Caspian Journal of Environmental Sciences
Online ISSN: 1735-3866
Print ISSN: 1735-3033

# A Moran's I autocorrelation and spatial cluster analysis for identifying Coronavirus disease COVID-19 in Iraq using GIS approach

**Adil Shnawa Jaber[1], Ali Kadhim Hussein[2], Nabeel Abbood Kadhim[1], Ali Abid bojassim[3*]**

*1. Ministry of Education, Najaf Education Office, Najaf, Iraq*
*2. Basic Science Department, College of Dentistry, University of Kufa, Kufa, Najaf, Iraq*
*3. Physics Department, Faculty of Science, University of Kufa, Najaf, Iraq*

**Corresponding Author's Email: ali.alhameedawi@uokufa.edu.iq*

## ABSTRACT

Iraq is one of the states in the world, affected with coronavirus. Mapping spatial patterns analysis distribution of disease incidence and danger can be assist as a suitable tool for detecting exposures of public health concern. A geographical information system (GIS)-based methodology to examine the relationship between the reported incidence of coronavirus and spatial patterns analysis in eighteen provinces of Iraq was analyzed in 2020. So, the study was applying spatial statistics to inspect the spatial patterns and areas of clustering detection to describe the pattern of coronavirus in Iraq. In this study, local Moran's I has been applied to measure spatial distribution of coronavirus in the study area and examined how provinces were spread or clustered. Spatial patterns statistics were used to apply Moran's I test and it estimated considerable negative spatial autocorrelation of coronavirus disease incidences from 24/02/2020 to 06/04/2020. The results described spatially random clustered and spatial pattern of this disease in the study area. The study determined that the coronavirus cases were increased in the northeastern- and southwestern-side provinces of Iraq.

**Keywords:** Coronavirus; Cluster analysis; Moran's I; Spatial statistics, Iraq.
**Article type:** Research Article.

## INTRODUCTION

In December 2019, a new virus appeared (first called "Novel Coronavirus 2019-nCoV" and was later renamed to SARS-CoV-2) causing severe acute respiratory syndrome (COVID-19) in Wuhan, Hubei Province, China (Al-Kindi, *et al*., 2020), and spread rapidly to other parts of China and other countries around the world, despite China's massive efforts to contain the disease inside Hubei. Compared to 2002/2003 SARS-CoV and 2012-2014 MERS-CoV (Coronavirus Associated with Middle East Respiratory Syndrome), the coronavirus spread with amazing speed. While it took about two and a half years to catch the Coronavirus, and SARS took about 4 months, the SARS-CoV-2 novel reached that number in just 48 days. On January 30, 2020, the World Health Organization (WHO) announced that the SARS-CoV-2 new coronavirus outbreak is a public health emergency of international concern (PHEIC) (Tsou *et al*. 2005). The Coronavirus pandemic of 2020 has been spread in Iraq from February 24, 2020 in the city of Najaf, when a sample of an Iranian religious student the nationality has been examined and the result was positive for his infection with the Coronavirus associated with severe acute respiratory syndrome type 2 (SARS-CoV-2). Then, other cases of coronavirus were revealed, and there were 1,055 confirmed cases in Iraq, including 64 deaths as of April 6, 2020. GIS is a computerized system capable of displaying, merging, modeling, querying and analyzing large amounts of spatial data (Vine *et al.* 1997). In the field of health, the application of geographic information systems helps to know the spatial difference, patterns and risk factors of

Caspian Journal of Environmental Sciences, Vol. 20 No. 1 pp. 55-60 Received: Aug. 15, 2021 Revised: Sep. 18, 2021 Accepted: Dec. 04, 2021
DOI: 10.22124/CJES.2022.5392

Publisher: University of Guilan, Iran

disease and improve health service delivery. Spatially common diseases are characterized by many diseases (Briggs & Elliott 1995). Spatial analysis techniques such as spatial statistics and geographic information systems make it easier for officials to deal with spatial dispersion and guess disease outbreaks more accurately (Chaput *et al.* 2002). For the disease, it is necessary to determine the spatial and temporal characteristics of the disease and its transmission. GIS plays a very important role in disease surveillance and control (Albrecht 2007). Recently, GIS and remote sensing methods play an important role in mapping diseases such as dengue and HIV (Pathirana *et al*. 2009). GIS cannot simply pave the way to increasing our understanding of the pattern of disease spread, but also support health officials and manufacturers strategies about diverse information are simply implicit. Therefore, in the present study, a survey was planned to assess the prevalence of coronavirus in Iraq. GIS was a powerful analysis tool, does not include essential epidemiological information on humans, times and places but also coz it acts as a shared interface for centralized reporting and tracking indicators from diverse fields (e.g., epidemiologic data georeferencing) (Al-Kindi *et al.* 2020). GIS is a main tool for a lot of realization consciousness programs cooperation with pandemic diseases (Sithiprasasna *et al.* 2004). Many studies have been used spatial technology for estimating spatial and transitory characteristics (e.g., central, density, hotspots, cold spots, direction, spread, the Arabic psychological, etc.) COVID-19 (Biswas *et al.* 2020; Danon *et al.* 2020; Kang *et al*. 2020; Zhou *et al*. 2020; Bendaif *et al.* 2020; Maliki *et al.* 2020). Such a guess is significant to muzzle spatial and transitory patterns of COVID19 dispersal, estimate its track changes over time, for determining the various demographic, environmental, and socioeconomic inconstant that may speed conveyance and infection average. From a policy standpoint, this estimate is wished for aid policy makers improve their plans and strategies in a more reliable method, to take into consider spatial difference of this health massive threat. In this condition, a GIS is a wide set of spatial statistics for playing a significant function in plan the spatial and transitory patterns of COVID-19 in Iraq. Specifically, advancements in GIS technology, especially spatial modeling and information mining, have make it is likely to spread an inclusive picture of the estimably spatial hots-pots of this virus at various spatial scales (Wang *et al.* 2020; Boulos & Geraghty 2020). The main aims of this study for identifying Coronavirus disease COVID-19 in Iraq Using GIS Approach.

## MATERIALS AND METHODS

### Study Area

Iraq is located within latitude 28º57'0"N to 37º5'0"N and longitude 38º35'0"E to 48º45'0"E (Fig. 1). It covers an area of 437125.41 km$^2$ and its boundary share with five adjacent lands. Syria situated in the west, Iran on the east, Saudi in the southwest, Turkey in the north and Kuwait in the southeast. Iraq is administratively subdivided into 18 units including Baghdad (capital), Ninawa, Dahuk, Najaf, Karbala, Tam'mim, Muthanna, Babil, Diwaniyah, Dha Qar, Basrah, Salah Ad Din, Anbar, Arbil, Diyala, Sulaymaniya, Wasit and Maysan. In 2015, the Iraq's government estimated population as 37,056,169. The number of males was over 19 million, contributed by 51%, while females, over 18 million, contributed by 49% of the total population, according to the statistics.

### Data collections

The epidemiologic data utilized for this analysis was collected by the Ministry of Health of Iraq from 24/02/2020 through 06/04/2020 for all provinces. The location of each case of coronavirus recorded during this period was digitalized by the authors with the geographical information system (GIS) Arcgis10.7.1. Population values of provinces, provided by the Republic of Iraq Ministry of Planning in 2001, served as the basic unit for the analyses (Fig. 2). These have been digitalized as polygons with the occurrence factors as attributes. All data were storing and managing as a geographical database with Arcgis10.7.1.

### Spatial autocorrelation statistics

Spatial patterns can be defined as the spatial arrangement style of features on a specific surface according to their location. Spatial statistics were used in a set of accessibility studies. Both global and local spatial correlation patterns were used to analyze the spatial autocorrelation pattern. As well as analyzing global patterns, more emphasis has been placed on analyzing local patterns of spatial correlation (Talen & Anselin 1998; Apparicio *et al.* 2007; according to Sarwar *et al.* 2020). The most widely used index for spatial autocorrelation is Moran's I. The global spatial statistic Moran's I measures the global spatial correlation based on distinct geographical locations and associated feature values (Algert *et al.* 2006). Moran's I measures proximity to locations according to the similar characteristics of their features. Moran's global I measures spatial autocorrelation without distinguishing between high or low value patterns. Looking at a set of related features and attributes, Moran's global I assesses whether a pattern is clustered, dispersed, or random on a scale from -1 to +1. A negative value indicates a spatial dispersion, while a positive value indicates a spatial correlation (Moran, notes on continuous stochastic phenomena, 1950; Tsou *et al.* 2005).

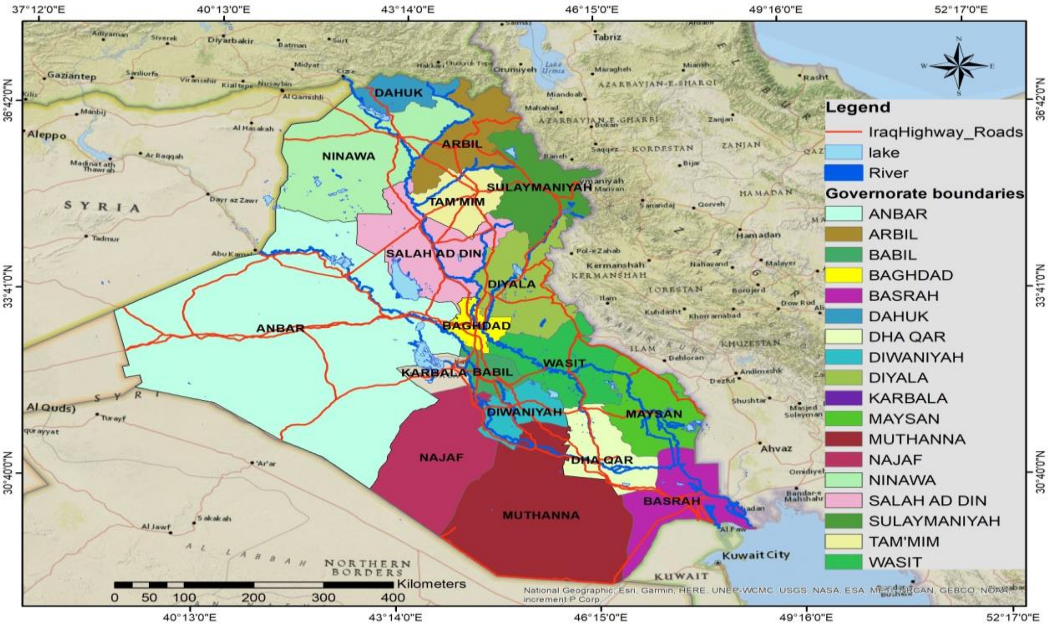


**Fig. 1.** Study area location.

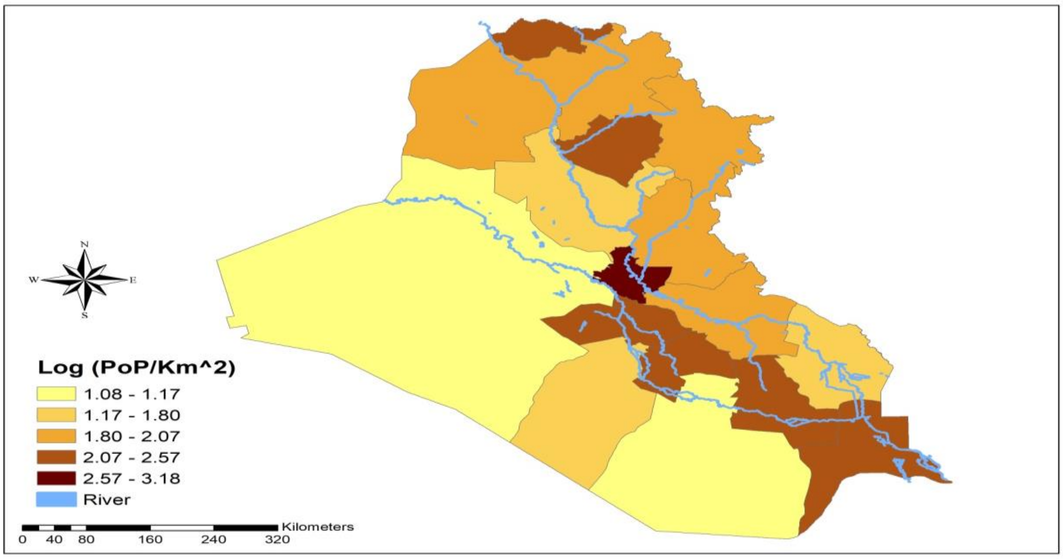


**Fig. 2.** Log population density (people per km$^2$).

**Spatial cluster analysis**

Anselin's Local Moran I statistics were first described in 1995 (Moran, The interpretation of statistical maps, 1948). The purpose of this technique is to identify groups of features with values of similar size and also for identifying outliers by comparing adjacent features and entire population. The ability to explicitly identify spatial extreme values is a Moran's I advantage regarding spatial survey statistics. Spatial analysis often requires knowledge of the degree of spatial correlation between variables. More recently, additional spatial statistic models have been developed, known as local spatial statistics, to measure the correlation between a single geographic region and its neighbors at a specified distance (Miller 2004). Local spatial association index is a local spatial statistic that measures the local spatial autocorrelation, for providing information about spatial clusters and spatial outliers (Anselin 1995; Getis 1996). The local Moran's I developed by Anselin (1995) indicates the similarity of geographical unit in relation to its neighboring units. A positive value of Moran's I is indicated that the feature is surrounding by features of similar values. This feature is a part of the cluster. A negative value of Moran's I indicates that the feature is surrounded by features of different values. Such a feature is an outlier. One of the limitations of this method is that the positive value of Moran's I can also be specified for regions with low values if they are surrounded by other regions with low values (Serra-Sogas *et al.* 2008; Murgante & Borruso 2012). Anselin's Local Moran I statistics have been applied using the cluster and outliers analysis tool in the ArcGIS 10.7.1 Spatial Statistics toolbox. Fixed Euclidean distance ranges including 750 km were analyzed in order to investigate potential distances at which clustering occurred.

## RESULTS AND DISCUSSION

Fig. 3 shows coronavirus cases appeared in southwestern-northwestern regions of Iraq. Analysis of coronavirus disease by Moran's I test between 24/02/2020 and 06/04/2020 indicated the spatial random of coronavirus disease incidence rate for every 100,000 population. The coronavirus disease random distribution during this year confirmed value of Moran's I test -0.055, z-score 0.37 and p-value 0.71. These results confirmed that the adjoining province are also facing the same situation of coronavirus distribution which may be low or high. Anselin's Local Moran's I statistic were used to analyze coronavirus incidence in the Iraq for each individual in 2020. Using this, method has been specifically suggested by the previous studies. Fig. 2 represents population density in the Iraq provinces.

The low and high population densities were found in the southwestern provinces (Najaf and Karbala) of the country, with high human coronavirus incidence rates, because these contain religious centers and airports. Middle population densities were observed in the northeastern ones (Erbil and Sulaymaniyah) with high human coronavirus incidence rates, because these provinces share common borders with neighboring Iran, which is affected by coronavirus. The low densely populated western and middle parts of the country exhibited low rates of human coronavirus incidence (Fig. 3). The Local Moran's I method exhibiting high-low, classified 4 provinces, the northeast and southwest regions, while 10 provinces, displayed the most low-high spatial clustering in the middle part (Fig. 4). In summary, the southwestern and northeastern regions were the most evident areas of human coronavirus clustering throughout the study.

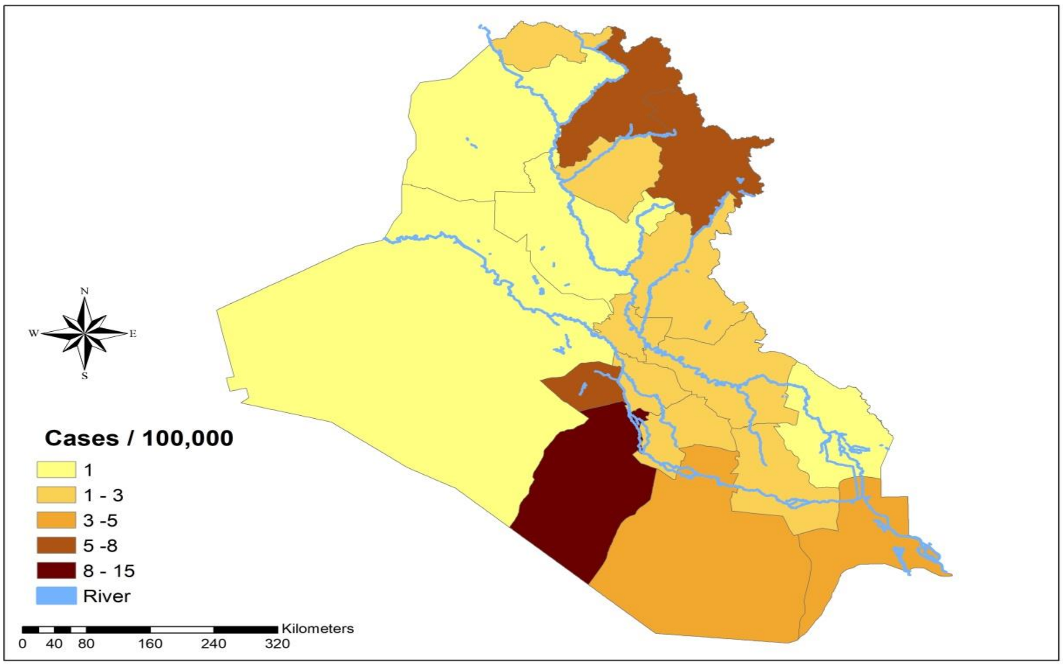


**Fig. 3.** Number of human coronavirus cases per 100,000 people between 24/02/2020 and 06/04/2020.

## CONCLUSION

Spatial patterns statistics were used to apply Moran's I test in order to estimate considerable negative spatial autocorrelation of coronavirus disease incidences during 24/02/2020 to 06/04/2020. Analysis using spatial autocorrelation statistics indicated a negative (random) spatial autocorrelation in the spatial distribution of coronavirus. Anselin's local Moran's I statistics were a particularly useful tool for detecting clusters and outlier values located but also to derive some understanding of the correlation between the spatial distribution of coronavirus. Clusters of coronavirus cases were found by means of spatial statistic methods in some sectors of northeastern and southwestern parts of the country. This study displayed that these methods and tools can be useful for the coronavirus assessment in the public health personnel. It explained that utilizing the current health data, spatial analysis and geographic information systems can provide an opportunity to determine the health burden of the coronavirus in the affected areas, and also lays the basis for further investigation into the associated factors responsible for an increased risk of disease.

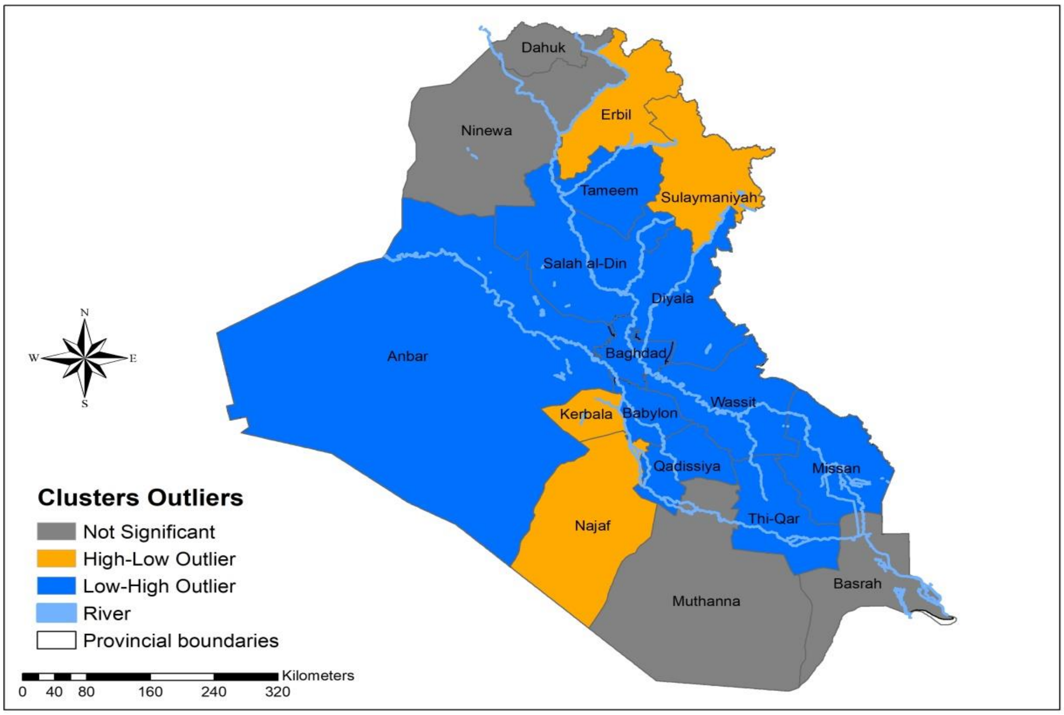


**Fig. 4.** Local Moran's I analyses between 24/02/2020 and 06/04/2020).

## REFERENCES


Albrecht, J 2007, Key concepts and techniques in GIS. Sage Publications, 101 p.

Algert, SJ, Agrawal, A & Lewis, DS 2006, Disparities in access to fresh produce in low-income neighborhoods in Los Angeles. *American Journal of Preventive Medicine*, 30: 365-370.

Al-Kindi, KM, Alkharusi, A, Alshukaili, D, Al Nasiri, N, Al-Awadhi, T, Charabi, Y, & El Kenawy, AM 2020, Spatiotemporal sssessment of COVID-19 spread over Oman using GIS techniques. *Earth Systems and Environment*, 4: 797-811.

Anselin, L 1995, Local indicators of spatial association-LISA. *Geographical Analysis*, 27: 93-115.

Apparicio, P, Cloutier, MS & Shearmur, R 2007, The case of Montreal's missing food deserts: Evaluation of accessibility to food supermarkets. *International Journal of Health Geographics*, 6: 1-13.

Bendaif, H, Hammouti, B, Stiane, I, Bendaif, Y, El Ouadi, MA & El Ouadi, Y 2020, Investigation of spread of novel coronavirus (COVID-19) pandemic in MOROCCO & estimated confinement duration to overcome the danger phase. *Caspian Journal of Environmental Sciences*, 18: 149-156.

Biswas, K, Khaleque, A & Sen, P 2020, Covid-19 spread: Reproduction of data and prediction using a SIR model on Euclidean network. arXiv preprint arXiv:2003.07063.

Boulos, MN & Geraghty, EM 2020, 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. BioMed Central.

Briggs, DJ & Elliott, P 1995, The use of geographical information systems in studies on environment and health. World health statistics quarterly. *Rapport Trimestriel de Statistiques Sanitaires Mondiales*, 48: 85-94.

Chaput, EK, Meek, JI & Heimer, R 2002, Spatial analysis of human granulocytic ehrlichiosis near Lyme, Connecticut. *Emerging Infectious Diseases*, 8: 943.

Danon, L, Brooks-Pollock, E, Bailey, M & Keeling, MJ 2020, A spatial model of CoVID-19 transmission in England and Wales: Early spread and peak timing. MedRxiv.

Getis, A 1996, Local spatial statistics: An overview. Spatial analysis: Modelling in a GIS environment. pp. 261-277.

Kang, L, Li, Y, Hu, S, Chen, M, Yang, C, Yang, BX *et al.* 2020, The mental health of medical workers in Wuhan, China dealing with the 2019 novel coronavirus. The Lancet Psychiatry. Mar;7(3):e14. DOI: 10.1016/S2215-0366(20)30047-X.

Maliki, I, Elmsellem, H, Hafez, B, EL Moussaoui, A, Reda Kachmar, M & Ouahbi, A 2020, The psychological properties of the Arabic BDI-II and the psychological state of the general Moroccan population during the mandatory quarantine due to the COVID-19 pandemic. *Caspian Journal of Environmental Sciences*, pp. 111-122.

Miller, HJ 2004, Tobler's first law and spatial analysis. *Annals of the Association of American Geographers*, 94: 284-289.

Moran, PA 1948, The interpretation of statistical maps. *Journal of the Royal Statistical Society: Series B (Methodological)*, 10: 243-251.

Moran, PA 1950, Notes on continuous stochastic phenomena. *Biometrika*, 37: 17-23.

Murgante, B & Borruso, G 2012, Analyzing migration phenomena with spatial autocorrelation techniques. International Conference on Computational Science and Its Applications, pp. 670-685.

Novel, CP *et al.* 2020, The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China. Zhonghua liu xing bing xue za zhi= Zhonghua liuxingbingxue zazhi, 41, 145.

Organization, WH *et al.* 2005, Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV). Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV).

Pathirana, S, Kawabata, M & Goonatilake, R 2009, Study of potential risk of dengue disease outbreak in Sri Lanka using GIS and statistical modelling. *Journal of Rural and Tropical Public Health*, 8: 8-17.

Sarwar, S, Waheed, R, Sarwar, S & Khan, A 2020, COVID-19 challenges to Pakistan: Is GIS analysis useful to draw solutions? *Science of the Total Environment*, 730: 139089.

Serra-Sogas, N, O'Hara, P, Canessa, R, Bertazzon, S & Gavrilova, M 2008, Exploratory spatial analysis of illegal oil discharges detected off Canada's Pacific Coast. International Conference on Computational Science and Its Applications, pp. 81-95.

Sithiprasasna, R, Patpoparn, S, Attatippaholkun, W, Suvannadabba, S & Srisuphanunt, M 2004, The geographic information system as an epidemiological tool in the surveillance of dengue virus-infected Aedes mosquitos.

Talen, E & Anselin, L 1998, Assessing spatial equity: an evaluation of measures of accessibility to public playgrounds. *Environment and Planning A*, 30: 595-613.

Tsou, KW, Hung, YT & Chang, YL 2005, An accessibility-based integrated measure of relative spatial equity in urban public facilities. *Cities*, 22: 424-435.

Vine, MF, Degnan, D & Hanchette, C 1997, Geographic information systems: Their use in environmental epidemiologic research. *Environmental Health Perspectives*, 105: 598-605.

Wang, D, Hu, B, Hu, C, Zhu, F, Liu, X, Zhang, J *et al.* 2020, Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in Wuhan, China. *Jama*, 323: 1061-1069.

Zhou, P, Yang, XL, Wang, XG, Hu, B, Zhang, L, Zhang, W *et al.* 2020, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 579: 270-273.

***Bibliographic information of this paper for citing:***

Jaber, AS, Hussein, AK, Kadhim, NA, bojassim, AA 2022, A Moran's I autocorrelation and spatial cluster analysis for identifying Coronavirus disease COVID-19 in Iraq using GIS approach. Caspian Journal of Environmental Sciences, 20: 55-60.