices of the cells whose centroid falls within a radius of 10 km[#] or 2 km*.	Unspecified	[34]
EVI year	EVI value between 0 (least vegetation) and 10000 (Most vegetation	The EVI was calculated by measuring the density of green leaves in the near-infrared and visible bands and describe as the average EVI of the cells whose centroid falls within a radius of 10 km[#] or 2 km*.	Unspecified	[35]
Proximity to water	Meters	Straight-line distance to the nearest major water body. Based on the World Vector Shorelines, CIA World Data Bank II, and Atlas of the Cryosphere	1:1,000,000	[36,37]
Rainfall	Millimeters per year	The average rainfall of the cells whose centroid falls within a radius of 10 km[#] or 2 km*.	30 minutes	[38,39]

rural points and * urban points.

4.3. Ethical Approval

Permission to use the data was obtained through a written request and subsequent approval from the DHS division of the USAID. During the DHS project, interviews and blood test analyses are conducted only if the respondent provides voluntary informed consent. Written informed consent was obtained from all participants.

4.4. Statistical Analysis

4.4.1 Getis-Ord Gi* statistics: The Getis-Ord Gi* statistics is a local statistic that allows us to discover new locations with significant clusters of hot and coldspots. It assesses each malaria household cluster (or feature) within the context of neighboring malaria households and compares the local situation to the global situation. The 2015 Malaria indicator survey data was analyzed for malaria hotspots.

The Getis-Ord Gi* local statistics is given as:

$$G_i^* = \frac{\sum_{i=1}^{n} w_{i,j}\, x_j - \bar{X}\sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{\left[n\sum_{j=1}^{n} w^2_{i,j} - \left(\sum_{j=1}^{n} w_{i,j}\right)^2\right]}{n-1}}} \qquad (1)$$

Where x_j is the attribute value for feature j; $w_{i,j}$ is the spatial weight between i and j; n is equal to the total number of features and:

$$\bar{X} = \frac{\sum_{j=1}^{n} w_j}{n} \qquad (2)$$

$$S = \sqrt{\frac{\sum_{j=1}^{n} x^2_j}{n} - (\bar{X})^2} \qquad (3)$$

4.4.2. Spatial regression analysis: To investigate the spatial relationship between the distribution of malaria hotspots and environmental covariates, we used the regression analysis technique. The mathematical computation applied to the global OLS is given as;

$$\gamma = \beta_0 + \sum_k \beta_k X_{ik} + \varepsilon_i \qquad (a)$$

γ is the dependent or response variable for the observation i (the process to be predicted or understood), for example, the malaria hotspot cases, X_{ik} is the value of the independent or explanatory variable kth for i used to model or predict the dependent variables. The explanatory variables include: population density (X_1), enhanced vegetation index (X_2), rainfall (X_3), aridity (X_4) drought episodes (X_5), and proximity to water (X_6). β_k is the regression coefficients for the variable k -they are values, one for each explanatory variable that represent the strength and type of relationship the explanatory variable has to the dependent variable. β_0 is the regression intercept- it represents the expected value for the dependent variable if all the independent variables are zero. The residuals ε_i represent the potion of the dependent variable that isn't explained by the model[14].

The global OLS method is based on the use of only one equation (a) to explore the relationship between variables. To account for the spatial component of the OLS variables for a specified model, the GWR local regression model creates an equation (b) for each element of the dependent variable data set, in order to capture geographic variations [24]. The GWR model is computed as;

$$\gamma = \beta_0\,(u_i, v_i) + \sum_k \beta_k(u_i, v_i)X_{ik} + \varepsilon_i \qquad (b)$$

where u_i and v_i represent the point coordinates of ith in space. Thus, the GWR equation (b) distinguishes that spatial variations in relationships might exist and provides a way in which they can be measured and this is based mainly in its spatial ability in a GIS environment with emphasis on the differences regarding space and local disaggregation of the local statistics as thoroughly illustrated by Fotheringham [40]. Although the GWR model can be advantageous in order to distinguish heterogeneity of the space itself, it also makes it easier to go from a global OLS perspective to a local analysis, thus obtaining a better grade of details and precision using the GIS environment [40]. Despite the better grade of details with the GWR model, some shortcomings of this technique include: it does not include spatial regime (dummy) variables, it does not include variables with little value variation, and some variables may be exhibiting a local multicollinearity making it hard for the model to be realized [41].

For this study, our regression models (OLS and GWR) was built using the modelling spatial relationship in ArcGIS. Multiple models were made using a set of candidate variables until selecting a model with high explanatory power, and that contained the six explanatory variables (X_1 to X_6) relevant for a properly specified model with respect to malaria transmission [24]. A positive coefficient of an OLS candidate variable means the explanatory **X** variables and the response or dependent **Y** variables changed in the same direction and if the environmental risk factor increases, then the number of confirmed malaria cases will increase. Similarly, a negative coefficient means **X** and **Y** changed in reverse directions [15]. To have a properly specified GWR model, the generated OLS summary result of the model variables should pass a statistically significant six-test check (data pre-processing background analysis).

The GWR model was constructed using a default fixed distance as the kernel type and AICc (Akaike's Information Criterion) as the bandwidth method to find the optimal distance for better model performance. The GWR generates a map that was represented as raster surfaces for the model predictions (combined strength of the relationship amongst the variables used), residuals, local R squared (model significance), and coefficients (explaining the strength of the relationship between the dependent and explanatory variables) and an output table demonstrating the strength of adjusted R^2 significance (from 0- 1).

5. Results

5.1. Hotspot Analysis

The analysis depicted that, there exist high malaria hotspots distribution in most areas (rural, urban, and urban-city centers) of the Northwest, Southwest, Littoral, Yaoundé, and Center regions and low hotspots location elsewhere, as shown in figure 2.

5.2. Ordinary Least Square (OLS)

The OLS global regression illustrates the statistical significance of the model variables and feasibility to be used in specifying the GWR model. The result demonstrated that two of the explanatory variable coefficients (aridity and proximity to water) had a negative relationship (negative sign) with the dependent variable, while all the other explanatory variables had a positive relationship with the dependent variable (table 2). The adjusted R-squared for the OLS performance is 22.3 % (table 3).

5.3. Geographically Weighted Regression (GWR)

In the current study, the GWR produces maps for each coefficient of the spatial association between the dependent and each independent variable. The coefficients (β) of the population density(X_1), enhanced vegetation index (X_2), rainfall (X_3) and drought episodes (X_5) exhibit high (strong) correlation with malaria hotspots in most areas in the western portion of the country and few elsewhere, while aridity (X_4) and proximity to water (X_6) showed a weak association as exemplified in figure 3.
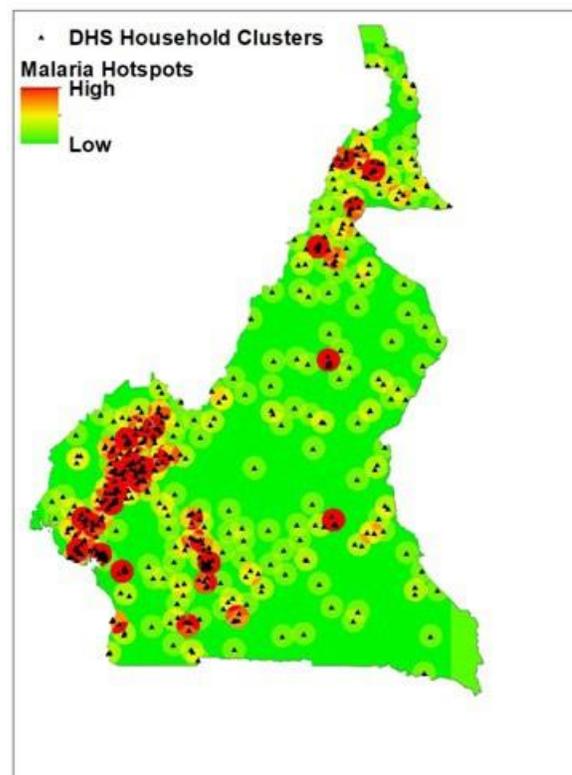

Figure 2: Map of malaria hotspots.

Table 2: Summary of OLS Model Variables

Variable	Coefficient	Std Error	t-Statistic	Probability	Robust SE	Robust t	Robust Pr	VIF
Intercept	-412.757664	126.9804	-3.250562	0.001234*	119.323296	-3.459154	0.000597*	-------
Population density	0.017716	0.006601	2.683856	0.007486*	0.00731	2.423448	0.015671*	1.406656
Aridity	-0.085866	0.009752	-8.804637	0.000000*	0.014935	-5.749357	0.000000*	2.970344
Droughts episodes	0.026051	0.008413	3.096489	0.002066*	0.016161	1.611959	0.107536	1.295128
Enhanced vegetation index	0.108842	0.029597	3.677506	0.000270*	0.067284	1.617653	0.106302	1.586277
Proximity to water	-0.00129	0.000575	-2.24251	0.025297*	0.000409	-3.152627	0.001716*	1.381426
Rainfall	0.746337	0.084883	8.792495	0.000000*	0.135624	5.502997	0.000000*	3.2782

Note: * An asterisk next to a number indicates a statistically significant p-value (p < 0.01) VIF (Variance Inflation Factor) checks for redundancy among explanatory variables.
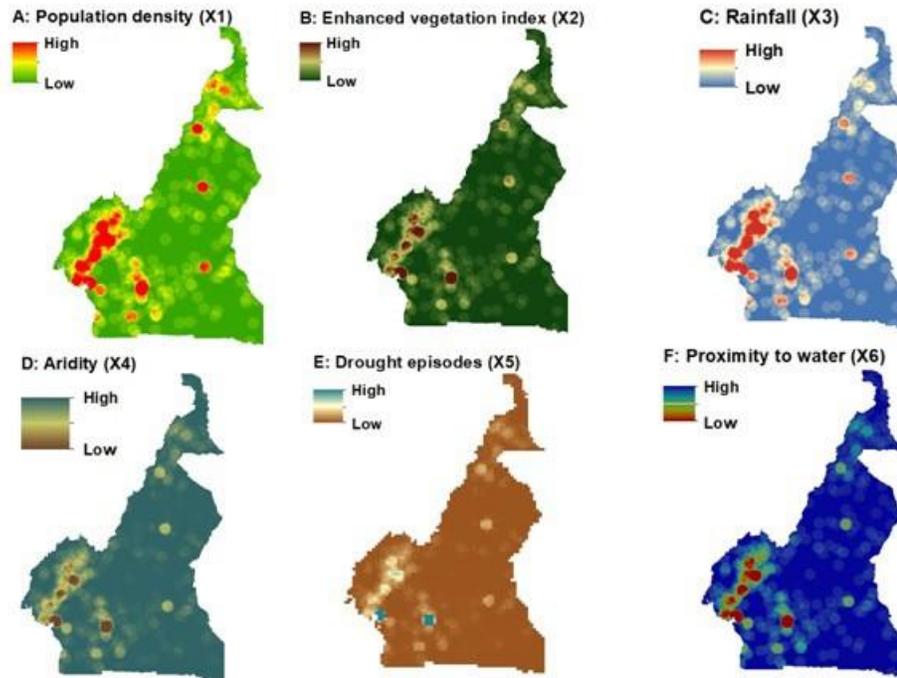
Figure 3: Distribution of GWR model coefficients showing the spatial relationship between malaria hotspot cases and explanatory variables.

The GWR output also produced a predicted malaria map, local R-squared (R^2) and residuals. The local R^2 in figure 4A shows that the model performance was high in regions of the western part of the country and some areas in the Central and Yaoundé DHS regions with a model performance of 0.243 (adjusted R^2 = 24.3%, table 3). The predicted map (the combined strength of the relationship amongst the variables used) showed high spatial association in areas of the Southwest, Northwest and the North regions; figure 4B, while the model residuals (areas not explained by the model) were seen in most parts in the South region; figure 4C.

The differences in model specification between OSL and GWR or how well the model has improved from the global OSL to the local GWR is summarized in table 3.
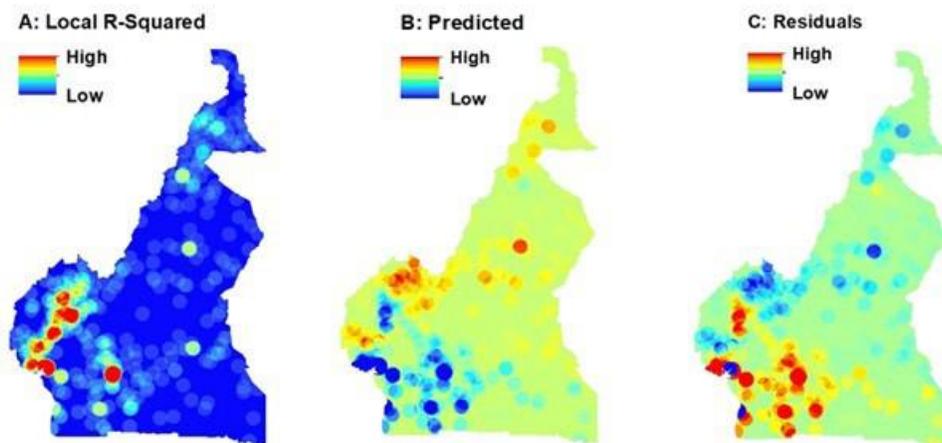
Figure 4: Map of GWR local R2(A), predicted malaria areas (B), and model residuals (C)

Table 3: Comparison between OLS and GWR models

Value	OLS	GWR
AICc*	9400.7	9386.5
Multiple R^2	0.231	0.252
R^2 adjusted	0.223	0.243

6. Discussion

Studies on the association of environmental factors with malaria cases in Cameroon have been illustrated [28,42]. In this study, we demonstrated the ability of the GWR technique as a facilitator tool to understand the spatial interaction of environmental factors with malaria hotspots. Though malaria clusters and hotspots have been reported in Cameroon [28], our analysis confirmed rural and urban-city centers of the Western, Littoral, and Central part of the country as targeted areas for further research and intervention programs.

For the model specification, the OLS model (table 2) illustrated that aridity and proximity to water had negative associations. For example, the negative coefficient for aridity in our model means the malaria cases decrease with lack of water since aridity is a deficiency of moisture probably due to the lack of rainfall while an increase in rainfall in a given rainy season in Cameroon will promote malaria cases and these periods can be targeted for malaria prevention programs since rainfall creates breeding sites for female *Anopheline* mosquitoes. This understanding can promote malaria prevention campaign such as getting rid of stagnant waters around habitable household clusters or discarding water cans to prevent the growth of malaria-causing mosquitoes. Moreover, filling up of potholes during the dry season in high-risk areas will help diminish the mosquito breeding sites. The Koenker statistics (background check-test for non-stationarity) was statistically significant (P< 0.01) and reflects that the relationships being modeled were consistent across the entire study area and thus nonstationary (except for drought episodes and EVI). Furthermore, the VIF values (< 7.5; table 2) indicates no redundancy among explanatory variables and hence no multicollinearity. The OLS model residuals were tested for clustering using the spatial autocorrelation (pre-analysis check) tool in ArcGIS as described elsewhere [43] and it indicated that the variables used were randomly distributed (no clustering). A statistically significant spatial autocorrelation in the model residuals would indicate that we neglected one or more key explanatory variables. This is a positive indicator of choosing a good model [44]. The Jarque-Bera Statistics was statistically significant (p < 0.01) indicating that our model predictions were biased (the residuals were not normally distributed); this may be due to the changing signs in some of the coefficients in the explanatory variable and thus causing variability. Although the test was biased, we proceeded to the GWR model because recent studies [15] have reported on spatial variations similar to our specified model variables and our main goal was to understand the behavior of these environmental variables with malaria cases for future research and intervention projects. The 22.3%

evaluation of the OLS model performance, indicates the explanatory variable coefficients were telling 22% of the spatial relationship story between the malaria cases and the environmental factors we were trying to model. This may seem low per the R^2 range (0-100%) where higher R^2 values depict good model performance. Hasyim [15], in the spatial modelling of malaria cases associated with environmental factors in South Sumatra, Indonesia, had low R^2 (6.2%) variation of malaria incidences by environmental factors for the OLS model. Moise [16] also had low R^2 (< 5%) in their OLS model variables in the seasonal and geographic variation of pediatric malaria in Burundi. This is to say that, the OLS model R^2 varies with the explanatory variables under investigation and lower R^2 does not always signify poor model performance [14].

The local GWR model was built based on the validated variables from the global OLS model. A validated OLS can promote a global policy for malaria control programs where similar statistically significant environmental or climatic predictors could be targeted across many malaria-endemic countries while a validated spatial relationship with GWR is an appropriate method to initiate prevention programs in local systems or village level within endemic countries [15]. The GWR output coefficients maps (figure 3) indicated that population density, EVI, rainfall, and drought episodes had a strong spatial correlation or positive influence on malaria cases in our study locations. The strong correlations were seen in areas of the western part of Cameroon and few in the north. The population of these localities should be considered for malaria control programs in high transmission seasons in Cameroon. Whereas, aridity and proximity to water had a weak association in the above-cited locations meaning there was little or no spatial interdependence between these factors and malaria cases. The spatial variability ability of the GWR model is observed in figure 4A where the generated local R^2 illustrated areas with 24.3% of the model performance between the environmental factors and malaria cases. The high spatial interdependence observed in the Northwest, Southwest, Littoral, Douala, Central and Yaoundé DHS regions calls for effective malaria surveillance. Likewise, active control measures are needed in areas (East, Adamawa and North regions) with low spatial variability. The necessity for the creation of efficient malaria surveillance systems can further be demonstrated from our predicted map (figure 4B) highlighting strong spatial variability in some household clusters of the Northwest, Southwest, Adamawa and North regions for which public health interventions should be prominent.

The GWR residuals (figure 4C) indicated that areas, where the model did not work, were common in the southern part of the country. This implies that our model was unable to explain the spatial variability story between malaria cases and environmental factors in the locations depicted by the GWR residuals. Though our model finds it difficult explaining the spatial interdependence in these areas, the over-

all condition numbers from the attribute table output (background analysis verification) of the GWR indicates the model did not have a hard time solving; since the condition numbers from the explanatory environmental variables were < 30 (>30 would mean the model had difficulties solving the spatial relationships [44]. Furthermore, table 3 illustrated that the GWR improved our understanding of the spatial relationship between malaria cases and environmental factors from the OLS as the AICc and R^2 were lower and higher respectively. This signifies that the GWR is a better indicator for explaining spatial variability at the local level [15,24].

We had the following limitations: Firstly, we used only the malaria hotspot locations to specify our GWR model. This was because running the model on entire malaria cases location for the whole country would have missed key explanatory variables. Moreover, this will help cut down on resources allocation by targeting vulnerable hotspot communities. Secondly, our OLS model was biased and failed to pass the six tests check at the level of the Jarque-Bera statistics that was significant (p< 0.01), indicating our explanatory predictors were not normally distributed in some locations. Thirdly, our model did not include other key environmental variables such as temperature and humidity and socioeconomic predictors for which data were limited or not available. Fourthly, due to the shortcomings of the ArcGIS software in controlling model parameters leading to estimation biases with the GWR technique, our analysis needs to be tested with other clear statistical packages such as the R software. Further investigations for practically available data for other key environmental and socioeconomic predictors are required to provide detailed spatial variability coverage for the malaria disease in Cameroon. Though our study had some remits, it has demonstrated for the first time, a rigorous understanding of the spatial interdependence between malaria cases and environmental risk factors and has provided new insights into the malaria disease in space at the local level in Cameroon by applying the GWR spatial modelling technique that was limited in most studies. Moreover, the methods in this study can be used to study other health outcomes in Cameroon that have been applied in available literature in other countries [17,22,24].

7. Conclusions

This study demonstrated that rainfall, EVI, drought episodes, and population density had a strong spatial association with malaria distribution and could be targeted as important risk predictors for control programs at the local level in Cameroon. Given the greater availability of spatial data [31] and desktop GIS packages and statistical techniques, the challenges faced in the malaria disease investigation will be improved in the future. The generated GWR maps suggest that for policymakers to eliminate malaria by 2030, there should be the creation of outreach programs that will target malaria hotspot locations and carry out further investigations in areas where the environmental variables showed strong spatial associations with malaria cases.

References

1. WHO; High burden to high impact: A targeted malaria response. 2019.

2. WHO Cameroon. 2018.

3. Ministry of Public Health. Sectorial health strategic plan 2001–2015. 2019.

4. Institut National de la Statistique - INS/Cameroun and ICF. International. Enquête Démographique et de Santé et à Indicateurs Multiples du Cameroun 2011. 2019.

5. World Health Organization, Global Malaria Programme. Global technical strategy for malaria, 2016-2030, 2015 [accessed 19 March 2019]

6. Bhatt S, Weiss DJ, Cameron E, Bisanzio D, Mappin B, Dalrymple U, et al. The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. Nature. 2015; 526: 207–11.

7. Gemperli A, Sogoba N, Fondjo E, Mabaso M, Bagayoko M, Briet OJT, et al. Mapping malaria transmission in West and Central Africa. Trop Med Int Health. 2006; 11:1032–46.

8. Gething PW, Casey DC, Weiss DJ, Bisanzio D, Bhatt S, Cameron E, et al. Mapping Plasmodium falciparum Mortality in Africa between 1990 and 2015. N Engl J Med. 2016; 375: 2435–45.

9. Omumbo JA, Noor AM, Fall IS, Snow RW. How Well Are Malaria Maps Used to Design and Finance Malaria Control in Africa? PLoS One. 2013; 8: e53198.

10. Booman M, Durrheim DN, Grange KL, Martin C, Mabuza AM, Zitha A, et al. Using a geographical information system to plan a malaria control programme in South Africa. Bull World Health Organ. 2000; 78: 1438-44.

11. Dalrymple U, Mappin B, Gething PW. Malaria mapping: understanding the global endemicity of falciparum and vivax malaria. BMC Med. 2015; 13:140.

12. Moukam Kakmeni FM, Guimapi RYA, Ndjomatchoua FT, Pedro SA, Mutunga J, Tonnang HEZ. Spatial panorama of malaria prevalence in Africa under climate change and interventions scenarios. Int J Health Geogr. 2018; 17: 2.

13. Efe SI, Ojoh CO. Climate Variation and Malaria Prevalence in Warri Metropolis. ACS. 2013; 3: 132–40.

14. Regression analysis basics. [accessed 13 April2019]

15. Hasyim H, Nursafingi A, Haque U, Montag D, Groneberg DA, Dhimal M, et al. Spatial modelling of malaria cases associated with environmental factors in South Sumatra, Indonesia. Malar J. 2018; 17: 87.

16. Moise I, Roy S, Nkengurutse D, Ndikubagenzi J. Seasonal and Geographic Variation of Pediatric Malaria in Burundi: 2011 to 2012. Int J Environ Res Public Health. 2016; 13: 425.

17. Rusk A, Highfield L, Wilkerson JM, Harrell M, Obala A, Amick B. Geographically-weighted regression of knowledge and behaviour determinants to anti-malarial recommending and dispensing practice among medicine retailers in western Kenya: capacitating targeted interventions. Malar J. 2016; 15: 562.

18. Ehlkes L, Krefis A, Kreuels B, Krumkamp R, Adjei O, Ayim-Akonor M, et al. Geographically weighted regression of land cover determi-

nants of Plasmodium falciparum transmission in the Ashanti Region of Ghana. Int J Health Geogr. 2014; 13: 35.

19. Delmelle E, Hagenlocher M, Kienberger S, Casas I. A spatial model of socioeconomic and environmental determinants of dengue fever in Cali, Colombia. Acta Trop. 2016; 164: 169–76.

20. Ren H, Zheng L, Li Q, Yuan W, Lu L. Exploring Determinants of Spatial Variations in the Dengue Fever Epidemic Using Geographically Weighted Regression Model: A Case Study in the Joint Guangzhou-Foshan Area, China, 2014. Int J Environ Res Public Health. 2017; 14: 1518.

21. Goovaerts P, Xiao H, Adunlin G, Ali A, Tan F, Gwede CK, et al. Geographically-weighted regression analysis of percentage of late-stage prostate cancer diagnosis in Florida. Appl Geogr. 2015; 62: 191-200.

22. Choi H, Kim H. Analysis of the relationship between community characteristics and depression using geographically weighted regression. Epidemiol Health. 2017; 39: e2017025.

23. Song C, Kwan M-P, Zhu J. Modeling Fire Occurrence at the City Scale: A Comparison between Geographically Weighted Regression and Global Linear Regression. Int J Environ Res Public Health. 2017; 14: 396.

24. Martínez Bascuñán M, Rojas Quezada C. Geographically weighted regression for modelling the accessibility to the public hospital network in Concepción Metropolitan Area, Chile. Geospat Health. 2016; 11: 451.

25. Michael P. Cameron, William Cochrane, Craig Gordon, Michael Livingston. Alcohol outlet density and violence: A geographically weighted regression approach. Drug Alcohol Rev. 2016; 35: 280-88

26. Legg R, Bowe T. Applying Geographically Weighted Regression to a Real Estate Problem. 2019.

27. Massoda Tonye SG, Kouambeng C, Wounang R, Vounatsou P. Challenges of DHS and MIS to capture the entire pattern of malaria parasite risk and intervention effects in countries with different ecological zones: the case of Cameroon. Malar J. 2018; 17: 156.

28. Tewara MA, Mbah-Fongkimeh PN, Dayimu A, Kang F, Xue F. Small-area spatial statistical analysis of malaria clusters and hotspots in Cameroon;2000–2015. BMC Infect Dis. 2018; 18:636

29. Puth M-T, Neuhäuser M, Ruxton GD. Effective use of Pearson's product–moment correlation coefficient. Anim Behav. 2014; 93: 183–9.

30. The DHS Program - Quality information to plan, monitor and improve population, health, and nutrition programs. [accessed 15 April 2018]

31. Spatial Data Repository, The Demographic and Health Surveys Program. ICF International.dhsprogram.com [accessed 25 March 2019]

32. Center for International Earth Science Information Network - CIESIN - Columbia University. 2016. Gridded Population of the World, Version 4 (GPWv4): Population Count Adjusted to Match 2015 Revision of UN WPP Country Totals. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). [accessed 30 April 2019]

33. CGIAR-CSI. 2009. "Global Aridity – Annual." [accessed 30 April 2019]

34. Center for Hazards and Risk Research - Columbia University, Center for International Earth Science Information Network - Columbia University, and International Research Institute for Climate and Society - Columbia University. 2005. "Global Drought Hazard Frequency and Distribution." [accessed 30 April 2019]

35. Kamel Didan. 2016. "NASA MEaSUREs Vegetation Index and Phenology (VIP) Phenology EVI2 Yearly Global 0.05Deg CMG" [accessed 30 April 2019]

36. Wessel P, Smith WHF. A global, self-consistent, hierarchical, high-resolution shoreline database. J Geophys 1996; 101: 8741–3.

37. Wessel Paul, Walter Smith. "A Global Self-consistent, Hierarchical, High-resolution Geography Database Version 2.3.7." 2017.

38. Climate Hazards Group. "Climate Hazards Group InfraRed Precipitation with Station data 2.0." 2017.

39. Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, et al. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci Data. 2015; 2: 150066.

40. Fotheringham AS, Brunsdon C, Charlton M. Geographically weighted regression: the analysis of spatially varying relationships. Wiley, Hoboken, NJ, USA; 2002.

41. Wheeler D, Tiefelsdorf M. Multicollinearity and correlation among local regression coefficients in geographically weighted regression. J Geograph Syst. 2005; 7: 161–87.

42. Tatem AJ, Adamo S, Bharti N, Burgert CR, Castro M, Dorelien A, et al. Mapping populations at risk: improving spatial demographic data for infectious disease modeling and metric derivation. Popul Health Metrics 2012; 10: 8.

43. How Spatial Autocorrelation (Global Moran's I) works—Help | ArcGIS Desktop. [accessed 26 March 2019]

44. Rosenshein L, Scott L, Pratt M. Finding a Meaningful Model. [2019].