

How is Coronavirus distributed in the world? A Spatial-Temporal Assessment Using Geographic Information System Approach

Sayede Shadi Nazari¹, Solmaz Norouzi^{2*}, Mohammad Asghari Jafar-abadi³

1. Department of Zanjan University of Medical Sciences, Zanjan, Iran
2. Biostatistics Department of Zanjan University of Medical Sciences, Zanjan, Iran
3. Road Traffic Injury Research Center, Faculty of Health, Tabriz University of Medical Sciences, Tabriz, Iran

Article Type:
Original Article

***Correspondence:**

Solmaz Norouzi,
Biostatistics Department of
Zanjan University of
Medical Sciences, Zanjan,
Iran
snorouzibiostatistics@gmail.com



Abstract

Background and objective: Prevalence and the spread of novel Coronavirus (2019-ncov) cause significant life and financial destruction worldwide and is the cause of severe respiratory infection in humans. The present study briefly reviews the latest information on how the virus is distributed around the world. The main question of the study are: 1- In which geographic regions of the world is the Coronavirus more concentrated? 2- Is the distribution of the Coronavirus geographically stable?

Methods: To answer these questions, we first began collecting and studying the available scientific resources. The required data was obtained from a daily report of confirmed, recovered, and deaths by the Coronavirus separated by state which was collected from January 22, 2020 to Jun 19, 2020. Based on analyzing available patterns in spatial statistics tool in ArcGIS and geostatistical models, we examined how the Coronavirus was distributed around the world.

Results: The spread of the disease is increasing all over the world. Using the results of Map 1, it is seen that the spread of Corona virus has a trend and starts in China and then spreads to the Middle East, Europe and the United States in a linear manner. The results also show that the prevalence of mortality is higher than that of recovery. Central mean and median for all types (Confirmed, Recovered and death) are close to each other. Death mean and median was close to Western countries and Recovered mean and median was close to Eastern countries, while confirmed mean and median was located in the center.

Conclusion: Based on spatial statistics tool in ArcGIS and geostatistical models, we examined how the Coronavirus was distributed around the world. Our results showed that the spread of Corona virus had a trend and started in China and then spread to the Middle East, Europe and the United States in a likely linear manner.

Keywords: COVID-19, pandemic, ArcGIS, Spatial Statistics Tools

Introduction

In late December 2019, a number of patients were admitted to hospitals in China's Hubei Province with an initial diagnosis of pneumonia and symptoms such as fever, cough and shortness of breath for unknown reasons. After review, it was determined that the patients were epidemiologically related to a wholesale market for seafood and wet animals in Hubei, China (1, 2). Preliminary reports of the spread and pathogenicity of the new corona virus (COVID-19, named by the World Health Organization (WHO) on Feb 11, 2020), predicted the possible outbreak of COVID19. (3, 4). The WHO announced the outbreak of COVID-19 to be a pandemic on March 11, 2020, with the announcement of the director-general, "This is not just a public health crisis, this is a crisis that will touch every sector. So every sector and every individual must be involved in the fights," (5)

The source of the infection that has been observed so far are mainly patients with coronavirus. But you need to know that a person who is asymptomatic can also be a source of infection (6). The virus can be passed from person to person (7), prevalence through respiratory tract secretions and contact (with surfaces or person to person) are from One of the most important causes of virus transmission. Airborne particles are also another way of transmitting the virus. The virus is generally easily transmitted during gatherings (8-10).

The COVID-19 pandemic virus is full of unknowns, and many of them (unknowns) have spatial dimensions that lead to the understanding of the phenomenon as geography and potential maps (11). Nowadays, it is very important to have timely data and extract the required information from

this data. In this regard, Geographic Information Systems (GIS) are considered as an important tool in land data management, which provides the possibility of integrating data from different sources, the possibility of extracting the required information and discovering complex relationships between different phenomena (12).

Therefore, in order to interpret the COVID-19 phenomenon as a global epidemic disease in terms of location and its geographical impact, it is necessary to use geospatial and statistical tools (13). GIS has become an important tool in analyzing and visualizing the spread of COVID-19. For instance, Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE) currently utilizes a GIS dashboard that provides live data of the world wide spatial distribution of COVID-19, including the total number of confirmed cases, mortalities, and recovered patients. A limited number of GIS-based studies have been published since the initial outbreak of COVID-19 (14).

The aim of this study is to implement, geographical and geo-spatial analysis in understanding locations and the distribution patterns of COVID-19.

Materials and Methods

This ecological study was conducted through coronavirus package provides a tidy format dataset of the 2019 Novel Coronavirus COVID-19 epidemic, reported by <https://github.com/RamiKrispin/coronavirus>. The raw data pulled from the Johns Hopkins University Center for Science and Engineering Systems (JHU CCSE) Coronavirus repository. The information reviewed daily reports of definite cases, definitively recovered and death from

coronavirus in countries affected by the disease worldwide divided by state / province from January 22, 2020 to Jun 19, 2020. Variables included location by latitude and longitude, country name, province name (in some countries), date (daily), type (confirmed, recovered and death) and the number of cases.

Inclusion and exclusion criteria

All reported cases of Corona virus during January 22, 2020 to Jun 19, 2020 were considered as inclusion criteria. Cases reported with a negative value and 0 were excluded.

Analysis method

Data were analyzed by ArcGIS software desktop version 10.2(<http://www.esri.com>) To understand the patterns and general trends in the data, the spatial statistics tool has been used. In terms of clustering, data were analyzed using the average nearest neighbor, Moran's I and multi-distance spatial cluster analysis. To measure geographical distribution and descriptive analysis, directional distribution, mean and median center and standard distance methods were used.

The average nearest neighbor method first measures the distance between the center point of each event and the center point of its nearest neighbor. If the calculated average distance is less than the average hypothetical random distribution, then it can be concluded that the distribution of the event under study in space is clustered. Moran's I spatial autocorrelation tool examines the spatial autocorrelation based on the location of two of the desired properties of geographical features. This tool shows that the distribution pattern of these geographical features,

considering the studied properties, has a clustered or scattered pattern. The multi-distance, spatial cluster analysis tool, also known as the K-Ripley function, is another very useful tool for statistically examining the spatial pattern of events in place and space. This tool shows the status of clustering events at different geographical distances. Directional distribution indicates whether the distribution of geographical features in space has been directional. Standard distance analysis measures the degree of concentration or distribution of events around the central Mean. Finally, to have an overview of the data, a prediction layer was obtained from places which there was no sample, using geostatistical analysis system. This layer was obtained using the inverse distance weighting, which is a precise method and has no presuppositions for data.

Ethical considerations

Due to the use of registration data and non-use of personal information, the ethical considerations of this study are not of concern.

Results

Table 1 shows the numerical results of the average nearest neighbor analysis. The average observed distance for confirmed, recovered, and death cases respectively were equal 80, 359.26 & 666.01 meters. However, the average value of the expected distance was calculated as 94436.7255, 106138.72 & 120603.20 respectively. The ratio of the nearest neighbor was 0.0008, 0.0034 and 0.0055, respectively. The calculated standard score was -251.55, -212.98 and -172.1, respectively.

Table 1 shows the numerical results of the Moran's I analysis. The Moran's Index for

confirmed, recovered, and death cases respectively were equal 0.78, 0.39 & 0.83. However, the expected index was calculated as -0.00006, -0.00008 & -0.00012 respectively.

The variance was 0.000001, 0.000002 and 0.000003 respectively. The calculated standard score was 673.72, 265.25 and 448.26 respectively.

Table 1: Results of the average nearest neighbor analysis

variable	average observed distance (Meters)	average value of the expected distance (Meters)	ratio of the nearest neighbor	Z score	P- Value
Confirmed	80.00	94436.7255	0.0008	-251.55	<0.0001
Recovered	359.26	106138.72	0.0034	-212.98	<0.0001
Death	666.01	120603.20	0.0055	-172.1	<0.0001

Table 2: Results of Moran's I analysis

variable	Moran's Index	Expected Index	variance	Z score	P- Value
Confirmed	0.7789	-0.00006	0.000001	673.72	<0.0001
Recovered	0.3866	-0.00008	0.000002	265.25	<0.0001
Death	0.8341	-0.00012	0.000003	448.26	<0.0001

Figure 1 shows the Average Nearest Neighbor and Moran's I results graphically. Fig1 (a), (b), (c) shows Moran's I results for confirmed, recovered and death cases respectively. Fig1

(d), (e), (f) shows Average Nearest Neighbor results for confirmed, recovered and death cases respectively.

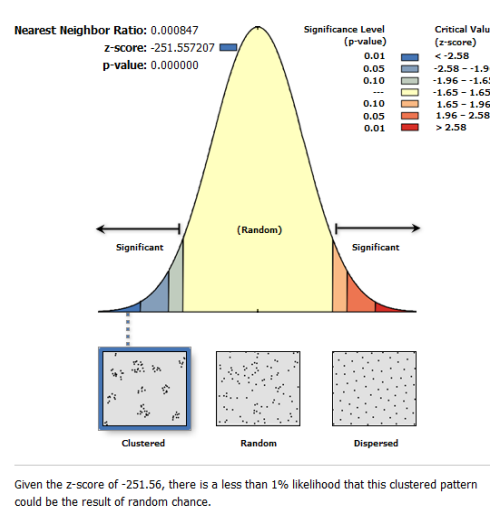
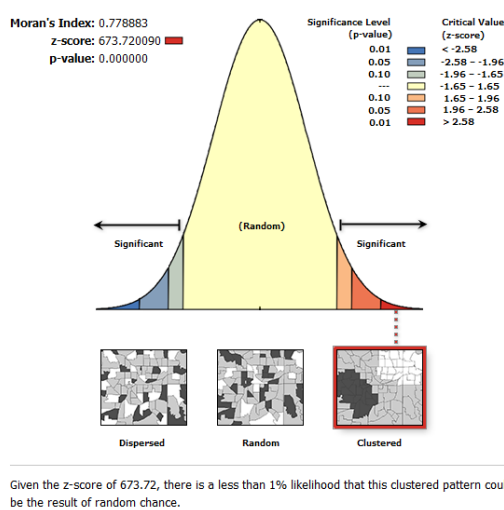


Fig1 a) Moran's Index result for Confirmed cases

Fig1 d) Nearest Neighbor result for Confirmed cases

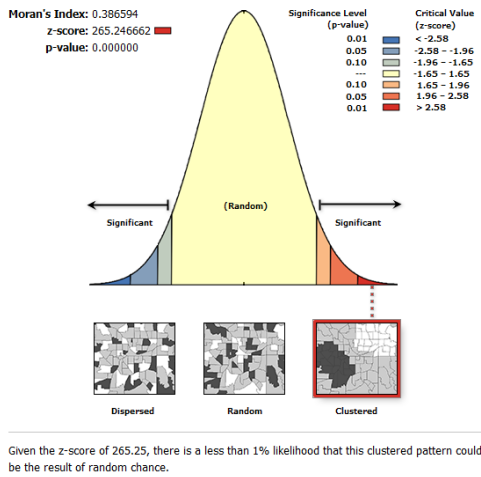


Fig1 b) Moran's Index result for Recovered cases

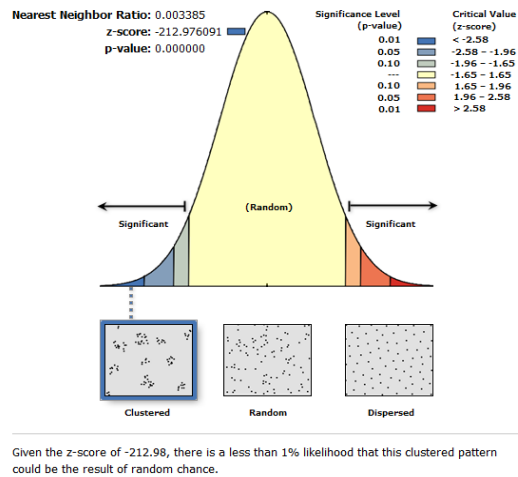


Fig1 e) Nearest Neighbor result for Recovered cases

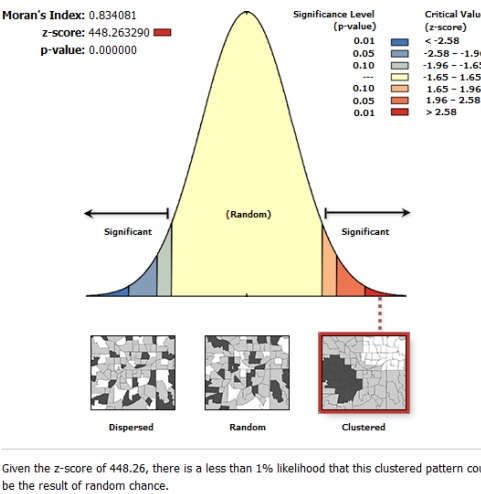


Fig1 c) Moran's Index result for death cases

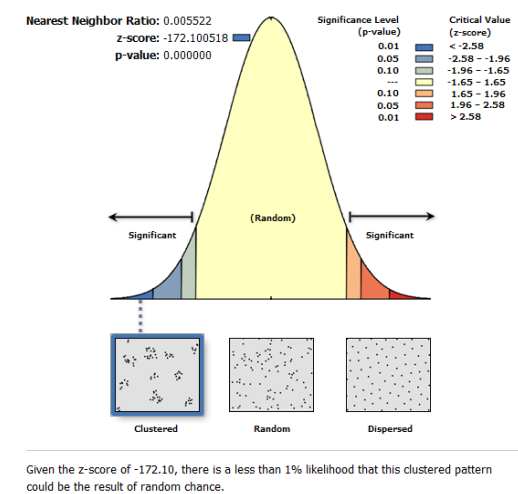


Fig1 f) Nearest Neighbor result for death cases

Figures 2 (a), (b) and (c) are the results of multi-distance, spatial cluster analysis for confirmed, recovered, and death cases, respectively.

In these figures, the blue line indicates the expected items and the red line indicates the observed items. The x-axis indicates an increase in distance

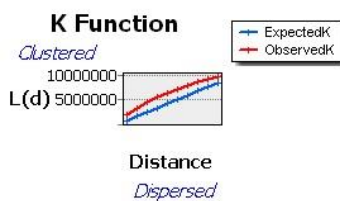


Fig2 a) Confirmed multi-distance, spatial cluster analysis

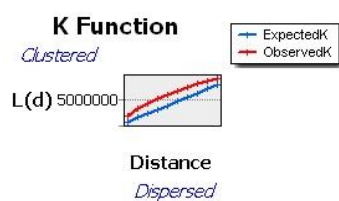


Fig2 b) Recovered multi-distance, spatial cluster analysis

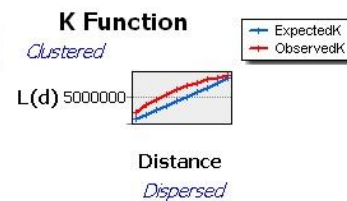
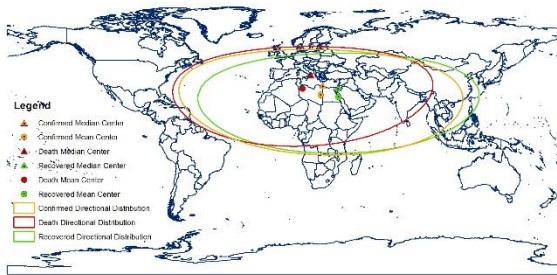


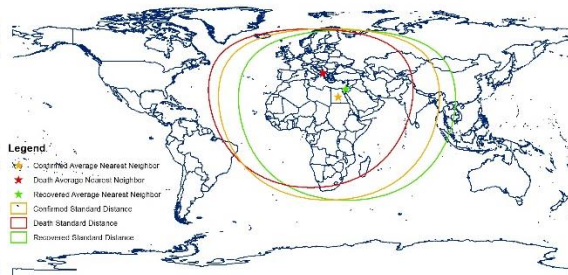
Fig2 c) Death multi-distance, spatial cluster analysis

Map 1 shows the global distribution of corona virus, a comparison of the directional distribution of confirmed, recovered and death cases. The central criteria (mean and median) are also shown.



Map1) The directional distribution of confirmed, recovered and death cases along their central points

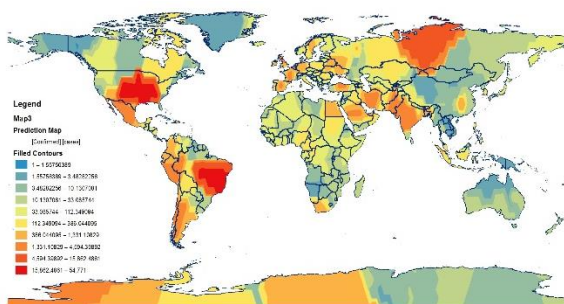
Map 2 shows the standard distance between confirmed, recovered, and death cases, as well as the average calculated by the nearest neighborhood method



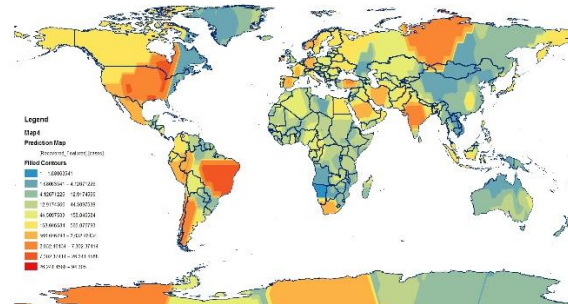
Map2) the standard distance of confirmed, recovered and death cases

Maps 3, 4, and 5 show an overview of the prevalence of confirmed, recovered, and death cases, respectively. Areas marked in red and orange indicates more infected areas, and areas marked in blue and green indicate fewer

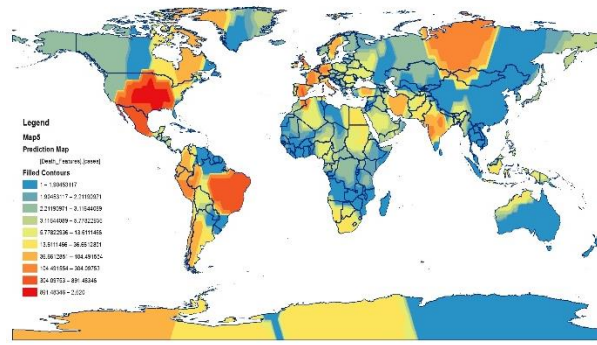
infection. These maps also show predictions for areas for which there were no samples. In fact, these maps provide an estimate of the corona virus.



Map 3) confirmed geostatistical map



Map 4) recovered geostatistical map



Map 5) death geostatistical map

Figure 3 (a) shows how the number of cases is distributed over time divided by type (confirmed, recovered and death), and

figure 3 (b) shows how the median of cases is distributed over time divided by type. (drawn by spss25)

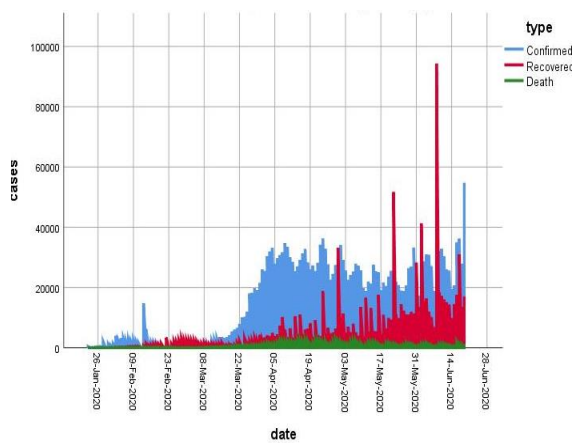


Fig3 a) number of cases over time

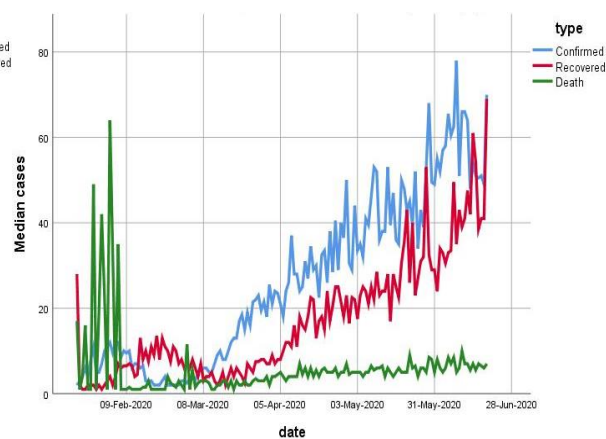


Fig3 b) median of cases over time

Discussion

This study focused on geospatial features that can predict the presence or absence of COVID-19. We examined a few spatial statistics tools to investigate the trend of coronavirus.

Our results showed that the spread of COVID-19 had a trend and started from

China and then spread to the Middle East, Europe and the United States more likely in a linear manner. The results also showed that the prevalence of mortality was higher than that of recovery. Central mean and median for all types (Confirmed, Recovered and death) were close to each other. The standard distance for confirmed, recovered and death was the same, only differing in location, which could mean that the spread of disease

growth and mortality and recovery were almost on the same level. Figures 1 and 2, as well as the results of Table 1 and 2, all suggested that the spread of Corona virus was clustered and not random.

Boulos and Geraghty (2020), presented how various GIS applications and dashboards such as JHU CSSE, WHO dashboard, Health Map, World Pop, and Epi Risk are able to provide a clear representation of the COVID-19 spread (15). Lakhani (2020) utilized GIS mapping to identify COVID-19 health care priority locations pertaining to vulnerable populations, including elderly, palliative, and disabled patients in Melbourne, Australia. The findings suggest potential improvements in quality of care in the midst of the pandemic (16). Gibson and Rush (2020), utilized GIS technology to outline dwelling boundaries to detect the probability of COVID-19 spread in Cape Town, South Africa. Their results suggest that COVID-19 spread can be reduced through social distancing measures as supported by their buffer analysis and cluster identifications (5).

Krivoruchko et al listed different statistical methods for aggregated data analysis that can be implemented within a GIS to provide a powerful set of interactive, analytical tools well suited to health studies(17). Tewara et al used spatial statistical tool in GIS to show distribution of Malaria in Cameron(18). Saxena et al highlighted the utility of GIS and spatial statistical tool in efficient processing of voluminous epidemiological data at micro level and analyzing with sound statistical base(19). A review of GIS studies presented by Chung et al , revealed some issues regarding the use of spatial statistical analyses in health field(20).

The pandemic caused by the spread of the virus has led to policies in various countries

which suit their circumstances. To better understand how coronavirus is spread and pathogenic in the future, we need to do more research on identifying a specific host, its specific mode of transmission, and risk factors for its transmission in medical centers. Appropriate treatment strategies to alleviate the disease also require future research in this area.

Conclusion

This study focused on geospatial features that can predict the presence or absence of COVID-19, because "should be expected to happen more frequently moving forward". Based on analyzing available patterns in spatial statistics tool in ArcGIS and geostatistical models, we examined how the Coronavirus was distributed around the world. The spread of the disease is increasing all over the world. Using the results, it is seen that the spread of Corona virus had a trend and started in China and then spread to the Middle East, Europe and the United States in a likely linear manner.

References

1. Lu H, Stratton CW, Tang YA-O. Outbreak of pneumonia of unknown etiology in Wuhan, China: The mystery and the miracle. 2020(1096-9071 (Electronic).
2. Bogoch, II, Watts A, Thomas-Bachli A, Huber C, Kraemer MUG, Khan K. Pneumonia of unknown aetiology in Wuhan, China: potential for international spread via commercial air travel. LID - 10.1093/jtm/taaa008.
3. Zhao S, Lin Q, Ran J, Musa SS, Yang G, Wang W, et al. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. 2020(1878-3511 (Electronic).
4. Rothan HA, Byrareddy SN. The epidemiology and pathogenesis of coronavirus disease (COVID-19)-outbreak. 2020(1095-9157 (Electronic).
5. Gibson LA-O, Rush DA-O. Novel Coronavirus in Cape Town Informal Settlements: Feasibility of Using Informal Dwelling Outlines to Identify High Risk Areas for COVID-19 Transmission From A Social Distancing Perspective. 2020(2369-2960 (Electronic).
6. Sun J, He W-T, Wang L, Lai A, Ji X, Zhai X, et al. COVID-19: Epidemiology, Evolution, and Cross-Disciplinary Perspectives. Trends in Molecular Medicine. 2020;26(5):483-95.
7. Bai Y, Yao L, Wei T, Tian F, Jin D-Y, Chen L, et al. Presumed Asymptomatic Carrier Transmission of COVID-19. JAMA. 2020;323(14):1406-7.
8. Liu T, Hu J, Kang M, Lin L, Zhong H, Xiao J, et al. Transmission dynamics of 2019 novel coronavirus (2019-nCoV). bioRxiv. 2020:2020.01.25.919787.
9. Gralinski LE, Menachery VA-O. Return of the Coronavirus: 2019-nCoV. LID - 10.3390/v12020135 [doi] LID - 135. 2020(1999-4915 (Electronic).
10. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019 .2020
11. Franch-Pardo I, Napoletano BM, Rosete-Verges F, Billa L. Spatial analysis and GIS in the study of COVID-19. A review. Science of The Total Environment. 2020;739:140033.
12. Pirnazar M, Ostad-Ali-Askari K, Eslamian S, Singh V, Dalezios N, Ghane M, et al. Change Detection of Urban Land Use and Urban Expansion Using GIS and RS, Case Study: Zanjan Province, Iran. International Journal of Constructive Research in Civil Engineering. 2018;4.
13. Lovett DA, Poots AJ, Clements JT, Green SA, Samarasundera E, Bell D. Using geographical information systems and cartograms as a health service quality improvement tool. 2014(1877-5853 (Electronic).
14. Mollalo A, Vahedi B, Rivera KM. GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. Science of The Total Environment. 2020;728:138884.
15. Kamel Boulos MN, Geraghty EM. Geographical tracking and mapping of coronavirus disease COVID-19/severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) epidemic and associated events around the world: how 21st century GIS technologies are supporting the global fight against outbreaks and epidemics. Int J Health Geogr. 2020;19(1):8.-

16. Lakhani A. Which Melbourne Metropolitan Areas Are Vulnerable to COVID-19 Based on Age, Disability, and Access to Health Services? Using Spatial Analysis to Identify Service Gaps and Inform Delivery. 2020(1873-6513 (Electronic)).

17. Krivoruchko K, Gotway CA, Zhigimont A. Statistical tools for regional data analysis using GIS. Proceedings of the 11th ACM international symposium on Advances in geographic information systems; New Orleans, Louisiana, USA: Association for Computing Machinery; 2003. p. 41–8.

18. Tewara MA, Mbah-Fongkimeh PN, Dayimu A, Kang F, Xue F. Small-area spatial

statistical analysis of malaria clusters and hotspots in Cameroon;2000-2015. BMC infectious diseases. 2018;18(1):636.

19. Saxena R, Nagpal BN, Das MK, Srivastava A, Gupta SK, Kumar A, et al. A spatial statistical approach to analyze malaria situation at micro level for priority control in Ranchi district, Jharkhand. The Indian journal of medical research. 2012;136(5):776-82.

20. Chung K, Yang D-H, Bell R. Health and GIS: Toward Spatial Statistical Analyses. Journal of medical systems. 2004;28:349-60.

How to cite:

Nazari S, Norouzi S, Asghari Jafar-abadi M. How is Coronavirus distributed in the world? A Spatial-Temporal Assessment Using Geographic Information System Approach. *Jorjani Biomedicine Journal*. 2020; 8(1): 24-33.