SPATIAL STRUCTURE AND CLIMATIC ASSOCIATIONS WITH COVID-19 CASES ACROSS THE GLOBE

Olanrewaju Lawal[1]
Anyiam Felix Emeka[2]

ABSTRACT

The study examined the spatial structure and the association between COVID-19 cases and selected climatic variables. Data on cases, deaths, recovery were obtained from the COVID-19 Resources website of the Environmental Systems Research Institute (ESRI). The climatic variables were selected included Land Surface Temperature (LST) and Water Vapour (WV) and collated from the NASA Earth Observations (NEO). Spatial and inferential statistics were used to examine spatial autocorrelation and associations with these variables. Results show that China, Italy, and Iran have the largest number of confirmed cases, the highest recovery (81%) was recorded in China. Confirmed cases have 7 clusters and 2 outlier locations. There are 21 and 17 spatial outliers for recoveries and deaths respectively. There are 2 natural clusters of the incidences and 98.7% of the locations belong to one of the groups. A weak but statistically significant ($P<0.05$) associations were observed for the incidence and the climatic variables. The analysis of spatial structure revealed more insight into the distribution of the disease, shedding more light on areas with needs for more investigation (outlier locations) and providing opportunities for mitigating spread and re-emergence.

Keywords: COVID-19, Spatial Clustering, Pandemic, Spatial Autocorrelation, Climatic Variables.

JEL Classification: I18

1. INTRODUCTION

The seasonal cycle and its ubiquitous association with infectious diseases have been well researched (Dowell, 2001; Pascual & Dobson, 2005; Buckee et al., 2017), and from measles, diphtheria and chickenpox of childhood origin to vector-borne diseases including malaria (Martinez, 2018). And the effects of this seasonality may vary according to geographic locations, thus when adequate measures are not adopted to prevent and curtail the spread of seasonal infections, such diseases can turn to an epidemic and sometimes a pandemic thus, making it pertinent that the current novel Corona Virus of 2019 (COVID-19) pandemic be understood from the dimension of its relationship with climatic and the spatial structure inherent in its cases. This would provide a basis for identifying potential climate influence and identification of spatial clusters across the world to inform the development of initiatives to further curtail the spread and future recurrence. To this end, this study examined the spatial structure formed by the incidence of COVID-19 at the global and local scales and examined the relationship between the attributes of the incidence and selected climatic variables.

Devastating pandemics have existed since 541 A.D when the world experienced the bubonic plague that wiped out 25-50 million people in one year (Morony, 2007). It was

[1] University of Port Harcourt, Port Harcourt, Nigeria (olalaw@hotmail.com)
[2] Centre for Health & Development, University of Port Harcourt, Port Harcourt, Nigeria (felix.anyiam@uniport.edu.ng)

one of the worst outbreaks the world experienced, and it lasted for another 225 years, sweeping throughout the Mediterranean world until 750CE. According to Cohn (2002) between 1347 and1351, the Black Plague killed more than 75 million people in the Middle Eastern lands of China, India, as well as Europe. Furthermore, the author reported that the Spanish Flu pandemic, caused by an H1N1 virus with genes of avian origin, killed over 50 million people in one year in 1918. In America, 675,000 deaths were recorded; 2,000,000 in Sub-saharan Africa, and 500,000 in Nigeria, out of a population of 18 million in less than 6 months (Ohadike, 1991). The 20th century experienced the smallpox pandemic which claimed 300-500 million lives and is presently the only human disease that has been eradicated, according to the World Health Organization (World Health Organization, 1980; Voigt et al., 2016). In recent times, the global pandemic of Tuberculosis (TB) continues to kill over 1.5 million people annually and approximately 1 death in every 21 seconds (TB Alliance, 2020) and despite the availability of effective treatment, the emergence of multi-drug resistance (MDR-TB) is thwarting any effort of slowing the number of deaths or complete eradication (Di Gennaro et al., 2017) and there are fears on how the Coronavirus (COVID-19) pandemic that is sweeping the world right now will increase the death toll amongst TB and Human Immuno-Deficiency Virus (HIV) patients with already underlying medical conditions coupled with low immune systems (Liu et al., 2020; Pai, 2020). From the foregoing, it is evident that influenza viruses have been with humanity longer than we have anticipated, with so much ability to mutate into varying forms in no time, thus producing new strains (Smith et al., 1951; Ohadike, 1991; Pauly et al., 2017). The World Health Organization (WHO) in 1999 made available its first influenza preparedness plan that consisted of 6 outlined phases, which served as a blueprint for countries to draw up their national strategies (Iskander et al., 2013). WHO in 2009 developed a revised version of the phases to differentiate appropriate preparedness and response to enable countries effectively handle the high mutation rate of viruses especially with its ability to move from animals to humans (World Health Organization, 2009). As at present, WHO has recorded 4 types of influenza viruses namely: A, B, C and D while A and B cause most of the seasonal epidemics (World Health Organization, 2018).

The seasonality of infectious disease has puzzled epidemiologists since the 18th century. Although it is worth noting at this time that seasonality alone may, in certain cases, not be completely responsible for all epidemics, as other factors such as the immunity of the host, the susceptibility of the environment, or simply a combination of both can play a role. For example, the seasonal transmission of measles has been seen to occur every two to five years rather than annually (Kesson, 2011). This is why great insight into spatiotemporal patterns can help clarifying many issues around the transmission of such ailments (Coletti et al., 2018). Several works have examined spatial structure across a wide range of topics e.g. Kim et al. (2003) - air quality and house prices; Grineski and Collins (2010) - environmental inequality/injustice; Longley and Tobón (2004) - urban deprivation; Ceccato and Uittenbogaard (2014) - crime pattern; Yu et al. (2007) - house prices; Zou (2014) - higher-priced mortgages; Fan and Myint (2014) - landscape fragmentation; (Lawal, 2015) - age dependencies, Lawal (2017) - dependency and socio-economic factors. Similarly, many studies have explored the spatial pattern of diseases and its association with climatic variables, which have helped in understanding disease patterns and their dynamics (Wu et al., 2016; Wangdi & Clements, 2017; Anwar et al., 2019). The understanding of infectious diseases, spatial dynamics and seasonality can improve preparedness and response plan thus mitigating its burden worldwide, especially with the event of the COVID-19 pandemic. The study, therefore, aims to explore the spatial pattern of the disease and its association with selected satellite-derived climatic variables.

Infectious disease outbreaks such as the Avian Human Influenza (AHI), Influenza A (H1N1), Severe Acute Respiratory Syndrome (SARS), Ebola, and now the novel Coronavirus (2019-nCoV) poses serious physical and economic losses for our world (Sands et al., 2016; Fan et al., 2017) and understanding the role seasonality play could support complete control, thus strengthening our ability to predict its occurrence in the future (Dowell & Ho, 2004; Fares, 2013). Some epidemics could reflect seasonal variations and understanding of tropical areas in terms of cold air or low humidity may contribute to knowledge on reoccurrences. The novel Coronavirus disease, designated as COVID-19 by the World Health Organization in February 2020 (World Health Organization, 2020a) has become a pandemic with continued rapid spread since December 2019 when it was first detected in Wuhan, China (The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team, 2020), with 439,940 confirmed cases, 19,744 deaths and 111,942 recoveries as at 25 March 2020 (3:25 PM) (Worldometer, 2020). The World Health Organization first declared COVID-19 as a public health emergency of international concern and now a pandemic as it has been found in almost all regions of the world (World Health Organization, 2020c). The virus, at its first discovery, was in patients exhibiting illness caused by pneumonia (Chen et al., 2020). The clinical features have ranged from mild to severe respiratory illnesses; which may include fever, dry cough, fatigue and difficulty in breathing, and spreads during close contacts with respiratory droplets when people sneeze or cough (World Health Organization, 2020b). There are no vaccines now and recommendations have been solely on preventive measures such as hand washing, social distancing, covering of mouth when coughing and self-isolation (Centre for Disease Control, 2020).

Not all viruses have a biological constant for pathogenic spread, as some may be environmentally and behaviorally escalated, as we have seen in COVID-19 and such environmental factors such as climate and seasonality may modulate transmissibility as have been observed in several studies (Grassly & Fraser, 2006; Coletti et al., 2018; Martinez, 2018).

Lowen and Steel (2014) in their study using the guinea pig as a model host observed that aerosol spread of the influenza virus is dependent upon both ambient relative humidity and temperature. A highly efficient transmission was observed at 5°C and slowed reasonably at 30°C. The dry conditions (20% and 35% RH) showed more transmission than humid (80% RH) conditions. Also, the authors of a study on environmental predictors of seasonal influenza epidemics across temperate and tropical climates simply observed that Human influenza infections exhibit a strong seasonal cycle in temperate regions (Tamerius et al., 2013). Based on information from modelling epidemiological and climatic information from 78 study sites, they concluded that seasonal influenza thrives on the: "cold-dry" and "humid-rainy" periods. Lowen and Steel (2014) in their study observed that the relative humidity and temperature simply increase the transmission of the respiratory droplets.

Analyses of spatial patterns and distribution of human activities and natural phenomena are often leveraged for the creation regions of across an area of interest. Thus, guiding activities and measures to address issues within the area. Spatial clustering is fundamentally based on our understanding of spatial autocorrelation. Just as economic, social, cultural and political attributes and activities clusters so also is disease occurrence. Observation in space often suffers from spatial autocorrelation - the exhibition of correlation between the values of a variable attributable to their location. This thus nullifies the independent observation assumption, of conventional statistics as alluded to by Tobler's First Law of Geography (Griffith, 1987). Human activities and habitation are directly impacted by location. This interrelationship and dependencies often lead to the increasing agglomeration of businesses, industries, and infections/diseases at specific locations (e.g., established urban centres). Similarly, there is a tendency for the segregation of people with a similar culture, tradition,

behaviour, social class, etc. across different regions and locations. Spatial clustering affects economic activity, thus leading to spatial unities and differences.

The study of the spatial structure formed by diseases could provide an understanding of the potential impact of location and its emergence, spread, and reoccurrence, thereby contributing to informed decision-making for planning and management. Also, regionalisation using multivariate clustering methods could provide new insights into the interaction among location, demographic, and socio-economic factors.

2. DATA AND METHODS

2.1 Data

Data about the cases, deaths, recovery were sourced from the COVID-19 Resources website of Environmental Systems Research Institute (ESRI) - https://coronavirus-resources.esri.com. The version of the dataset used was last updated on the 13th of March 2020. This dataset is point vector data. Two climatic variables were selected to represent temperature and rainfall. Monthly Land Surface Temperature (LST) for day (LSTD) and night (LSTN), as well as the monthly Water Vapour (WV), were collated from the NASA Earth Observations (NEO) website - https://neo.sci.gsfc.nasa.gov/ for December (2019), January (2020), February (2020). These datasets were downloaded in GEOTIFF floating raster format at a resolution of 0.1 degrees.

2.2 Methods

Confirmed, recovery and death incidence were all subjected to exploratory spatial data analysis. This was carried out to understand the spatial structure within the dataset. Moran's I statistic and Local Indicator of Spatial Association (LISA) or Cluster and Outlier Analysis were computed to analyze the existence of global spatial autocorrelation (Moran's I) and determine the presence of local spatial clusters around locations, thus allowing for inferences on the stationarity of global spatial autocorrelation. This was carried out within ArcGIS (ESRI, 2020).

Examination of the relationship between climatic variables selected (LSTN, LSTD, and WV) and the attributes of the diseases was preceded by the extraction of values from the raster data to the point data with a GIS environment. A correlation analysis (Spearman Rank Correlation) was carried out to examine the linear relationship between climatic variables selected and the attributes of the diseases (confirmed, deaths, recovered), with SPSS version 23 (IBM, 2015).

A two-step cluster analysis (IBM, 2016) was carried out to examine natural groupings from the combination of the confirmed, recovery, and death cases from infection by COVID-19. This method was executed by carrying out a pre-clustering procedure to build a data structure (Cluster feature tree) with nodes (branch) and leaf entry (sub-clusters); for these, the log-likelihood measure of distance was used. This process was followed by the resolution of atypical values (outliers) while the final clustering was carried using the initial pre-clustering sub-clusters excluding the outliers (noise) via a hierarchical clustering method. The optimal number of clusters was determined automatically using Schwarz's Bayesian information criterion (BIC), essentially the number of clusters with the highest ratio of distance measure - this is based on the current number of clusters as against the previous number of clusters.

3. RESULTS AND DISCUSSION

3.1 Comparison of Aggregated Country-Level Data

The aggregation of the cases across countries showed that China, Italy, and Iran have the largest number of confirmed cases (Table 1) based on the data collated. The top 10 countries for confirmed cases were also all in the Northern hemisphere. As the confirmed cases vary so also does the rate of recovery, with a higher proportion (81%) in China and the worst recovery within the period under investigation in Norway (0.1%). Andorra, Jordan, and Nepal recorded 100% recovered, this is because they all have one confirmed patient who has recovered. Among the top 10 countries, the death rate was highest in Italy (7.2%) followed by Iran and China at 4.5% and 3.9% respectively. Guyana and Sudan have the highest death rate, this could be attributed to them having only one confirmed case resulting in death.

Table 1. Aggregated Distribution of COVID-19 Cases for the Top 10 Countries with Confirmed Cases

Country	Confirmed	Recovered	Deaths
China	80,973	65,634	3,193
Italy	17,660	1,439	1,266
Iran	11,364	2,959	514
Korea, South	8,086	510	72
Spain	5,232	193	133
Germany	3,675	46	8
France	3,667	12	79
US	2,174	12	47
Switzerland	1,139	4	11
Norway	996	1	1

Source: COVID-19 Resources Website of Environmental Systems Research Institute (ESRI)

Kruskal-Wallis test was conducted testing the hypothesis that the distribution of the cases (confirmed, recovered, and deaths) is the same across the countries. The result (Table 2) showed a statistically significant difference across countries, and what this means is that all the case attributes are not the same across the country i.e. the count of confirmed, recovered and death cases across the countries are significantly different ($p<0.05$).

Table 2. Kruskal-Wallis Test Result for the Comparison of Case Distribution

Statistics	Cases		
	Confirmed	Recovered	Death
Test Statistics	175.204	207.401	170.711
Asymptotic Sig. (2-sided test)	0.018	0.000	0.031*

Degree of freedom = 138

Source: Own Elaboration

The confirmed, recovery, and death cases differ significantly from one country to another, this could be attributed to various factors, from demographics of the infected people to population density as well as the response actions taken by the countries. Such differences will lead to different outcomes.

3.2 Spatial Autocorrelation Across Incidence Locations

Testing for the spatial autocorrelation, the Global Moran's I observed is 0.002 (Expected index of -004) with a Z-score of 0.49 (p = 0.620), indicating that the pattern of the distribution of confirmed cases is not significantly different than random. This indicated that the distribution of the confirmed cases across the world appears to be random.

For the recovery, the Moran's I index value of 0.0098 was recorded (Expected value = -0.004) with a Z-score of 4.009 (p = 0.000). Thus, indicating that this is a less than 1% likelihood that this clustered pattern could be the result of random chance. The recovery count is not random, indicating spatial autocorrelation in the recovery rate.

The number of death was also tested for spatial autocorrelation, and the result shows as indicated by the observed Moran's I (-0.01) compared to the Expected value (-0.004) that this is not different from random with a Z-score of -0.401; the pattern does not appear to be statistically significantly different from random.

The Local spatial autocorrelation test was carried to identify local clusters and outliers for the COVID-19 incidences. The inverse distance conceptualisation of spatial relationship was adopted using Euclidean distance and 4,475km was identified as the default neighbourhood search distance threshold. The result for confirmed cases shows that there are 9 locations of statistical importance concerning confirmed cases (Table 3) - 7 clusters (High-High and Low-Low Clusters) while there are 2 outlier locations (Low-High).

Table 3. Cluster and Outlier Locations for Confirmed Cases of COVID-19

OBJECTID	Country	Province/State	Confirmed	Cluster Type
235	US	Ohio	13	LL
236	US	Minnesota	14	LL
237	US	New Jersey	29	LL
238	Venezuela		2	LL
44	China	Chongqing	576	LH
110	Holy See	Vatican	1	LH
51	China	Henan	1273	HH
53	China	Jiangxi	935	HH
71	China	Hunan	1018	HH

LL-Low-Low, HH – High-High, LH – Low-High

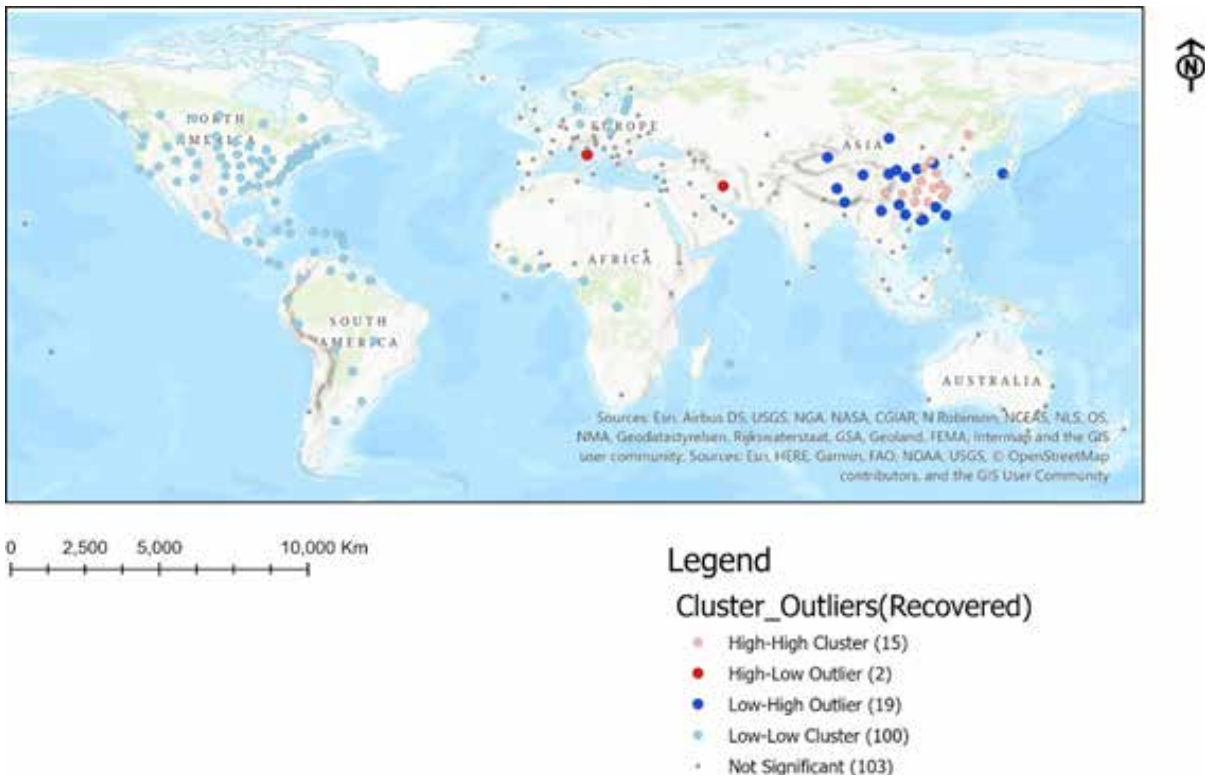
Source: Own Elaboration

The cases High-High clusters were all found in China (Table 3) – Hunan, Jiangxi, and Henan. One of the outliers (Low value surrounded by high values) was found in China (Chongqing) and the other in Rome (See of Rome), these represent usually a low number of confirmed cases amid of a high number of confirmed cases. Four locations were found to belong to a statistically significant Low-Low cluster, with 3 in the USA and 1 in South America (Venezuela). All the other locations were not found to be statistically significant. Despite global analysis showing no spatial autocorrelation, the local analysis revealed there are local patterns of spatial autocorrelation for confirmed cases.

For the cases where recoveries have been recorded, the results of the local spatial autocorrelation are presented in Figure 1. There are 100 locations identified as Low-Low clusters. These locations spread across (a) North and South America; (b) Northern Europe; and (c) West Africa. The High-High cluster is made up of 15 locations across China (Shandong,

Chongqing, Jiangsu, Heilongjiang, Hubei, Henan, Shanghai, Jiangxi, Guangdong, Beijing, Anhui, Sichuan, Zhejiang, Hunan, and Hebei).

Figure 1. Distribution of Cluster and Outlier Locations for COVID-19 Recovery Cases

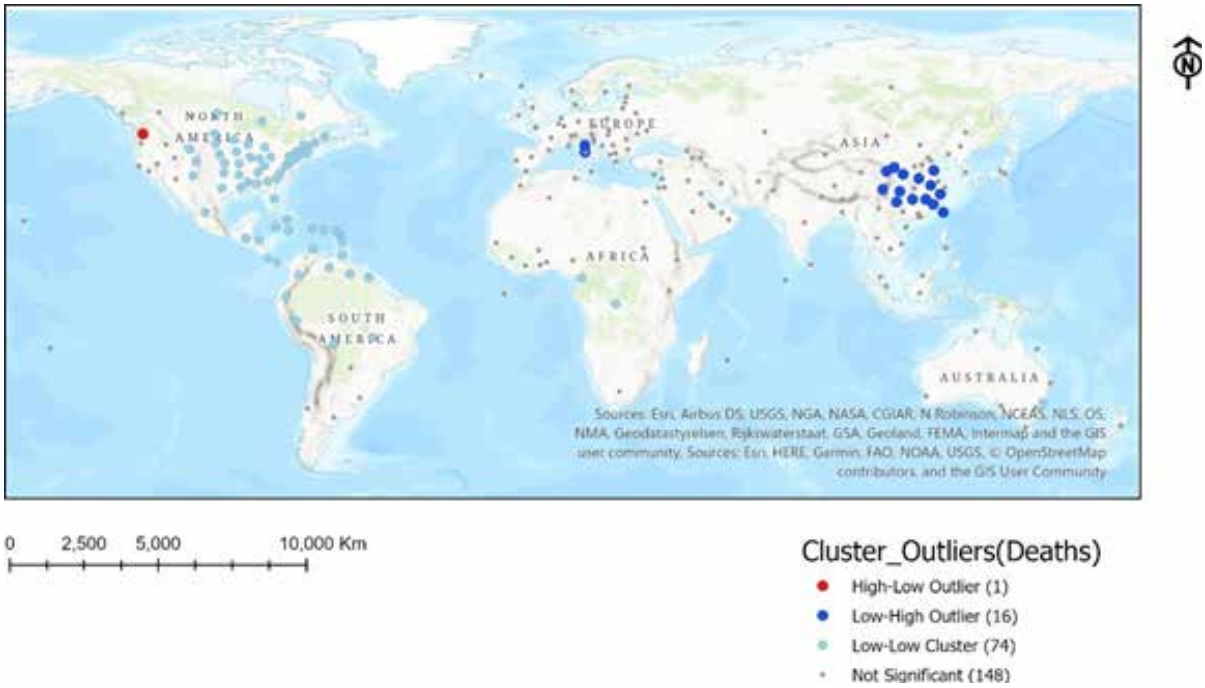


Source: Own Elaboration

Two locations (Iran and Italy) were identified as High-Low outlier i.e. those locations have unusually high recovery counts amid low recovery counts. Nineteen (19) locations displayed attributes of Low-High outlier, spread across Mostly across China and other places such as Bhutan, Holy Sea, Japan, Mongolia, and Taiwan. Evidently, most of these outliers are also located within China.

Analysis of local spatial autocorrelation for deaths shows that there is only one cluster type (Figure 2). The Low-Low clusters could be found across countries in North and South America, and Central Africa. The only High-Low outlier was found in Washington State on the east coast of the USA. All the Low-High outliers were located across China, Taiwan, and Italy (Holy See). Generally, most of the locations did not show any indication of spatial autocorrelation.

Figure 2. Distribution of Cluster and Outlier Locations for COVID-19 Death Cases



Source: Own Elaboration

3.3 Relationship between Climatic Variable Proxies and COVID-19 Cases

The correlation analysis results (Table 4) shows that monthly average water vapour content in the atmosphere - WV19_12, WV20_01 and WV20_02 (December 2019, January and February 2020, respectively) have a statistically significant ($P<0.05$) negative relationship with confirmed and recovery cases. However, the relationships were all found to be weak with R-square ranging between 0.043 and 0.054 for confirmed cases and between 0.021 and 0.029 for recovery cases, clearly indicating that a very small percentage of the variation in the confirmed and recovery cases can be explained by the average amount of water vapour across the three months. There was no statistically significant relationship between the death cases and the water vapour content of the atmosphere for the three months.

Confirmed counts were found to be negatively related to the Land surface temperature for the night in December 2019 (LSTN_1912) and January 2020 (LSTN_2001). The relationship identified was statistically significant albeit being very weak (Rho ranges between - 0.152 and -0.142). This indicated that the nighttime surface temperatures have very little explanatory power for predicting confirmed case of COVID-19. Death and recovery cases show no statistically significant relationship with the nighttime surface temperatures across the months.

The daytime land surface temperatures were also compared with death and recovery cases, and the result indicated no statistically significant relationship. There is a weak but negative correlation between these case attributes and the surface temperatures (LSTD_1912 - December 2019, LSTD _2001 - January and LSTD_2002 - February 2020). The Rho values indicated a very weak explanatory power with R-Square value ranging between 0.022 (LSTD_2002) and 0.026 (LSTD_1912).

Table 4. Correlation Analysis Result for Climatic Variable Versus COVID-19 Cases Count

Climatic Variables	Statistics	Confirmed	Recovered	Deaths
WV19_12	Correlation Coefficient	**-0.233****	**-0.169***	-0.118
	Sig. (2-tailed)	0.000	0.010	0.074
	N	231	231	231
WV20_01	Correlation Coefficient	**-0.208****	**-0.156***	-0.095
	Sig. (2-tailed)	0.001	0.017	0.148
	N	233	233	233
WV20_02	Correlation Coefficient	**-0.208****	**-0.145***	-0.076
	Sig. (2-tailed)	0.001	0.026	0.245
	N	236	236	236
LSTN_1912	Correlation Coefficient	**-0.152***	-0.081	-0.057
	Sig. (2-tailed)	0.030	0.251	0.419
	N	204	204	204
LSTN_2001	Correlation Coefficient	**-0.142***	-0.094	-0.061
	Sig. (2-tailed)	0.043	0.183	0.387
	N	204	204	204
LSTN_2002	Correlation Coefficient	-0.129	-0.026	-0.024
	Sig. (2-tailed)	0.065	0.713	0.736
	N	204	204	204
LSTD_1912	Correlation Coefficient	**-0.161***	-0.015	-0.051
	Sig. (2-tailed)	0.021	0.831	0.470
	N	204	204	204
LSTD_2001	Correlation Coefficient	**-0.162***	-0.039	-0.041
	Sig. (2-tailed)	0.021	0.577	0.564
	N	204	204	204
LSTD_2002	Correlation Coefficient	**-0.147***	0.036	-0.033
	Sig. (2-tailed)	0.036	0.611	0.635
	N	204	204	204

** - Correlation is significant at the 0.01 level (2-tailed)

* - Correlation is significant at the 0.05 level (2-tailed)

Source: Own Elaboration

3.4 Multivariate Cluster Analysis

To examine the natural groupings that may exist within the COVID-19 dataset, the two-step cluster analysis was carried out. The internal consistency of the members within the groups identified was examined using the Silhouette measure of cohesion and separation (Rousseeuw, 1987). For this measure, clustering with silhouette measure value greater than 0.5 is considered to have a good cluster quality while less than 0.5 but greater than 0.2 is considered fair.

Table 5. Summary of the Auto-clustering Diagnostic for the Two-step Clustering Analysis

Number of Clusters	Schwarz's Bayesian Criterion (BIC)	BIC Change[a]	Ratio of BIC Changes[b]	Ratio of Distance Measures[c]
1	528.344			
2	104.806	-423.538	1.000	16.251
3	109.581	4.775	-0.011	5.521
4	137.353	27.772	-0.066	1.402
5	166.584	29.231	-0.069	4.037
6	198.544	31.960	-0.075	1.782
7	230.898	32.354	-0.076	1.318
8	263.375	32.476	-0.077	2.054
9	296.047	32.673	-0.077	1.868
10	328.806	32.759	-0.077	1.890
11	361.612	32.806	-0.077	1.937
12	394.444	32.832	-0.078	1.119
13	427.278	32.834	-0.078	1.047
14	460.114	32.836	-0.078	1.319
15	492.955	32.841	-0.078	1.146

[a] The changes are from the previous number of clusters in the table; [b] The ratios of changes are relative to the change for the two-cluster solution; [c] The ratios of distance measures are based on the current number of clusters against the previous number of clusters.

Source: Own Elaboration

The summary for the auto-clustering operation is presented in Table 5. The result indicated that two clusters are the optimal number of clusters for the dataset (incidence counts). This is because the highest ratio of distance measure (Table 5) is greatest at 16.251 when the number of clusters is 2 compared to 5.521 for 3 clusters and 1.402 for 4 clusters.

Based on this result the cluster distribution showed that 98.7% of the locations belong to Cluster 1, while 1.3% (3 locations) belong to Clusters 2. This clustering was found to exhibit good cluster quality (internal consistency) with a silhouette measure of 1.0. Examination of the variable importance (Table 6) for the grouping exercise shows that importance is in the order Deaths > Confirmed > Recovery. In the grouping of the locations, the most important criterion was the number of deaths recorded while the least important is the recovery counts.

Table 6. Variable Importance Result

Nodes	Importance
Recovered	0.4243
Confirmed	0.7866
Deaths	1

Source: Own Elaboration

Looking at the centroid of the natural grouping (Table 7), Cluster 2 has high mean values for deaths, confirmed, and recovery cases when compared to Cluster 1. The members of Clusters 2 area Hubei (China), Iran, and Italy while all the other locations belonging to Cluster 2.

Table 7. Centroid Attributes for the Natural Groupings

Cluster	Confirmed		Recovered		Deaths	
	Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
1	206	737	61	202	2	12
2	32271	30921	19114	29307	1618	1316
Combined	608	4623	300	3432	23	217

Source: Own Elaboration

3.5 Discussion

The most affected country could be found around the northern hemisphere (around a range of latitude) with most of the confirmed cases occurring at the epicenter of the disease (China). The recovery was also high among the countries and locations where the confirmed cases are high as well as some places where the infection is low. The highest death rate among the high infection countries was recorded in Italy. SARS Corona Virus has caused a major pandemic in this millennium and its origin is China - Guangdong Province (Ksiazek et al., 2003), so it is not surprising that a new strain emerged from China. This rate of infection from China the top 10 countries with a high number of confirmed cases indicated a point-source outbreak with a high number at the source (Ksiazek et al., 2003) and as one moves away from the source the number thins out.

Examination of aggregated counts of all cases across countries shows that there are statistically significant variations in the count of confirmed, recovery, and death cases for the COVID-19 infection. This could be attributed to the distance away from the source, demography, preventive and response measure across the countries, thus the manifestation of the disease is impacted by various constraints (Hagerstrand, 1968) while culture and tradition influenced its emergence (Cheng et al., 2007).

At the global level, spatial autocorrelation showed up only for the recovery rate as such there is a spatial process at those places across the globe. At the local scale, spatial autocorrelation emerged for all the infection attributes, thus indicating that there is a spatial process taking place. There are local clusters and outliers across the globe for confirmed, recovery, and death cases, thus showing that infection, recovery, and deaths have a spatial dimension to their occurrence. Large populations near the epicenter and nearby locations to confirmed cases are potentially still at risk. Since the disease is highly transmissible, contact with a large population may result in reemergence even after this current wave.

While the disease has been more prominent across a region, the result from the correlation analysis shows that there is a weak negative but statistically significant association between the cases and the climatic variables. This is in agreement with the findings of Tan et al. (2005) which identified 16°C to 28°C as the optimum environmental temperature associated with the SARs virus outbreak in 2003-2004. Essentially, as land surface temperature and water vapour content increases, confirmed case decreases. But it should be noted that these associations are weak and due to the coarseness (low spatial resolution) local variations may not be properly captured. However, with the outbreak coinciding with the previous episodes at the beginning of the millennium, it is logical to deduce that the period is conducive for the virus to grow and spread. This is in partial agreement with the findings of Tamerius et al. (2013) indicating that seasonal influenza thrives during the “cold-dry” and “humid-rainy” periods.

Two types of clusters emerged with the attributes of the incidence at different locations across the globe. Thus, locations could be distinguished most importantly based on the number of death and the number of confirmed cases. There is a natural group with a high

number of deaths, confirmed cases, and recovery cases with very few members - 3 locations (Hubei China, Italy, and Iran) indicating the most impacted so far. While the other locations (98.7%) are characterised by comparatively lower counts of confirmed, recovery, and death cases.

4. CONCLUSION

Based on the collated data, we can conclude that there is a latitudinal dimension to the prevalence of COVID-19, but since this is an ongoing event, we will have to wait and see how the event pans out. As such for now, there is evidence showing that the most impacted locations in terms of confirmed cases are around the northern hemisphere.

From the correlation analysis is it possible to conclude that there is a weak negative association between confirmed cases count and the selected climatic variables. The weakness could be attributed to the coarseness (low spatial resolution) of the climatic data. However, there is evidence of association with climatic conditions, this is also reflected in the latitudinal dimension highlighted earlier.

Spatial analyses revealed the level of spatial association across event locations. From the results, we can conclude that patterns formed at the global scale indicate majorly a random formation. However local level analysis shows that there are cluster and outlier locations across the globe, thus identifying locations for further studies and investigations.

The multivariate cluster analysis (non-spatial) reveals the natural grouping for the locations with the incidence of the disease, as such, we can conclude that there are two categories of location based on their counts of confirmed, death, and recovery. Essentially, the identification of area with differences could reveal underlying factors creating the observed figures.

The study has showcased the relevance of spatial structure in understanding the emergence phase of COVID-19 pandemic, while also exploring the relationship between the climatic variables and confirmed cases. With the pandemic is still evolving, our current conclusions might only be relevant for this initial phase of the disease. However, from the foregoing, it is recommended that spatial structure and organisation be considered in studying the evolution of the COVID-19 disease to help mitigate the spread and re-emergence of the disease. Consequently, further research needs to consider the spatial structure in the study of the evolution of the disease.

REFERENCES

Anwar, M. Y., Warren, J. L., & Pitzer, V. E. (2019). Diarrhea Patterns and Climate: A Spatiotemporal Bayesian Hierarchical Analysis of Diarrheal Disease in Afghanistan. *Am J Trop Med Hyg, 101*(3), 525-533. doi:10.4269/ajtmh.18-0735

Buckee, C. O., Tatem, A. J., & Metcalf, C. J. E. (2017). Seasonal Population Movements and the Surveillance and Control of Infectious Diseases. *Trends in Parasitology, 33*(1), 10-20. doi:https://doi.org/10.1016/j.pt.2016.10.006

Ceccato, V., & Uittenbogaard, A. C. (2014). Space–Time Dynamics of Crime in Transport Nodes. *Annals of the Association of American Geographers, 104*(1), 131-150. doi:10.1080/00045608.2013.846150

Centre for Disease Control. (2020). How Coronavirus Spreads. Retrieved from https://www.cdc.gov/coronavirus/2019-ncov/prepare/transmission.html

Chen, L., Liu, W., Zhang, Q., Xu, K., Ye, G., Wu, W., Sun, Z., Liu, F., Wu, K., Zhong, B., Mei, Y., Zhang, W., Chen, Y., Li, Y., Shi, M., Lan, K., & Liu, Y. (2020). RNA based mNGS approach identifies a novel human coronavirus from two individual pneumonia cases in 2019 Wuhan outbreak. *Emerg Microbes Infect, 9*(1), 313-319. doi:10.1080/22221751.2020.1725399

Cheng, V. C. C., Lau, S. K. P., Woo, P. C. Y., & Yuen, K. Y. (2007). Severe Acute Respiratory Syndrome Coronavirus as an Agent of Emerging and Reemerging Infection. *Clinical Microbiology Reviews, 20*(4), 660. doi:10.1128/CMR.00023-07

Cohn, S. K., Jr. (2002). The Black Death: End of a Paradigm. *The American Historical Review, 107*(3), 703-738. doi:10.1086/ahr/107.3.703

Coletti, P., Poletto, C., Turbelin, C., Blanchon, T., & Colizza, V. (2018). Shifting patterns of seasonal influenza epidemics. *Scientific Reports, 8*(1), 12786. doi:10.1038/s41598-018-30949-x

Di Gennaro, F., Pizzol, D., Cebola, B., Stubbs, B., Monno, L., Saracino, A., Luchini, C., Solmi, M., Segafredo, G., Putoto, G., & Veronese, N. (2017). Social determinants of therapy failure and multi drug resistance among people with tuberculosis: A review. *Tuberculosis, 103*, 44-51. doi:https://doi.org/10.1016/j.tube.2017.01.002

Dowell, S. F. (2001). Seasonal variation in host susceptibility and cycles of certain infectious diseases. *Emerging Infectious Diseases, 7*(3), 369-374. doi:10.3201/eid0703.010301

Dowell, S. F., & Ho, M. S. (2004). Seasonality of infectious diseases and severe acute respiratory syndrome–what we don't know can hurt us. *The Lancet Infectious Diseases, 4*(11), 704-708. doi:https://doi.org/10.1016/S1473-3099(04)01177-6

ESRI. (2020). ArcGIS Pro (Version 2.5). Redlands, CA: Environmental Systems Research Institute.

Fan, C., & Myint, S. (2014). A comparison of spatial autocorrelation indices and landscape metrics in measuring urban landscape fragmentation. *Landscape and Urban Planning, 121*(0), 117-128. doi:http://dx.doi.org/10.1016/j.landurbplan.2013.10.002

Fan, V. Y., Jamison, D. T., & Summers, L. H. (2017). The loss from pandemic influenza risk. In D. T. Jamison, H. Gelband, S. E. Horton, P. K. Jha, R. Laxminarayan, C. N. Mock, & R. Nugent (Eds.), *Disease Control Priorities: Improving Health and Reducing Poverty* (3rd ed., Vol. 9, pp. 347-358). Washington DC: The International Bank for Reconstruction and Development/The World Bank.

Fares, A. (2013). Factors influencing the seasonal patterns of infectious diseases. *International Journal of Preventive Medicine, 4*(2), 128-132. Retrieved from https://pubmed.ncbi.nlm.nih.gov/23543865

Grassly, N. C., & Fraser, C. (2006). Seasonal infectious disease epidemiology. *Proceedings. Biological Sciences, 273*(1600), 2541-2550. doi:10.1098/rspb.2006.3604

Griffith, D. A. (1987). Spatial autocorrelation. *A Primer. Washington DC: Association of American Geographers*.

Grineski, S. E., & Collins, T. W. (2010). Environmental injustices in transnational context: urbanization and industrial hazards in El Paso/Ciudad Ju rez. *Environment and Planning A, 42*(6), 1308-1327. Retrieved from http://www.envplan.com/abstract.cgi?id=a42392

Hagerstrand, T. (1968). *Innovation Diffusion as a Spatial Process*. Chicago: University of Chicago Press.

IBM. (2015). IBM SPSS Statistics (Version 23). Armonk, New York: IBM Corporation.

IBM. (2016). Knowledge Centre - Two Step Cluster Analysis. Retrieved from http://www.ibm.com/support/knowledgecenter/SSLVMB_21.0.0/com.ibm.spss.statistics.help/idh_twostep_main.htm

Iskander, J., Strikas, R. A., Gensheimer, K. F., Cox, N. J., & Redd, S. C. (2013). Pandemic influenza planning, United States, 1978-2008. *Emerging Infectious Diseases, 19*(6), 879-885. doi:10.3201/eid1906.121478

Kesson, A. M. (2011). Measles. In E. C. Jong & D. L. Stevens (Eds.), *Netter's Infectious Diseases* (pp. 42-46). Philadelphia: Elsevier Health Sciences.

Kim, C. W., Phipps, T. T., & Anselin, L. (2003). Measuring the benefits of air quality improvement: a spatial hedonic approach. *Journal of Environmental Economics and Management, 45*(1), 24-39.

Ksiazek, T. G., Erdman, D., Goldsmith, C. S., Zaki, S. R., Peret, T., Emery, S., Tong, S., Urbani, C., Comer, J. A., Lim, W., Rollin, P. E., Dowell, S. F., Ling, A.-E., Humphrey, C. D., Shieh, W.-J., Guarner, J., Paddock, C. D., Rota, P., Fields, B., DeRisi, J., Yang, J.-Y., Cox, N., Hughes, J. M., LeDuc, J. W., Bellini, W. J., & Anderson, L. J. (2003). A Novel Coronavirus Associated with Severe Acute Respiratory Syndrome. *New England Journal of Medicine, 348*(20), 1953-1966. doi:10.1056/NEJMoa030781

Lawal, O. (2015). Geodemographic Analysis of Age Dependencies in Nigeria. *Port Harcourt Journal of Social Sciences, 6*(1&2), 115-132.

Lawal, O. (2017). Mapping Economic Potential Using Spatial Structure of Age Dependency and Socio-economic factors. *African Journal of Applied and Theoretical Economics, Special Edition*(November), 32-49.

Liu, Y., Bi, L., Chen, Y., Wang, Y., Fleming, J., Yu, Y., Gu, Y., Liu, C., Fan, L., Wang, X., & Cheng, M. (2020). Active or latent tuberculosis increases susceptibility to COVID-19 and disease severity. *medRxiv*, 2020.2003.2010.20033795. doi:10.1101/2020.03.10.20033795

Longley, P. A., & Tobón, C. (2004). Spatial Dependence and Heterogeneity in Patterns of Hardship: An Intra-Urban Analysis. *Annals of the Association of American Geographers, 94*(3), 503-519. doi:10.1111/j.1467-8306.2004.00411.x

Lowen, A. C., & Steel, J. (2014). Roles of humidity and temperature in shaping influenza seasonality. *Journal of Virology, 88*(14), 7692-7695. doi:10.1128/JVI.03544-13

Martinez, M. E. (2018). The calendar of epidemics: Seasonal cycles of infectious diseases. *PLOS Pathogens, 14*(11), e1007327. doi:10.1371/journal.ppat.1007327

Morony, M. G. (2007). 'For Whom Does the Writer Write?': The First Bubonic Plague Pandemic According to Syriac Sources. In L. K. Little (Ed.), *Plague and the End of Antiquity: The Pandemic of 541-750* (Reprint ed., pp. 59-86). New York: Cambridge University Press.

Ohadike, D. C. (1991). Diffusion and physiological responses to the influenza pandemic of 1918–19 in Nigeria. *Social Science & Medicine, 32*(12), 1393-1399. doi:https://doi.org/10.1016/0277-9536(91)90200-V

Pai, M. (2020). COVID-19 Coronavirus And Tuberculosis: We Need a Damage Control Plan. Retrieved from https://www.forbes.com/sites/madhukarpai/2020/03/17/covid-19-and-tuberculosis-we-need-a-damage-control-plan/#77668cbc295c

Pascual, M., & Dobson, A. (2005). Seasonal patterns of infectious diseases. *PLoS Medicine, 2*(1), e5-e5. doi:10.1371/journal.pmed.0020005

Pauly, M. D., Procario, M. C., & Lauring, A. S. (2017). A novel twelve class fluctuation test reveals higher than expected mutation rates for influenza A viruses. *eLife, 6*, e26437. doi:10.7554/eLife.26437

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics, 20*, 53-65. doi:http://dx.doi.org/10.1016/0377-0427(87)90125-7

Sands, P., El Turabi, A., Saynisch, P. A., & Dzau, V. J. (2016). Assessment of economic vulnerability to infectious disease crises. *The Lancet, 388*(10058), 2443-2448. doi:10.1016/S0140-6736(16)30594-3

Smith, W., Westwood, M. A., Westwood, J. C. N., & Belyavin, G. (1951). Spontaneous mutation of influenza virus A during routine egg passage. *British Journal of Experimental Pathology, 32*(5), 422-432. Retrieved from https://pubmed.ncbi.nlm.nih.gov/14886504

Tamerius, J. D., Shaman, J., Alonso, W. J., Bloom-Feshbach, K., Uejio, C. K., Comrie, A., & Viboud, C. (2013). Environmental predictors of seasonal influenza epidemics across temperate and tropical climates. *PLOS Pathogens, 9*(3), e1003194-e1003194. doi:10.1371/journal.ppat.1003194

Tan, J., Mu, L., Huang, J., Yu, S., Chen, B., & Yin, J. (2005). An initial investigation of the association between the SARS outbreak and weather: with the view of the environmental temperature and its variation. *Journal of Epidemiology and Community Health, 59*(3), 186-192. doi:10.1136/jech.2004.020180

TB Alliance. (2020). Global Pandemic. Retrieved from https://www.tballiance.org/why-new-tb-drugs/global-pandemic

The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team. (2020). The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China. *China CDC Weekly, 2*, 1-10. doi:10.3760/cma.j.issn.0254-6450.2020.02.003

Voigt, E. A., Kennedy, R. B., & Poland, G. A. (2016). Defending against smallpox: a focus on vaccines. *Expert Review of Vaccines, 15*(9), 1197-1211. doi:10.1080/14760584.2016.1175305

Wangdi, K., & Clements, A. C. A. (2017). Spatial and temporal patterns of diarrhoea in Bhutan 2003–2013. *BMC Infectious Diseases, 17*(1), 507. doi:10.1186/s12879-017-2611-6

World Health Organization. (1980). *The Global Eradication of Smallpox*. Geneva, Switzerland.

World Health Organization. (2009). The WHO Pandemic Phases. In *Pandemic influenza preparedness and response: a WHO guidance document* (pp. 24-27). Geneva: World Health Organization.

World Health Organization. (2018). Influenza (Seasonal). *Fact Sheet.* Retrieved from https://www.who.int/news-room/fact-sheets/detail/influenza-(seasonal)

World Health Organization. (2020a). Director-General's remarks at the media briefing on 2019-nCoV on 11 February 2020. *Spechees.* Retrieved from https://www.who.int/dg/speeches/detail/who-director-general-s-remarks-at-the-media-briefing-on-2019-ncov-on-11-february-2020

World Health Organization. (2020b). *Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19).* Geneva.

World Health Organization. (2020c). WHO Director-General's opening remarks at the media briefing on COVID-19 - 13 March 2020. Retrieved from https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-mission-briefing-on-covid-19-13-march-2020

Worldometer. (2020). Coronavirus Cases. Retrieved from https://www.worldometers.info/coronavirus/

Wu, X., Lu, Y., Zhou, S., Chen, L., & Xu, B. (2016). Impact of climate change on human infectious diseases: Empirical evidence and human adaptation. *Environment International, 86*, 14-23. doi:https://doi.org/10.1016/j.envint.2015.09.007

Yu, D., Wei, Y. D., & Wu, C. (2007). Modeling spatial dimensions of housing prices in Milwaukee, WI. *Environment and Planning B: Planning and Design, 34*(6), 1085-1102. Retrieved from http://www.envplan.com/abstract.cgi?id=b32119

Zou, Y. (2014). Analysis of spatial autocorrelation in higher-priced mortgages: Evidence from Philadelphia and Chicago. *Cities, 40, Part A*(0), 1-10. doi:http://dx.doi.org/10.1016/j.cities.2014.04.003