

Covid-19 in West & East Africa, a Geographical Weighted Regression Exploration with <http://mygeoffice.org/>

Joao Negreiros, Samia Loucif, Mohammed Amin Kuhail, Ahmed Seffah

College of Technological Innovation, Zayed University, Abu Dhabi, UAE

Email: joao.negreiros@zu.ac.ae, samia.loucif@zu.ac.ae, mohammad.kuhail@zu.ac.ae, ahmed.seffah@zu.ac.ae

How to cite this paper: Negreiros, J., Loucif, S., Kuhail, M. A., & Seffah, A. (2021). Covid-19 in West & East Africa, a Geographical Weighted Regression Exploration with <http://mygeoffice.org/>. *Journal of Geoscience and Environment Protection*, 9, 20-33. <https://doi.org/10.4236/gep.2021.99002>

Received: August 2, 2021

Accepted: September 5, 2021

Published: September 8, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Understanding the dynamics that affect the spread of Covid-19 is critical for the development of government measures to stop and reverse this nowadays disease propagation. Like in any epidemiological study, it is essential to analyze the spatial data to account for the inherent spatial heterogeneity within the data (spatial autocorrelation). This paper uses Geographically Weighted Regression (GWR) to identify the factors that influence the outbreak of Covid-19 in Western and Eastern countries of Africa. The analyses include traditional linear regression (including descriptive statistics, hierarchical clustering and correlations were not forgotten either) to reveal the importance of eight risk factors (population density, median age, aged over 65 years, GDP per capita, cardiovascular death rates, diabetes prevalence, female and male smokers) regarding Covid-19 dissemination. It is believed that this is the first attempt to explore possible causes associated with the spread of the Covid-19 pandemic in these disadvantage countries, where some intriguing clues are presented for further research such as the positive relationship between the financial purchase power of nations and the total number of infected people or the smoker's gender impact on Covid-19.

Keywords

Covid-19, Statistics, Spatial Analysis, Geographical Weighted Regression, myGeoffice[©]

1. Introduction

Infectious diseases are caused by biological agents such as viruses, bacteria, fungi or prions and are transmitted either directly or indirectly from one individual to another (Ewald, 1994). The emerging infectious diseases have appeared from

time to time and contributed to a crucial global burden impacting both the public health and economy around the world. The global trend explicitly indicates an upsurge and inflation in the total number of outbreaks, disease richness (number of unique diseases), disease diversity (richness and outbreak evenness) and per capita cases (Smith et al., 2014).

The Black Death was a devastating global epidemic of bubonic plague that struck Europe and Asia in the mid-1300s. Over five years, the Black Death would kill more than 20 million people in Europe, almost one-third of the continent's population (History.com Editors, 2010b).

The Spanish flu pandemic of 1918, the deadliest in history, infected an estimated 500 million people worldwide, about one-third of the planet's population, and killed an estimated 20 million to 50 million victims. The 1918 flu was first observed in Europe, the United States and parts of Asia before swiftly spreading around the globe. At the time, there were no effective drugs or vaccines to treat this killer flu strain. Schools, theaters and businesses were shuttered, citizens were ordered to wear masks while bodies piled up in makeshift morgues before the virus ended (History.com Editors, 2010a).

Severe acute respiratory syndrome (SARS) is a viral respiratory illness caused by a coronavirus. SARS was first reported in Asia in February 2003. Over the next few months, the illness spread to more than two dozen countries in North America, South America, Europe and Asia before the SARS global outbreak of 2003 was contained.

The 2014 outbreak of Ebola virus disease in West Africa was the largest, most severe and most complex Ebola epidemic in history, according to the World Health Organization. More than 28,000 people were infected and over 11,000 people died before the international public health emergency ended in June 2016. Most of the cases occurred in three countries: Guinea, Sierra Leone and Liberia. Without an effective early-warning system, the virus spread rapidly within the region, revealing the failures of disjointed and under-resourced healthcare systems (World Vision, 2014).

In recent years, more than 84 countries have reported and experienced unprecedented resurfacing of Zika virus outbreaks. Zika is an enveloped single-stranded, positive-polarity RNA virus classified under the Flaviviridae family of viruses and was first isolated from the Zika Forest of Uganda (Quereshi, 2017). Its symptoms include fever, rash, headache, joint pain, conjunctivitis and muscle pain.

As of 2018, approximately 37.9 million people are infected with HIV globally and there were about 770,000 deaths (Wikipedia, 2021).

Coronavirus (Covid-19) is an infectious disease that spreads primarily through droplets of saliva or discharges from the nose when an infected person coughs or sneezes and it affects different people in different ways. Most infected people will develop mild to moderate illness and recover without hospitalization. Its most common symptoms include fever, dry cough, tiredness, aches and pains, sore throat, diarrhea, conjunctivitis, headache and loss of taste or smell.

Elderly and those with underlying medical problems like cardiovascular disease, diabetes, chronic respiratory disease and cancer are more likely to develop serious illness. The best way to prevent and slow down transmission is by washing your hands or using an alcohol based rub frequently and not touching your face. This includes physical distancing, wearing a mask, keeping rooms well ventilated, avoiding crowds and cough into a bent elbow or tissue.

At present, the world is on the verge of an unprecedented pandemic disease, as many cities go on lockdown in response to the spread of Covid-19. The consensus about the global economic shutdown is believed to be the primary determinant of stock market volatility that could lead to the biggest stock market crash in the 21st century. According to Estrada et al. (2021), the SARS outbreak caused more than 50 billion dollars of damage to the global economy. In parallel, the MERS outbreak caused irreparable damage in South Korea's tourist industry, translating into 2.6 billion dollars in losses. Furthermore, empirical evidence on seasonal epidemics such as Influenza, Dengue (8.9 billion dollars), Zika (US\$7-18 billion in Latin America) and Ebola (\$2.2 billion lost in 2015) virus also had a considerable economic impact in the affected areas (Estrada et al., 2021).

Confirmed by Estrada et al. (2021), exports in durable goods, mining and oil production are being severely hurt, shifting from a short-term recession to a medium depression by the end of 2021. In contrast, agriculture sector economies will enforce restrictions or even ban exports in an attempt to secure domestic demand first. As well, tourism, transportation, entertainment, retail and small businesses are devastated, followed by durable goods value chain. Logically, some sectors may benefit, such as ecommerce, food retail and the healthcare industry.

The aim of this article is to compare and to analyze the impact of 8 structural environmental variables (population density, median age, aged over 65 years, GDP per capita, cardiovascular death rates, diabetes prevalence and female/male smokers) in Western (Tunisia, Algeria, Morocco, Mali, Niger, Burkina Faso, Togo, Ghana, Liberia) and Eastern (Djibouti, Ethiopia, Uganda, Kenya, Tanzania, Zambia, Mozambique, Zimbabwe, Botswana, South Africa) countries of Africa when facing the total number of Covid-19 effected people. GWR (Geographical Weighted Regression) will be used in order to take into account spatial autocorrelation of the location data. After all, space itself also explains patterns of epidemiology, the systematic study of the distribution and determinants (causes, risk factors) of health-related states and events in specified populations (neighborhood, school, and city, state, global). The date of the given dataset (check Appendix) is February 8, 2021. The dependent variable regards the total number of Covid-19 infected population in these nations.

Besides this introduction, the present writing is divided into four major sections: Section two reviews major theoretically concepts of Geographical Weighted Regression (GWR) using myGeoffice[®]. Section three highlights several Eastern and Western country patterns of Africa by using conventional statistical ap-

proaches. Before the conclusion, section four stresses the spatial exploration of both sub-datasets based on 2D location, (x, y) coordinates, plus 4 (first sub-dataset) and 4 (second sub-dataset) independent variables.

2. Eastern and Western Africa: Conventional Statistics

Africa is the world's second-largest and second-most populous continent (**Figure 1**), after Asia in both cases. The coastline is 26,000 km long. With 1.3 billion people as of 2018, it accounts for about 16% of the world's human population. Africa's population is the youngest amongst all the continents. Despite a wide range of natural resources, Africa is the least wealthy continent per capita, in part due to geographic impediments, legacies of European colonization and the Cold War, predatory/neo-colonialist activities by Western nations and China and undemocratic rule and deleterious policies. Africa's largest country is Algeria, and its smallest country is Seychelles, an archipelago off the east coast.

In this study which intent to evaluate the impact of several infrastructure settings such as GDP per capita, cardiovascular death rates or diabetes prevalence of some Eastern and Western states of Africa when facing the total number of Covid-19 affected people using GWR happens due to the shift away on its traditional trade and investment partners in the West towards the East, particularly China and to a lesser extent, India and Malaysia. Western firms remain dominant in the telecoms and banking sectors, with Vodafone, France Telecom, Standard Chartered and Barclays Bank active in many countries. By contrast,



Figure 1. Africa map with the countries labels and myGeoffice[®] numbers used in this study.

Indian pharmaceutical companies are eroding the market share of Western firms. However, it is in the mining and oil and gas sectors that competition has become fiercest for established players including Exxon-Mobil, Shell, BHP Billiton and Rio Tinto. In the oil and gas sector, China National Petroleum Corporation, GAIL India, Malaysia's Petronas and PTT of Thailand all have active African operations (Africa Business, 2013). A second reason copes to take into account spatial autocorrelation of the epidemiological data.

Regarding the 10 Eastern countries (Botswana, Djibouti, Ethiopia, Kenya, Mozambique, South Africa, Tanzania, Uganda, Zambia, Zimbabwe), Females Smokers, Males Smokers, Age 65 Older, Cardio Death Rate and Diabetes Prevalence variables do not follow a statistically Gaussian distribution, according to Shapiro-Wilk and Kolmogorov-Smirnov tests. South Africa is an extremely positive outlier when facing the total number of infected population (817,878 VS 110,676 average). The same rule happens to Uganda when it faces population density (213 VS 66 average). Both South Africa and Botswana neighbors are considered a positive outlier when GDP is taken into account (12,295 and 15,807, respectively, VS 4663 average) while Mozambique suffers from a high value of Cardio Vascular death rate (330 VS 226 average).

For the 9 Western countries (Algeria, Burkina Faso, Ghana, Liberia, Mali, Morocco, Niger, Togo, Tunisia), Population Density, Median Age, Females Smokers, Males Smokers and Diabetes Prevalence variables do not follow a statistically Gaussian distribution. Morocco is an extremely positive outlier when facing total number of infected population (381,188 VS 71,322 average). The same rule happens when it faces Cardio Vascular death rate (419 VS 280 average) too. Tunisia is considered a positive outlier when Male Smokers is taken into account (65 VS 23 average).

When comparing both sub-datasets, the T independent test of SPSS[®] stresses that the average of Age 65 Older (4.194 against 3.345, respectively) and Female Smokers (0.956 against 3.36, respectively) variables are statistically significantly different for a 95% level of confidence. Males Smokers (27.28 against 24.89, respectively) and Median Age (22.24 against 20.74, respectively) hold a Sig. (2-tailed) value of 0.065 and 0.061, respectively.

Globally, strong positive Pearson correlations may be found between GDP Per Capita with Median Age (0.808), Age 65 Older (0.716), Male Smokers (0.59) and Diabetes Prevalence (0.514). Definitely, the power purchase of individuals is closely related with some health diseases and long life factor. Not surprisingly, Diabetes Prevalence correlates with Median Age and Age 65 Older. In a further detail, Cardio Vascular death ratio highly correlates with Age 65 Older (0.456) and Male Smokers (0.59) as well. Yet, no significant statistical correlation with Female Smokers (0.388) was found.

In terms of linear OLS regression, the best found F-Anova significant model (Sig. = 0.003) became as follows: Total Cases = $-238,839 + 109,339 \times \text{Age 65 Older} + 56,508 \times \text{Female Smokers} - 7850 \times \text{Male Smokers}$. The adjusted R Square equals 50.9% with no error autocorrelation (Durbin-Watson = 2.029)

while the three VIFs (Variance Inflation Factor) of each independent variable vary between 1.123 and 2.989, confirming a low collinearity issue among them.

The curiosity at this point regards the positive and negative side-effect of smoking for both genders when it comes to Covid-19 total number of infections. Hierarchical clustering (it underlines Morocco and South Africa versus the 17 remaining countries) and Principal Component Analysis (KMO = 0.738, Varimax rotation with Kaiser Normalization) confirms this view where Female Smokers and Cardio Vascular Death Rate belongs to a different rotated component of Median Age, Aged 65 Older, GDP Per Capita, Diabetes Prevalence and Male Smokers variables (18.83% against 48.241%, respectively). The third axis is represented by Population Density only (13.9% of the total variability). Do gender and smoke influence the contagious process of Covid-19? Certainly, further research is needed to confirm this hypothesis.

When considering the 10 Eastern African countries, the best OLS linear regression (Adjusted R Square = 49.2%) with a significant F-Anova of 0.039 was as follows: Total Cases = $-491,509 + 144,810 \times \text{Age 65 Older} + 35,074 \times \text{Female Smokers}$. Although both T independent tests were not significant for a 95% level of confidence, no residual autocorrelation (Durbin-Watson = 1.915) and no (VIF = 1.436) collinearity were found.

At last, the best OLS linear regression (Adjusted R Square = 94.7%) for the 9 Western African states, with a significant F-Anova of 0.000 was as follows: Total Cases = $-605,128 + 2432 \times \text{Cardio Vascular Rate} - 574 \times \text{Population Density}$. Both T independent tests were significant for a 95% level of confidence whereas no residual autocorrelation (Durbin-Watson = 1.937) and no (VIF = 1.113) collinearity were found.

Consequently, a key highlight is the independent variables comparison between Eastern and Western Covid-19 infected countries, that is, Age 65 Older & Female Smokers VS Cardio Vascular Rate & Population Density, respectively. Why such a difference of factors between both African sub-regions?

3. Geographical Weighted Regression

GWR is a spatial regression technique that provides a local model to understand/predict from a set of independent variables by fitting a regression equation to every feature in the spatial dataset (Negreiros & Diakite, 2019). However, the novelty of this GIS approach regards the estimation of the dependent variable for each location by also depending on their closer space neighbors and defined by the Kernel bandwidth and a decay spatial function (respecting First Law of Tobler, that is, everything is related to everything else but near things are more related than distant things).

For instance, Wang, Zhao and Chen (2018) use GWR to examine the spatial effect of urbanization, energy intensity, energy structure and income on HCE (household CO₂ emissions) in China. Their results indicate an obvious spatial effect on carbon emissions in the studied provinces whose urbanization impact presented an increasing trend from the south-eastern coast to the north-west

from 2000 to 2015. As well, energy intensity had a remarkably positive effect on HCE for the same time period, although it had a negative effect in all provinces in 2005 and in some provinces in 2010. Likewise, income was a powerful explanatory factor for growth in household CO₂ carbon emissions in all years.

When doing spatial estimations with GWR, the available data are weighted according to their distance to any (x, y) location. Mathematically, observations close to the center of the local window are weighted more heavily than observations farther away, respecting Tobler's Law. For GWR, every beta (vector of parameter estimates) equals $(Y^T W Y)^{-1} Y^T W Z$, where W represents the diagonal matrix of weight factors, Y is the matrix of independent observations variable while Z denotes the vector of dependent observations variable. In a detail perspective, W is a zero content matrix with the exception of its diagonal which represents the geographical weights of observations around any location and given by a weighting function. myGeoffice[®] follows the Gaussian decay function which obtains the weight of each i observation to a generic location based on the following math expression: $W(X_i) = \text{EXP}(-0.5 \times (d/r)^2)$, where r equals the kernel bandwidth and d represents the Euclidean distance between the observation and the site itself. For a 30 sample dataset with 2 independent variables, for instance, the number of rows and columns for each matrix will be as follows: $Y = [30, 2]$, $Y^T = [2, 30]$, $Z = [30, 1]$ and $W(X_i) = [30, 30]$.

It is, then, up to the practitioner to set the key factors under this formula. As a rule of thumb, if the sample points are reasonably regularly spaced within the study area, a kernel with a fixed bandwidth is a suitable choice for modelling. If the sample points clustered, it is generally desirable to allow the kernel to accommodate this irregularity by increasing its size when the sample points are sparser and decreasing its size when the sample points are denser.

Since this dynamic bandwidth technic is not accomplished by myGeoffice[®], the user may run GWR with different bandwidth kernels and compare the measure of goodness of fit, the corrected Akaike Information Criterion ($\text{Num-Samples} \times \text{LOG}(\text{Residual Sum of Squares}) + 2 \times K + (2 \times K \times (K + 1))/(\text{Numsamples} - K - 1)$), where K denotes the total number of parameters of the model). The optimal value for the bandwidth presents the lowest AICc although it is up to the analyst to choose the final best value (based on prior reasons, for instance).

As a final note, each parameter of GWR has a sign and a magnitude according to each location (the essence of spatial heterogeneity, that is, the structure of the model changes from place to place across the study area as the parameter estimates change towards each other inside the model). If the sign is positive, an increase of the variable value to which the parameter refers will produce an increase in the dependent variable. If the sign is negative, a decrease will be induced. As stated by Charlton & Fotheringham (2009), parameter estimates for a variable that are close to zero often tend to be spatially clustered indicating that in these sub-regions of the study area, changes in this variable do not influence changes to the dependent variable. This is potentially interesting and encourages

further curiosity about the process, the data, the model and the outcome.

For reference, IBM-SPSS and myGeoffice[®] will be used under this research. If the first software is well known, the second one does not follow this trend. As a reference, myGeoffice[®] is not a comprehensive statistical package in the tradition of solving everyone's problems (Negreiros et al., 2009). Written for an Internet Information Service[®] setting, it was developed with the philosophy that spatial autocorrelation, Kriging, deterministic interpolators and other spatial techniques is needed as a learning tool by individuals with limited geostatistical knowledge (Negreiros & Painho, 2006). myGeoffice[®] deals with spatial interpolation in conjunction with spatial association measures in a Web continuum process, instead of a loose local spatial function (Negreiros, 2018). From this view point, an element of its originality and innovation can, thus, be appreciated (Negreiros et al., 2011).

3.1. Population Density, Median Age, Aged 65 Older and GDP Per Capita: GWR

The average of Covid-19 infections in these nineteen countries equals 87,691 with a standard deviation of 197,776. This happens because of the huge range (817, 366 = 817, 875-509) of values in this highly positive skewness distribution of 3.27. Certainly, this dependent variable does not follow a Gaussian distribution.

To embrace GWR, a key parameter must be set up: the kernel bandwidth of the Gaussian function in order to compute each Beta weight of each sample on the estimation at (x_0, y_0) site. As expected, locations that are far behind this cutoff distance, their contribution weight equals zero. Otherwise, their Beta weights are given by the Gaussian decay function where high weights correspond to closer distances to (x_0, y_0) .

For a kernel bandwidth of 250, GWR output of myGeoffice[®] (Figure 2) stresses three categories (high, medium, low) of classes regarding each Beta estimator's weights. Therefore and under this study, the goal is to unmask possible independent variables relationships when explaining the total number of Covid-19 infectious people in Western and Eastern African countries and, if possible, clarify the causes of these trends. Still, it is crucial to clearly state that GWR can also be used for interpolation purposes.

Population Density (Beta 1), Aged 65 Older (Beta 3) and GDP Per Capita (Beta 4) hold a strong and similar spatial distribution pockets, that is, highest and medium Beta (strength of the impact of these three factors on Covid-19 total infections) values in Eastern countries with low ones in the West African ones. The exception comes from Liberia and Ghana West nations that hold medium values like Kenya, Uganda, Ethiopia and Djibouti of the Eastern coast. On the other hand, the strength impact on Covid-19 presents a reverse map layout when facing Medium Age (Beta 2) independent variable, that is, West nations hold higher Beta weight values. Again, Ghana, Togo and Niger present similar medium values of their Eastern counter partners of Kenya, Uganda, Ethiopia and Djibouti.

GWR Final Estimation Mapping (dependent variable)

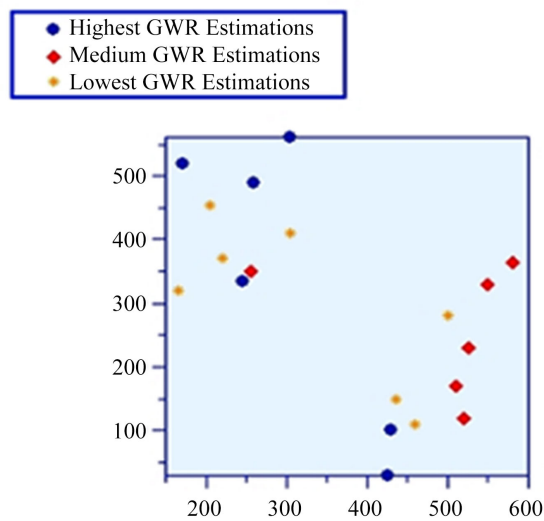


Figure 2. Statistical calculated values including the Kernel bandwidth distance necessary for GWR (left) and residuals/GWR estimations categorization of myGeoffice© (right). For model's comparison purposes (the lowest, the better), the Corrected Akaike Information Criteria equals 557.2.

Certainly, spatial autocorrelation plays a key role in these patterns. This concept describes the presence of systematic spatial variation in a variable. A positive spatial autocorrelation, which is most often encountered in practical situations, is the tendency for areas or sites that are close together to have similar values. Somehow, it can be interpreted as a descriptive index, measuring aspects of the way things are distributed in space (clustering) but, at the same time, it can be seen as a casual process measuring the degree of influence exerted by something over its neighbors (correlation). In this particular case, there is a spatial positive correlation corridor for all Betas between West/Central and East/Central sub-regions. On the other hand, a negative spatial autocorrelation may be found between West/North and East/South sub-regions.

Eastern Africa presents a higher total number of infected people (close to the double) than Western nations. From the population density point of view (Beta 1), the same global pattern can be found. However and in a regional evaluation, the red countries presents low relative cases with high densities of populations. Both South Africa and Morocco perform worst on the sense of a huge total number of infected individuals and with a comparative low population density. The large-scale trend between the total number of cases and median age (Beta 2) is close to a 45 degrees line. The same statement may be reproduced for the two regional sub-regions, with the exception of Morocco and South Africa, the positive outliers of this dataset. For Beta 3, there is an expected strong direct relationship between the number of elderly people and the total number of infected

with the wonderful exception of Tunisia which rejects this trend. In a paradox way, wealthier countries hold a greater propensity for higher infected individuals.

3.2. Cardiovascular Death Rate, Diabetes Prevalence, Female and Male Smokers: GWR

The present aim of this sub-section involves the analysis of new four independent variables and their possible hidden effects and factors for a possible explanation of the Covid-19 infected people in West and East sub-regions of Africa. Again, this investigation is considered spatial since GWR also includes the location factor (**Figure 3**).

The three local pockets of high, medium and local impact of Cardio Vascular Death Rate (Beta 1) factor follows a completely different spatial pattern from the previous independent variables, where Northern/West nations hold a higher Beta1 weights than their counter partners of Southern/East ones. Beta 2 (Diabetes Prevalence) and Beta 3 (Female Smokers) behaves on the reverse way. Curiously, the East nations of Djibouti and Ethiopia hold the same low classification of five their West countries. Again, South Africa is the major positive outlier for Cardio Vascular Death Rate and Diabetes Prevalence factors and Morocco for Female Smokers. As well, the Male Smokers (Beta 5) variable holds a quite different spatial pattern regarding their direct influence on Covid-19 disease. The positive outliers are both Western neighbors of Morocco and Tunisia. In a humble way, the present authors cannot explain this smoker's gender impact on Covid-19, only suggesting that further studies must be accomplished to understand this suspicious GWR outcome.

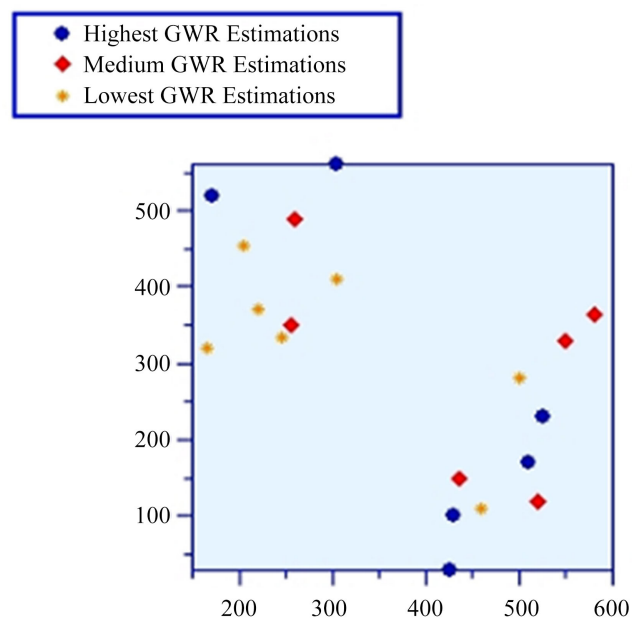


Figure 3. A slightly altered view of **Figure 2**.

4. Conclusion

Covid-19 is a complicated multifactorial disease, including the association between risk factors and disease occurrence at the group level worldwide, in general, and in Africa, in particular. Governments continuously set up regional policies for more tailored interventions to fit the local disease conditions. This approach is likely to improve the effectiveness of Covid-19 controls in a short term, but requires policy makers to have an understanding of local drivers such as population density or elderly people percentage if long term goals need to be achieved.

An association identified by GWR regards Population Density (Beta 1), Aged 65 Older (Beta 3) and GDP Per Capita (Beta 4) whose influence is stronger in Africa East rather than the West coast. It is not by chance that the total number of Covid-19 cases in the East is almost the double of Western nations (1,106,763 and 641,902, respectively). Confirmed by the highest Beta 3 weights of Western African countries, this difference is not farther significant because the average median age in the West is greater than the East coast of Africa (20.74 and 22.24 years old, respectively). However, the puzzle explanation issue concerns Beta 4 weights spatial distribution of GWR. How nations with higher financial power contribute to an increase of the total number of infected people?

Another unknown involves the medium classification similarity of GWR weights towards the West/East middle corridor. In a direct comparison, the total number of infected people in Djibouti (5708), Ethiopia (113,735), Uganda (23,200), for instance, is, at least, the double of Niger (1894), Ghana (52,274) and Togo (3111). A possible explanation is closely related with the non-capability of some of these governments to keep the right track of infected Covid-19 records.

The number of infections caused by male and female smokers' factors is an intriguing topic. Globally, the ratio between both factors is 8.13, an expected value due to the fact of male's propensity to smoke more often around the world. Yet, this ratio becomes 7.4 and 28.5 for Eastern and Western countries, respectively, leading to the conclusion of a heavier existence of male smokers in West African nations. Looking at Beta 3 and Beta 4 weights in **Figure 2** and **Figure 3**, the spatial distribution is slightly different, giving an indication of spatial fragmentation and contiguity, which may raise questions on the influence (or not) of smoking and gender on the spread of Covid-19. Still, if we add the fact that there is a negative relationship between the linear trend of total Covid-19 cases and the linear trend of smokers (with the exception of total males smokers in the West African coast), this raises questions on the positive smoking effect against Covid-19 spreading. Further studies need to be conducted to confirm (or not) this suspicious.

Fitting a global model to Covid-19 is a complicated process, not least because of the spatial nature of many of the variables of interest, but also because the spread into new areas can only be detected when testing takes place. Simultaneous autoregressive model and variography to account for spatial dependence

could be another possibility. Yet and again, GWR is a useful tool for descriptive and exploratory data analysis (Brunton et al., 2017).

Because of the large number of variables under consideration, the individual associations between all variables were not examined in great detail except where strong correlations were observed. Instead, more importance was placed on the contribution each variable made to the fit of the model rather than the magnitude of its effect.

At last and even though GWR was used as a more robust method for accounting for spatial dependence, the relationships identified were complex with regression Betas coefficients switching between negative and positive values in different nations, which indicate that while using independent variables to focus on a particular risk factor might be beneficial in one area, it may actually be detrimental in another area. As with any regression method, they cannot prove causality and other drivers may not have been retained as a result of the deliberately stringent method of covariate selection (Brunton et al., 2017).

Conflicts of Interest

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

References

- Africa Business (2013, May 2). *East vs West—Who Is Winning in Africa?*
<http://African.Business/2013/05/economy/east-vs-west-who-is-winning-in-africa/>
- Brunton, L., Alexander, N., Wint, W., Ashton, A., & Broughan, J. M. (2017). Using Geographically Weighted Regression to Explore the Spatially Heterogeneous Spread of Bovine Tuberculosis in England and Wales. *Stochastic Environmental Research and Risk Assessment*, 31, 339-352. <https://doi.org/10.1007/s00477-016-1320-9>
- Charlton, M., & Fotheringham, A. (2009). *Geographically Weighted Regression—White Paper* (14 p.). Science Foundation of Ireland.
http://www.geos.ed.ac.uk/~gisteac/fspat/gwr/gwr_arcgis/GWR_WhitePaper.pdf
- Estrada, R., Mario & Koutronas, Evangelos & Lee, Minsoo. (2021). Staggression: The Economic and Financial Impact of COVID-19 Pandemic. *SSRN Electronic Journal*, 15, 19-33. <https://doi.org/10.5709/ce.1897-9254.433>
- Ewald, P (1994). *Evolution of Infectious Disease*. Oxford University Press on Demand.
- History.com Editors (2010a, May 19). *Spanish Flu*. A & E Television Networks.
<http://www.history.com/topics/world-war-i/1918-flu-pandemic>
- History.com Editors (2010b, September 17). *Black Death*. A & E Television Networks.
<http://www.history.com/topics/middle-ages/black-death>
- Negreiros, J. (2018). *Spatial Analysis Techniques with MyGeoffice®* (409 p). IGI Global.
<https://doi.org/10.4018/978-1-5225-3270-5>
- Negreiros, J., & Diakite, A. (2019). Ten Spatial Problems with MyGeoffice® for Teaching Purposes. *Open Journal of Social Sciences*, 7, Article ID: 93971.
- Negreiros, J., & Painho, M. (2006). SAKWeb®—Spatial Autocorrelation and Kriging Web Service (Part II), Internet Research 7.0: Internet Convergences. *WIT Transactions on Ecology and the Environment*, 89, 11 p. <https://doi.org/10.2495/GEO060091>

- Negreiros, J., Costa, A., Painho, M., & Santos, J. (2011) Autocorrelation, Autoregression and Kriging: Appraisal for the Spatial Interpolation Issue. *International Statistic Institute 56th Conference*, Lisboa, Ago 07.
- Negreiros, J., Painho, M., Aguilar, M., & Aguilar, F. (2009). The Local and Global Confidence Uncertainty Plumes of SAKWeb[®]. *12th AGILE International Conference on Geographic Information Science*, Hanover, 2-5 June 2009, 23-33.
<https://doi.org/10.2495/RISK100031>
- Negreiros, J., Painho, M., Aguilar, M., & Aguilar, F. (2010). The Local and Global Confidence Uncertainty Plumes of SAKWeb. In J. Haunert, B. Kieler, & J. Milde, (Eds.), *AGILE 2009*. Hannover, Germany.
- Quereshi, A. (2017). *Zika Virus Disease: From Origin to Outbreak*. Academic Press.
- Smith, K. F., Goldberg, M., Rosenthal, S., Carlson, L., Chen, J., & Chen, C. (2014). Global Rise in Human Infectious Disease Outbreaks. *Journal of the Royal Society Interface*, 11, Article ID: 20140950. <https://doi.org/10.1098/rsif.2014.0950>
- Wang, Y., Zhao, M., & Chen, W. (2018). Spatial Effect of Factors Affecting Household CO₂ Emissions at the Provincial Level in China: A Geographically Weighted Regression Model. *Carbon Management*, 9, 187-200.
<https://doi.org/10.1080/17583004.2018.1451964>
- Wikipedia (2021). *HIV*. http://en.wikipedia.org/wiki/Epidemiology_of_HIV/AIDS
- World Vision (2014). *2014 Ebola Virus Outbreak: Facts, Symptoms, and How to Help*.

Appendix

Country Number	Coutry Side	Country	XX	YY	Total Cases	Population Density	Median Age	Aged 65 Older	Gdp Per Capita	Extreme Poverty	Cardiovasc DeathRate	Diabetes Prevalence	Female Smokers	Male Smokers
2	Eastern	Botswana	430	100	12,058	4.044	25.8	3.941	15,807.37	20	237.372	4.81	5.7	34.4
18	Eastern	Zambia	435	150	17,931	22.995	17.7	2.48	3689.251	57.5	234.499	3.94	3.1	24.7
11	Eastern	Mozambique	520	120	16,326	37.728	17.7	3.158	1136.103	62.9	329.942	3.3	5.1	29.1
4	Eastern	Djibouti	580	365	5708	41.285	25.4	4.213	2705.406	22.5	258.037	6.05	1.7	24.5
19	Eastern	Zimbabwe	460	110	10,839	42.729	19.6	2.822	1899.775	21.4	307.846	1.82	1.6	30.7
13	Eastern	SouthAfrica	425	30	817,878	46.754	27.3	5.344	12,294.88	18.9	200.38	5.52	8.1	33.2
14	Eastern	Tanzania	525	230	509	64.699	17.7	3.108	2683.304	49.1	217.288	5.75	3.3	26.7
8	Eastern	Kenya	240	70	88,579	87.324	20	2.686	2993.028	36.8	218.637	2.92	1.2	20.4
5	Eastern	Ethiopia	550	330	113,735	104.957	19.8	3.526	1729.927	26.7	182.634	7.47	0.4	8.5
17	Eastern	Uganda	500	280	23,200	213.759	16.4	2.168	1697.707	41.6	213.333	2.5	3.4	16.7
9	West	Mali	205	495	5290	15.196	16.4	2.519	2014.306	43	268.024	2.42	1.6	23
12	West	Niger	305	410	1894	16.955	15.1	2.553	926	44.5	238.339	2.42	0.1	15.4
1	West	Algeria	260	490	88,825	17.348	29.1	6.211	13,913.84	0.5	278.364	6.73	0.7	30.4
7	West	Liberia	165	320	1676	49.127	19.2	3.057	752.788	38.6	272.509	2.42	1.5	18.1
3	West	BurkinaFaso	220	370	3315	70.151	17.6	2.409	1703.102	43.7	269.048	2.42	1.6	23.9
16	West	Tunisia	305	560	104,329	74.228	32.7	8.001	10849.3	2	318.991	8.52	1.1	65.8
10	West	Morocco	170	520	381,188	80.08	29.6	6.769	7485.013	1	419.146	7.14	0.8	47.1
6	West	Ghana	245	335	52,274	126.719	21.1	3.385	4227.63	12	298.245	4.97	0.3	7.7
15	West	Togo	255	350	3111	143.366	19.4	2.839	1429.813	49.2	280.033	6.15	0.9	14.2