Journal of Geographic Information System, 2023, 15, 110-121
https://www.scirp.org/journal/jgis
ISSN Online: 2151-1969
ISSN Print: 2151-1950

Geospatial Coronavirus Vulnerability Regression Modelling for Malawi Based on Cumulative Spatial Data from April 2020 to May 2021

Emmanuel Chinkaka[1,2*], Kyle F. Davis[1], Dawnwell Chiwanda[2], Billy Kachingwe[2], Stanley Gusala[2], Richard Mvula[2], Francis Chauluka[3], Julie Michelle Klinger[1]

[1]Department of Geography and Spatial Sciences, University of Delaware, Newark, USA
[2]Department of Earth Sciences, Malawi University of Science and Technology, Limbe, Malawi
[3]Department of Water Resources Management, Malawi University of Science and Technology, Limbe, Malawi

Email: *echinka@udel.edu

How to cite this paper: Chinkaka, E., Davis, K.F., Chiwanda, D., Kachingwe, B., Gusala, S., Mvula, R., Chauluka, F. and Klinger, J.M. (2023) Geospatial Coronavirus Vulnerability Regression Modelling for Malawi Based on Cumulative Spatial Data from April 2020 to May 2021. *Journal of Geographic Information System*, **15**, 110-121.
https://doi.org/10.4236/jgis.2023.151007

Received: January 22, 2023
Accepted: February 20, 2023
Published: February 23, 2023

Abstract

In the past two to three years, the world has been heavily affected by the infectious coronavirus disease and Malawi has not been spared due to its interconnection with neighboring countries. There is no management tool to identify and model the vulnerabilities of Malawi's districts in prioritizing health services as far as coronavirus prevalence and other infectious diseases are concerned. The aim of this study was to model coronavirus vulnerability in all districts in Malawi using Geographic Information System (GIS) to monitor the disease's cumulative prevalence over the severely affected period between 2020 and 2021. To achieve this, four parameters associated with coronavirus prevalence, including population density, percentage of older people, temperature, and humidity, were prepared in a GIS environment and used in the modelling process. A multiscale geographically weighted regression (MGWR) model was used to model and determine the vulnerability of coronavirus in Malawi. In the MGWR modelling, the Fixed Spatial Kernel was used following a Gaussian distribution model type. The Results indicated that population density and older people (age greater than 60 years) have a more significant impact on coronavirus prevalence in Malawi. The modelling further shows that Malawi, between April 2020 and May 2021, Lilongwe, Blantyre and Thyolo were more vulnerable to coronavirus than other districts. This research has shown that spatial variability of Covid-19 cases using MGWR has the potential of providing useful insights to policymakers for targeted interventions that could otherwise not be possible to detect using non-geovisualization techniques.

Keywords

Malawi, Geospatial, Spatial Dependency, Coronavirus, Vulnerability, Spatial Variability, Prevalence, MGWR, GIS

1. Introduction

Coronavirus disease (Covid-19) is a global health concern due to its rapid spread [1]. It is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and was declared as a public health emergency of international concern on 30th January 2020 and a global pandemic on 11th March 2020 by the World Health Organization [2]. According to Zhu *et al.*, [3] the coronavirus outbreak began on 12th January 2020, and rapidly became a sudden public health crisis which put a threat to lives in almost all parts of the world. Shortly after the initial cases in China, Iran and some European countries detected the coronavirus disease including Italy as the most hit country experienced a significant increase in the number of cases and deaths [2]. Covid-19 commonly targets the human respiratory system, and in its severe case, results in fatal lung injury and consequently loss of life [4]. Malawi registered her first Covid-19 cases on the 2nd of April 2020, and the Government of Malawi through the Ministry of Health declared a state of emergency in the country [5]. On 8th April 2020, the Government of Malawi officially launched the Covid-19 national preparedness and response plan with the aim of controlling and stopping further spreading of the virus. UNICEF Malawi, with funding from UK Aid, ventured into supporting the Ministry of Health to train about 80 health workers on Covid-19 case management in the most affected districts within Malawi [5].

Wu *et al.* [6] argue that the spread of Covid-19 is a function of so many factors including air pollution, smoking and environmental conditions. Wu *et al.*, [6] further indicate that long-term exposure to polluted air has a great potential of exacerbating the health outcomes of Covid-19 cases. Using Covid-19 cases in China, Wu *et al.* [6] indicated that environmental conditions such as humidity and temperature have a great influence on the transmission of Covid-19 when compared to other respiratory viruses. These conditions are important natural factors that control transmission of most infectious diseases [3]. Previous studies have shown that the transmission of influenza such as Covid-19 is seasonal and mostly affected by temperature and humidity variations [7]. According to [8], the transmissions of most other human coronaviruses that result in mild respiratory symptoms, such as OC43 (HCoV-OC43) and HCoV-HKU1 spread rapidly based on seasonal variations. The seasonality of such viruses has been modelled to determine long-term simulation of the transmission of SARS-CoV-2 [9]. The geospatial modelling technique is an essential tool in examining the spatial distribution of most infectious diseases [10]. Geographic Information System (GIS) has become a vital assessment tool in analyzing and geo-visualization of the spread of Covid-19. For example, the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE) since the on-set of Covid-19, utilizes a Web-based GIS dashboard that provides real-time data of the worldwide spatial distribution of Covid-19 [11]. This geo-visualization includes the total number of confirmed cases, mortalities, and recovered patients. A GIS-based representation of coronavirus prevalence reveals geographic patterns that show hotspots that require interventions to prevent further spread of the disease. The Malawian government's health system was overwhelmed and stretched to its limit with a huge number of coronavirus cases and hospitalized patients [12]. Currently, there is no management tool to monitor and model vulnerability of Malawi's districts in prioritizing health services as far as coronavirus prevalence and other infectious diseases are concerned. Thus, this work provides a more effective and easy tool for developing coronavirus vulnerability while considering climatic and social factors that influence its spread. Therefore, the aim of this study was to use spatial data and a Multi-scale Geographic Weighted Regression algorithm to model the vulnerability of corona virus disease in Malawi based on cumulative Covid-19 cases' data from April 2020 to May 2021. Using the available and influential data on coronavirus such as population density, percentage of older people, temperature, and humidity, we propose a coronavirus vulnerability model for Malawi.

2. Materials and Methods

2.1. Study Area

Malawi is a country located in the southeastern Africa with an area that spans over 118,484 Km^2 with a population of about 18 million according to the 2018 National Census [13]. Malawi is bordered by Zambia, the United Republic of Tanzania and Mozambique. Malawi is located at latitude 13.25˚S and longitude 34.30˚E of the Greenwich Mean Time and has a total of 28 administrative districts as shown in **Figure 1**.

2.2. Data Collection and Preprocessing

The Public Health Institute of Malawi (PHIM) continues to monitor and publish district level data of the Coronavirus disease on a daily basis across Malawi. In this study, we used a three-fold steps approach to develop the coronavirus vulnerability model for Malawi. First of all, cumulative data on the number of COV-SAR 2 patients from April 2020 to May 2021 was collected from Public Health Institute of Malawi website https://covid19.health.gov.mw/ [12]. In the second step, based on recent work by [8] and [14], four coronavirus influencing factors shown in **Table 1**, including population density, temperature, humidity, and number of older people (age above 60) were prepared and visualized in ArcGIS Pro 2.8. In the third and final step, geospatial analysis and MGWR model were run to develop the coronavirus vulnerability model from April 2020 to December 2021. A comprehensive research summary is shown in **Figure 2**.

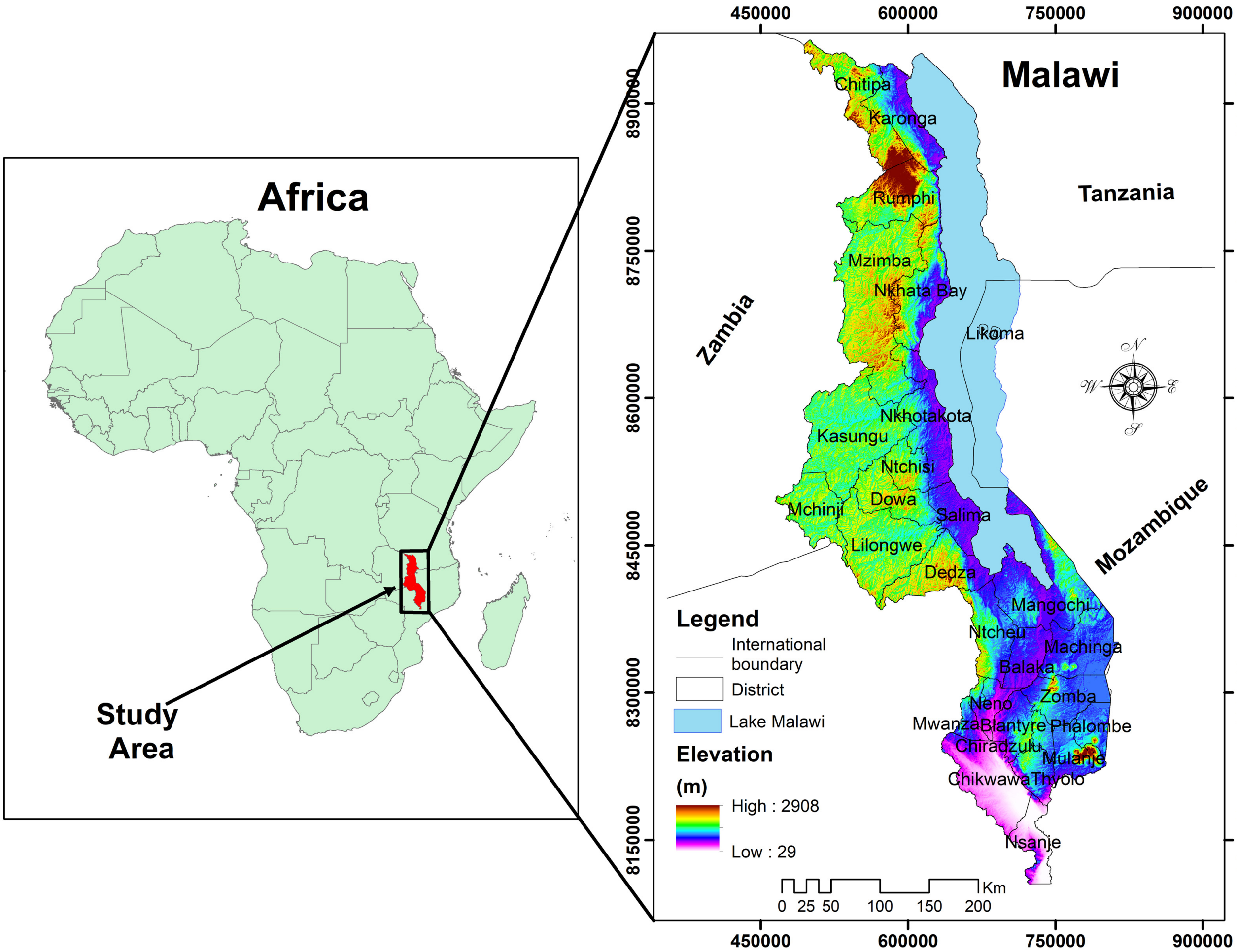

Figure 1. Location of the study area.

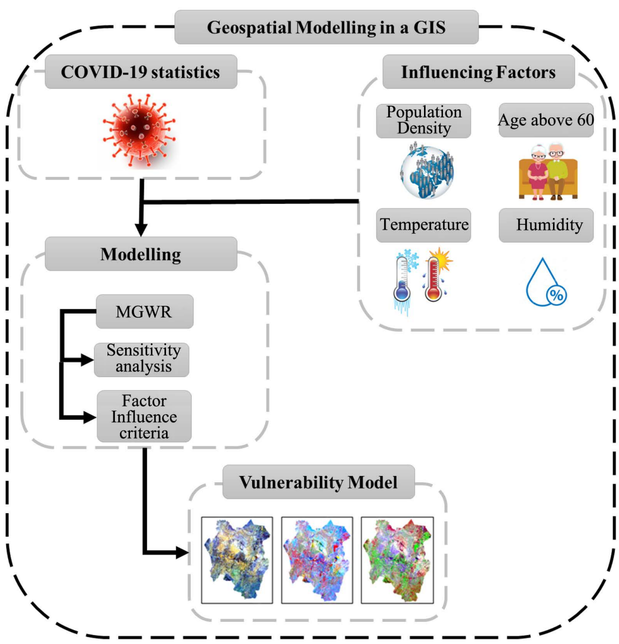

Figure 2. Research modelling methodology.

Table 1. COV-SAR 2 influencing factors and variables used in this study.

Theme	Variable used	Description	Source
Population	Population density	Proportion to the number of people living in a specific district	Malawi National Statistical Office [13]
Number of Elderly	Percentage of Elderly	Proportion of the elderly people age greater than 60 living in each district	Malawi National Statistical Office [13]
Humidity	Minimum, maximum and average humidity	Humidity values recorded in each district	Department of Climate Change and Meteorological Services (DCCMS) (www.metmalawi.gov.mw)
Temperature	Minimum, maximum and average temperature	Temperature values recorded in each district	Department of Climate Change and Meteorological Services (DCCMS) (www.metmalawi.gov.mw)

2.3. Malawi's Coronavirus Statistics

Based on PHIM reports, from April 2, 2020 to May 31, 2021, approximately 34,099 people were infected with the coronavirus in Malawi as cumulative statistics for all the districts, with most cases occurring in Blantyre (10,832 patients), Lilongwe (7881) and Mzimba (2785) [12]. Out of these cases, there were 1148 deaths countrywide. **Figure 3(a)** shows the spatial distribution of the number of confirmed cumulative coronavirus cases and **Figure 3(b)** shows the daily confirmed from 2nd April 2020 to 31st May 2021.

2.4. Common Influencing Factors of Coronavirus Disease

Coronavirus transmission and prevalence are a function of several factors. However, this research is based on the influencing factors as reported by WHO. Four factors, including, population density, number of elderly people (age > 60 years), humidity and temperature were identified as factors influencing coronavirus disease [15]. Most influenza diseases such as coronavirus have high prevalence and transmission due to increased temperature and humidity [14].

Temperature and humidity data records from the Malawi's Department of Climate Change and Meteorological Services (DCCMS) for the period of April 1, 2020 to May 31, 2021 were preprocessed and used in the vulnerability modeling. The average recordings from DCCMS stations were computed from all the districts and used in the MGWR modelling to determine the vulnerability of coronavirus infection in Malawi. According to Public Health Institute of Malawi and WHO [2] [12], high coronavirus cases and deaths occur among the elderly people and prevalence is mostly related to the population densities within the countries. As such, this research utilized population density and number of elderly people in each district based on the report by National Statistical Office 2018 census [13]. The GIS data for the four coronavirus influencing factors were created in ArcGIS Pro 2.8 for modeling (**Figure 4**).

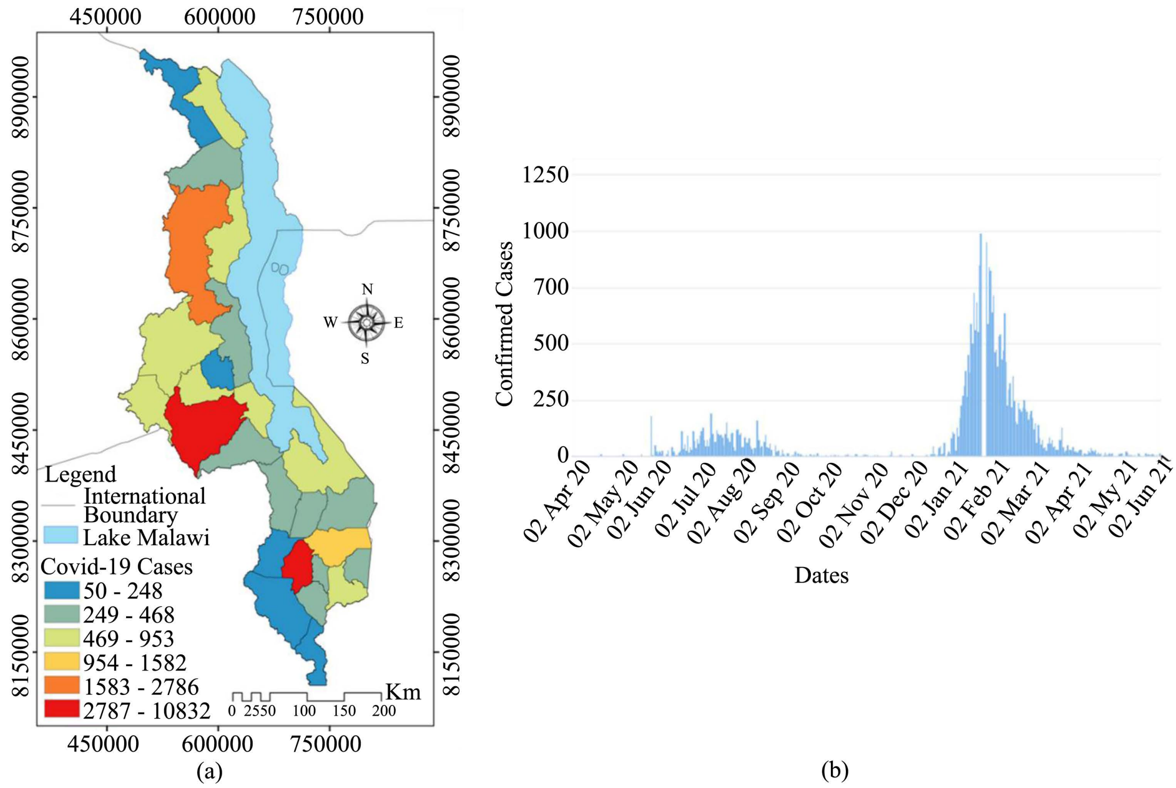

Figure 3. Spatial distribution of Covid-19 cases per district (a) Confirmed cases by date (b) (April 2021-May 2021).

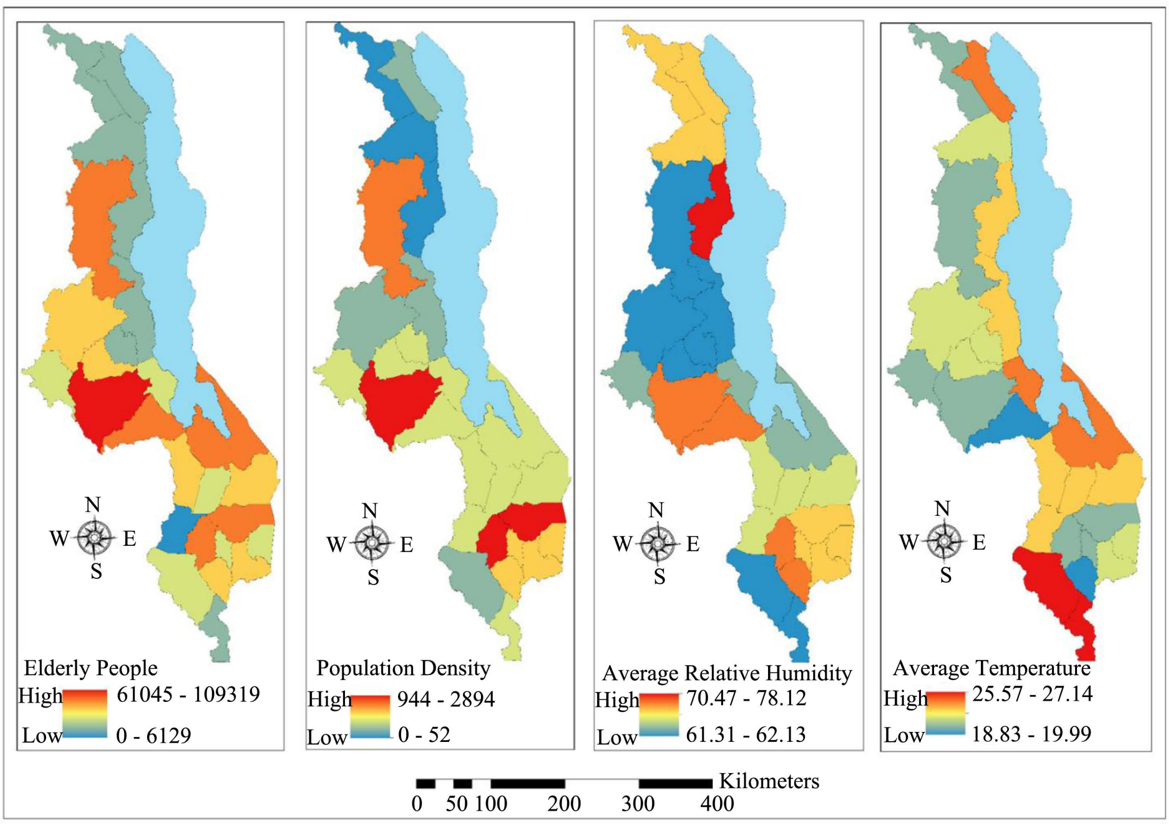

Figure 4. Coronavirus influencing factors.

2.5. MGWR Model

The Multi-scale Geographic Weighted Regression (MGWR) model is an extension that is built and improved from Geographic Weighted Regression model (GWR) that is used in modelling to achieve higher accuracy in analyzing location-based relationships of spatially referenced data [16]. The GWR model functions on the basis that the model parameters are estimated at any point in a study area location. The Fixed Spatial Kernel with a Gaussian distribution was used. It assigns more relative weights to closer observations and less to those that are more distant, following Tobler first law of geography [17]. However, the GWR model has the limitation of assuming that all relations in the analysis have a constant scale. This usually affects the modelling process and that is why this research applied the MGWR model in order to rectify this problem in the modelling of coronavirus vulnerability since MGWR takes into consideration variations of spatial scales of the relations [18]. It is considered as a type of regression model with geographically varying parameters [19].

Normally, the MGWR model has one variable as a dependent variable and one or more others as independent variables. Equation (1) shows this relationship.

$$y_i(\mu)=\beta_{0i}(\mu)+\sum_{j=1}^{m}\beta_{bwj}(\mu)x_{ij} \tag{1}$$

where y is the dependent variable at any point u, x is the independent variable at any point u, bwj stands for the bandwidth that has been used in the model calibration of the conditional relationship of jth in βbwj, and β represents model estimator. Equation (2) shows how this is calculated.

$$\hat{\beta}(\mu)=\left(x^T w(\mu)x\right)^{-1}x^T w(\mu)y \tag{2}$$

$x^T w(\mu)x$ is the weighted covariance matrix with a spatial component, $w(\mu)$ represents the square matrix of weights, and these weights are assigned values depending on any u point position in the study area Finally y denotes the value of the dependent variable at any point u. In this research, the coronavirus vulnerability modelling was developed using a python implemented code by [19] in MGWR 2.2 (https://sgsup.asu.edu/sparc/mgwr). The number of coronavirus cases in each district was used as dependent data, and the values of the four influencing factors: population density, number elderly people, humidity and temperature for each district were used as independent variables.

3. Results and Discussion

Coronavirus disease statistics in each of the Malawi's 28 districts and the four influencing factors which included population density, number of elderly people, humidity and temperature were used to model the vulnerability to coronavirus prevalence in Malawi using the MGWR model. **Table 2** shows summary statistics of the coronavirus disease modelling using these four influencing factors.

According to the results shown in **Table 2**, the impact of the MWGR in allowing spatial dependency of the influencing factors improves the performance

Table 2. Summary statistics for MGWR Model influencing factor estimates.

Variable	Coefficient	Standard Error	z-Score	p-Value
Population density	0.782	0.125	6.274	0.000
Number of elderly people	0.175	0.129	1.352	0.176
Humidity	−0.039	0.311	−0.126	0.900
Temperature	−0.012	0.299	−0.038	0.969

of spatial modelling of Covid-19 incidence rate in Malawi. This is very critical in true reflection modelling by incorporating the spatial scale of the independent factors in the spatial modelling. The coefficients of the influencing factors population density and number of elderly people were found to be strongly significant with positive values. But population density showed a stronger influence on the modelling with 0.782 coefficient (**Table 2**). This indicates that the prevalence of coronavirus in Malawi from April 2020 to May 2021 depended on the population density of the district which could influence more interaction of people within the district result in further spread of the virus. The number of elderly people came second as an influencing factor with coefficient value of 0.175. The results in **Figure 5** show that the coronavirus related deaths increased among people with the age greater than 60 years. This is normally the case since most elderly people have weak immune system due to underlying old-age health related conditions [15]. According to [20], coronavirus infection was observed to be three times more common in the elderly due to their weakened immune system. Thus, the prevalence of the disease among them is high.

The results in this study shows temperature and humidity as minor influencing factors in the prevalence of the disease within the study period. However, based on raw data in **Figure 3(b)**, it is still observed that the prevalence of coronavirus in the country was minimal in the summer from September 2020 to December 2020 and high cases were observed in winter season from June 2020 to August 2020 and also in the rainy season between January 2021 to March 2021. This analysis indicates that high temperature and high humidity significantly reduce the prevalence of coronavirus in Malawi data. This pattern of coronavirus prevalence corresponds with what [21] observed using China coronavirus data, that during the summer season, coronavirus cases dropped drastically in China and [14] for most common influenza diseases. As such, seasonality in coronavirus prevalence is certain Malawi. **Table 3** indicates the measure of goodness-of-fit in the MGWR in modelling prevalence of Covid-19 incidence rate in Malawi.

The adjusted R^2 of 0.746 meaning that the model used in this study could explain 74.6% of the total variations of Covid-19 incidence rates among the districts. This indicates a significant best fit in the modelling. This is due to the spatial variability scale that MGWR allows in the modeling process [19]. This best-fit modelling is also evident with the lower AICc value of 53.915, which still supports the best modelling result of the Covid-19 incidence rate. Coronavirus vulnerability model in Malawi's districts is shown in **Figure 6**. According to the

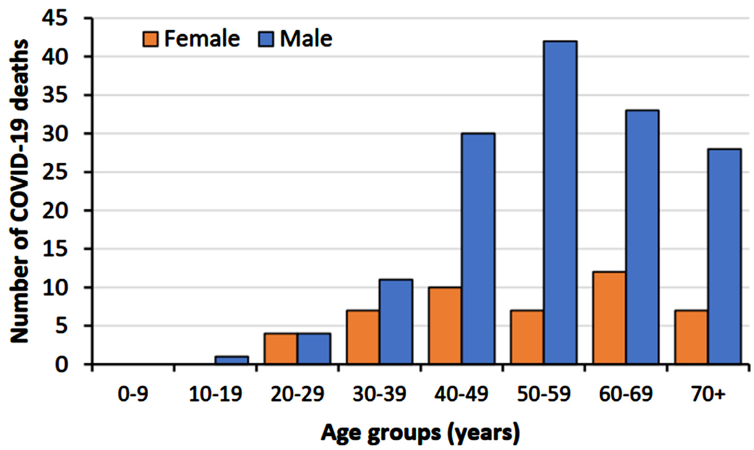

Figure 5. Age and sex distribution of Covid-19 related deaths in Malawi between April 2020 and June 2021 [12].

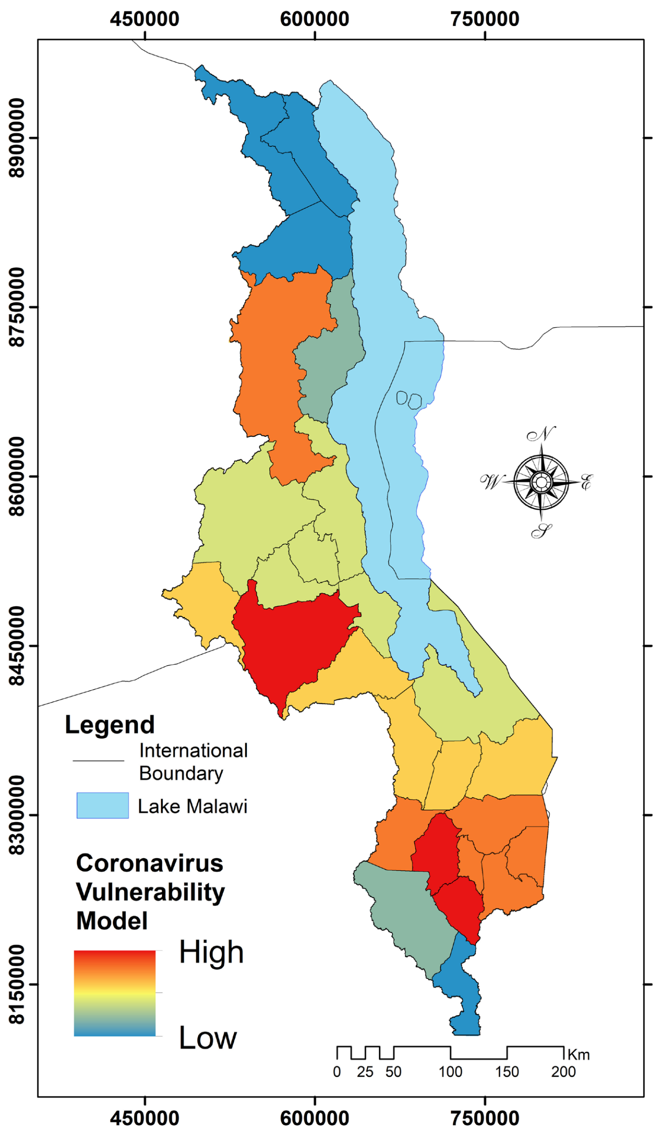

Figure 6. Coronavirus vulnerability model using MGWR.

Table 3. Measure of goodness-of-fit for MGWR in modeling Covid-19 incidence rate.

AIC	48.097
AICc	53.915
R^2	0.782
Adjusted R^2	0.746

MGWR model results, Lilongwe, Blantyre and Thyolo districts are more vulnerable than other parts of Malawi. These are followed by districts such as Mzimba, Zomba, Mwanza, Neno, Mulanje, Chiradzuru and Phalombe. Most of these districts have this high prevalence of coronavirus disease due to high population density and higher number of the elderly, contributing to vulnerability.

4. Conclusion

The rapid spread of coronavirus disease throughout the world, not spearing Malawi as a country, has resulted in a significant public health crisis and is a global concern. In this worse situation of a health crisis due to the Covid-19 pandemic, more researches are required to analyze spatial location impacts of Covid-19. It is necessary to seriously consider management of Covid-19 cases and indeed plan to address and reduce its further spread. In order to achieve this, we developed a Covid-19 vulnerability model of Malawi by using four influencing factors: population density, number of older people, humidity and temperature using MGWR in a GIS. Based on the MGWR model results from this study, population density and a number of elderly people have greatly influenced coronavirus prevalence in Malawi. It shows that between April 2020 and May 2021, Lilongwe, Blantyre and Thyolo were more vulnerable to coronavirus than other districts in Malawi followed by districts like Mzimba, Zomba, Mwanza, Neno, Mulanje, Chiradzuru and Phalombe. Based on these results, these districts are referred to as Malawi's Covid-19 hotspots where health officials need to prepare an action plan immediately to react, respond and recover the affected groups and make strategies for such vulnerable areas. Therefore, measures that would reduce the crowding of people such as avoiding travel and attendance of public functions in such districts could help reduce the spread of the disease. The spatial locations of the hotspots in this research study will help the Government, district administrative, and local authorities in these districts, including the residents of vulnerable districts to take intervention actions and precautions against this novel infectious disease. Using our approach, the Ministry of Health and other decision-makers in health sector can identify high-risk districts (in case a future variant emerges) and provide the necessary interventions to reduce the spread of the disease.

Acknowledgements

The authors would like to thank the Malawi University of Science and Technology (MUST), the Department of Climate Change and Meteorological Services (DCCMS) for providing the climatic data and Public Health Institute of Malawi (PHIM) for making available cumulative coronavirus statistics used in this study. The open-source python package for mgwr implemented in this work can be accessed at https://github.com/pysal/mgwr on GitHub.

Author Contribution

Emmanuel Chinkaka: Conceptualization, Data analysis, post-processing analysis, Drafting, Map Design, **Dawnwell Chiwanda**: Data Pre-processing, **Stanley Gusala**: Data Pre-processing, **Billy Kachingwe**: Data Pre-processing, **Francis Chauluka**: Reviewing, Editing, **Richard Mvula**: Editing, Review, **Julie Michelle Klinger**: Augmentation, Review, **Kyle F. Davis**: Augmentation, Review.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Parvin, F., Ali, S.A., Hashmi, S. and Ahmad, A. (2021) Spatial Prediction and Mapping of the Covid-19 Hotspot in India Using Geostatistical Technique. *Spatial Information Research*, **29**, 479-494. https://doi.org/10.1007/s41324-020-00375-1

[2] World Health Organization (2020) Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (Covid-19). 16-24.

[3] Zhu, Y., Chen, L., Ji, H., Xi, M. and Fang, Y. (2020) The Risk and Prevention of Novel Coronavirus Pneumonia Infections among Inpatients in Psychiatric Hospitals. *Neuroscience Bulletin*, **36**, 299-302. https://doi.org/10.1007/s12264-020-00476-9

[4] Rohleder, S. and Bozorgmehr, K. (2021) Monitoring the Spatiotemporal Epidemiology of Covid-19 Incidence and Mortality: A Small-Area Analysis in Germany. *Spatial and Spatio-Temporal Epidemiology*, **38**, Article ID: 100433. https://doi.org/10.1016/j.sste.2021.100433

[5] UNICEF Malawi (2021) UNICEF Malawi Covid-19 Situation Report 3 February. https://www.unicef.org/malawi/reports/unicef-malawi-covid-19-situation-report-3-february

[6] Wu, X., Nethery, R.C., Sabath, M.B., Braun, D. and Dominici, F. (2020) Exposure to Air Pollution and Covid-19 Mortality in the United States: A Nationwide Cross-Sectional Study. https://doi.org/10.1101/2020.04.05.20054502

[7] Wallinga, J. and Teunis, P. (2004) Different Epidemic Curves for Severe Acute Respiratory Syndrome Reveal Similar Impacts of Control Measures. *American Journal of Epidemiology*, **160**, 509-516. https://doi.org/10.1093/aje/kwh255

[8] Krishnamoorthy, S., Swain, B., Verma, R. and Gunthe, S. (2020) SARS-CoV, MERS-CoV, and 2019-nCoV Viruses: An Overview of Origin, Evolution, and Genetic Variations. *VirusDisease*, **31**, 411-423. https://doi.org/10.1007/s13337-020-00632-9

[9] Ayanshina, O., *et al.* (2020) Seasonal Variations in Nigeria: Understanding Covid-19 Transmission Dynamics and Immune Responses. *Journal of Global Health Reports*, **4**, e2020084. https://doi.org/10.29392/001c.14600

[10] Mollalo, A., Sadeghian, A., Israel, G.D., Rashidi, P., Sofizadeh, A. and Glass, G.E. (2018) Machine Learning Approaches in GIS-Based Ecological Modeling of the Sand Fly *Phlebotomus papatasi*, a Vector of Zoonotic Cutaneous Leishmaniasis in Golestan Province, Iran. *Acta Tropica*, **188**, 187-194. https://doi.org/10.1016/j.actatropica.2018.09.004

[11] JHU CSSE (2020) Johns Hopkins University Center for Systems Science and Engineering. https://www.arcgis.com/apps/dashboards/bda7594740fd40299423467b48e9ecf6

[12] PHIM (2021) Public Health Institute of Malawi Cumulative Data: Covid-19 Situation Report. 1-5.

[13] NSO (2019) 2018 Malawi Population and Housing Main Report, National Statistical Office.

[14] Zhang, Y., *et al.* (2015) The Impact of Temperature and Humidity Measures on Influenza A (H7N9) Outbreaks—Evidence from China. *International Journal of Infectious Diseases*, **30**, 122-124. https://doi.org/10.1016/j.ijid.2014.11.010

[15] Jin, Y., *et al.* (2020) A Rapid Advice Guideline for the Diagnosis and Treatment of 2019 Novel Coronavirus (2019-nCoV) Infected Pneumonia (Standard Version). *Military Medical Research*, **7**, Article No. 4.

[16] Yu, H., Fotheringham, A., Li, Z., Oshan, T., Kang, W. and Wolf, L. (2019) Inference in Multiscale Geographically Weighted Regression. *Geographical Analysis*, **52**, 87-106. https://doi.org/10.31219/osf.io/4dksb

[17] Waters, N. (2017) Tobler's First Law of Geography. https://doi.org/10.1002/9781118786352.wbieg1011

[18] Yu, J., *et al.* (2021) Superpixel Segmentations for Thin Sections: Evaluation of Methods to Enable the Generation of Machine Learning Training Data Sets. https://doi.org/10.31223/X55S65

[19] Oshan, T.M., Li, Z., Kang, W., Wolf, L.J. and Stewart Fotheringham, A. (2019) MGWR: A Python Implementation of Multiscale Geographically Weighted Regression for Investigating Process Spatial Heterogeneity and Scale. *ISPRS International Journal of Geo-Information*, **8**, 269. https://doi.org/10.3390/ijgi8060269

[20] Lin, J., *et al.* (2020) Containing the Spread of Coronavirus Disease 2019 (Covid-19): Meteorological Factors and Control Strategies. *Science of the Total Environment*, **744**, Article ID: 140935. https://doi.org/10.1016/j.scitotenv.2020.140935

[21] Wang, J., Tang, K., Feng, K. and Lv, W. (2020) High Temperature and High Humidity Reduce the Transmission of Covid-19.