Research Article

Spatial statistics and influencing factors of the novel coronavirus pneumonia 2019 epidemic in Hubei Province, China

Yongzhu Xiong [1,2*], Yunpeng Wang [3], Feng Chen [4,5] and Mingyong Zhu [1,2*]

[1] School of Geography and Tourism, Jiaying University, Meizhou 514015, China;
[2] Institute of Resources and Environmental Informatics Systems, Jiaying University, Meizhou 514015, China;
[3] Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou 510640, China;
[4] College of Computer and Information Engineering, Xiamen University of Technology, Xiamen 361024, China;
[5] Big Data Institute of Digital Natural Disaster Monitoring in Fujian, Xiamen University of Technology, Xiamen 361024, China
* Correspondence: yzxiong@gmail.com (Y.X.); Tel.: +86-753-218-6956 (Y.X.); laozhu77@163.com (M.Z.)

Received: date; Accepted: date; Published: date

Abstract: An in-depth understanding of the spatiotemporal characteristics of infectious diseases could be helpful for epidemic prevention and control. Based on the novel coronavirus pneumonia (NCP) data published on official websites, GIS spatial statistics and Pearson correlation methods were used to analyze the spatial autocorrelation and influencing factors of the 2019 NCP epidemic from January 30, 2020 to February 18, 2020. The following results were obtained. (1) During the study period, Hubei Province was the only significant cluster area and hotspot of cumulative confirmed cases of NCP infection at the provincial level in China. (2) The NCP epidemic in China had a very significant global spatial autocorrelation at the prefecture-city level, and Wuhan was the significant hotspot and cluster city for cumulative confirmed NCP cases in the whole country. (3) The cumulative confirmed NCP cases had a very significant global spatial autocorrelation at the county level in Hubei Province, and the county-level districts under the jurisdiction of Wuhan and neighboring Huangzhou district in Huanggang City were the significant hotspots and spatial clusters of cumulative confirmed NCP cases. (4) Based on Pearson correlation analysis, the number of cumulative confirmed NCP cases in Hubei Province had very significant and positive correlations ($p<0.01$) at the prefecture-city and the county levels with four indexes of registered population, resident population, regional GDP and total retail sales of consumer goods, respectively, during the study period. (5) The number of the cumulative confirmed NCP cases in Hubei Province also had a very significant and positive correlation ($p<0.01$) at the prefecture-city level with the Baidu migration index and population density but not with land area, whereas that in Hubei Province had a significant and positive correlation ($p<0.05$) at the county level with land area but not with population density from January 30, 2020, to February 18, 2020. It was found that the NCP epidemic in Hubei Province had distinctive characteristics of a significant centralized outbreak, a significant spatial autocorrelation and complex influencing factors, and that the spatial scale had a significant effect on the global spatial autocorrelation of the NCP epidemic. The findings help to deepen the understanding of spatial distribution patterns and transmission trends of the NCP epidemic and may also benefit scientific prevention and control of epidemics such as COVID-19.

Keywords: novel coronavirus pneumonia (NCP); spatial autocorrelation; influencing factor; spatial statistics; Wuhan city

1. Introduction

In December 2019, acute respiratory diseases (ARD) due to the 2019 novel coronavirus (2019-nCoV) emerged in Wuhan city, rapidly spread throughout China [1]. The sudden outbreak and spread of this novel coronavirus pneumonia (NCP, namely COVID-19) has attracted great attention from the Chinese government and the global public [2-3]. The Party Central Committee and the State Council recognized the magnitude of the outbreak. They promptly implemented comprehensive arrangements for the prevention and control of the epidemic and demanded "having confidence, helping one another, scientific prevention and control, and

precise implementation of policies." Chinese governments at all levels are taking urgent steps to prevent the spread of the epidemic. The World Health Organization (WHO) declared the outbreak of COVID-19 a Public Health Emergency of International Concern (PHEIC) on January 31, 2020, and characterized it as a pandemic due to its worldwide spread on March 11, 2020. During the NCP epidemic outbreak at first, the Chinese Lunar New Year occurred, and a large number of people gathered intensified the transmission and spread of the novel coronavirus, posing a serious threat to the lives and health of the public. As of April 2, 2020, a cumulative total of 82,773 confirmed NCP cases and 3,327 deaths were reported in China, and that of 869,581 confirmed cases and 45,001 deaths were also reported in more than 200 countries or regions around the world outside China, with the rapid growth of the cumulative confirmed NCP cases in North America and Europe. declares COVID-19 a pandemic. At present, the growth of the NCP epidemic in China has been slowing down. However, the global NCP epidemic has been accelerating, and prevention and control are urgent and paramount.

Shortly after the NCP outbreak, scholars conducted extensive research on NCP in terms of etiology, epidemiology, molecular biology, genomics, imaging, and clinical characteristics [4-14]. The results have offered an important scientific basis for the diagnosis and treatment of NCP and epidemic prevention and control. Some universities, research institutions, and online platforms have also launched daily epidemic dynamic information services based on epidemic maps and statistical data, which have provided essential information support for governments and the public to quickly and intuitively understand the progress of the epidemic. However, there is still a lack of research on the spatial statistical characteristics and temporal and spatial evolution of the NCP epidemic, and the spatial evolution mechanism of the NCP epidemic is still unclear. Spatial statistics can help reveal the spatiotemporal characteristics and evolution mechanism of the epidemic, consequently assisting epidemic prevention and control and scientific decisions [15]. Previous studies investigated the transmission network [16], spatial correlation [17-19], space-time transmission dynamics [20], epidemic characteristics, and law of spatial-temporal transmission [21] of the 2003 SARS epidemic, providing a useful reference for the study of spatial statistics of the NCP epidemic. NCP is a SARS-like infectious disease [22] that mainly achieves human-to-human transmission through respiratory droplets, person-to-person contact, and fecal-mouth contact[1]. Will the NCP epidemic be as spatially correlated as the SARS epidemic? Because the NCP epidemic is still spreading, prevention and control are urgent, in particular the spatial statistical characteristics and influencing factors. Hubei Province was seriously affected by NCP, and the lack of temporal and spatial evolution analyses of the NCP epidemic at the city and county levels at the early stage caused certain difficulties in the deployment of epidemic prevention and control materials as well as in precision medical support services in some cities and counties. This study uses GIS spatial statistics and Pearson correlation analysis methods to systematically analyze the spatial autocorrelation characteristics and influencing factors of the NCP epidemic from three scales (i.e. the provincial, prefecture and county levels) in order to provide a helpful decision-making reference for study on the temporal and spatial evolution and for scientific prevention and control of the NCP epidemic.

2. Materials and Methods

2.1. study area

Hubei Province is located in central China (108°21′42 ″ ～ 116°07′50″ E, 29°01′53″ ～ 33°6′47″ N) (Figure 1). The province's land area is 18.59×10^4 km², accounting for 1.94% of the country's total land area. The terrain is mountainous in the east, west, and north and low in the central region, with an incomplete basin opening slightly to the south. Most area within the province has a subtropical monsoon humid climate, with an average annual temperature of 15 °C to 17 °C and average annual precipitation of 800 mm to 1,600 mm. At present, Hubei Province has 12 prefecture-level cities (Wuhan, Huangshi, Shiyan, Yichang, Xiangyang, Ezhou, Jingmen, Xiaogan, Jingzhou, Huanggang, Xianning, Suizhou), one autonomous prefecture (Enshi Tujia and Miao Autonomous Prefecture), three municipalities directly under the administration (Xiantao, Qianjiang, and Tianmen) and one forest area (Shennongjia forest area), collectively referred to as prefecture-level cities in the following analysis. The county-level administrative units in Hubei Province include 39 municipal districts, 22 county-level cities (excluding the three municipalities directly under the provincial government), 36 counties, and two autonomous counties (Figure 1). As of the end of 2018, the province's total resident population was 59 million, with an urbanization rate of 60.3%; the total road distance was 2.75×10^5 km, of which high-grade highways accounted for 3.56×10^4 km, and the total

passenger traffic was 9.87×10^8. The province's total GDP in 2018 was RMB 394 billion yuan with a growth ratio of 7.8%.

Wuhan city, the provincial capital, is located in the eastern part of the Jianghan Plain and the middle reach of the Yangtze River. It is known as "the thoroughfare of nine provinces" due to its national comprehensive transportation hub. The land area of Wuhan is 8.57×10^4 km^2. As of the end of 2019, Wuhan had a registered population of 9.08×10^6 and a floating population of 5.10×10^6. By the end of 2018, there were 6.34×10^3 health institutions in Wuhan, 9.59×10^4 beds, 7.48 hospital beds per thousand people, 10.67×10^4 health care personnel, and 3.42 doctors per thousand people. The reported incidence rate of class A and B infectious diseases was 18/10^4 people per year.

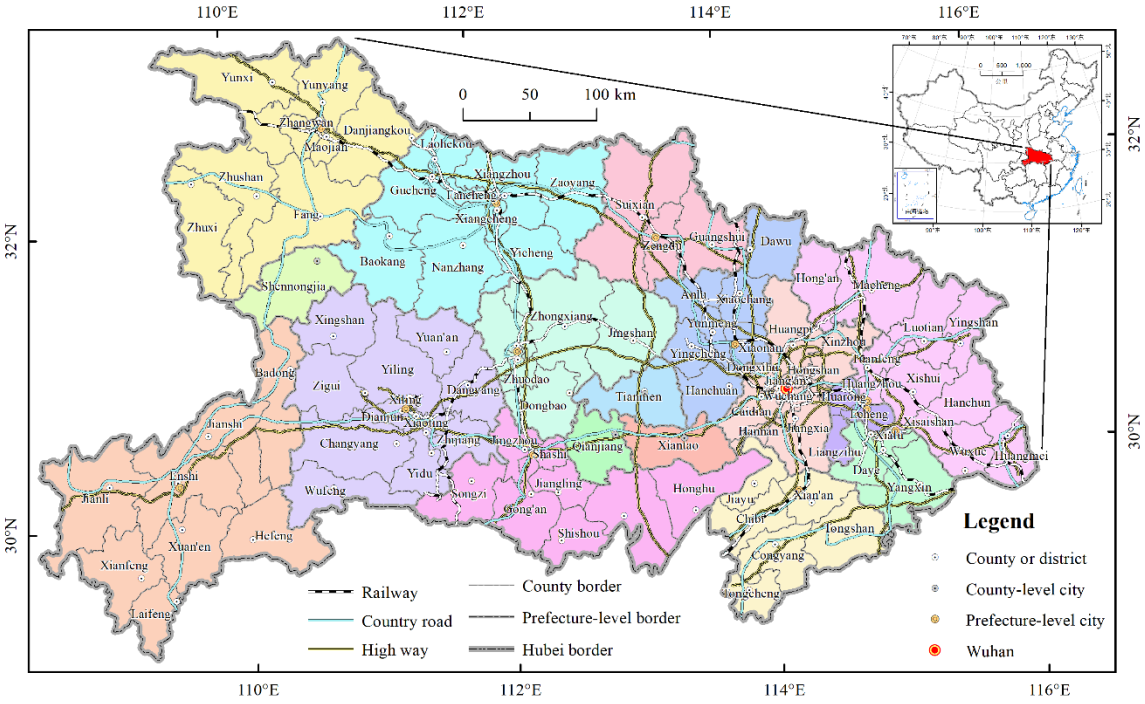

Figure 1. Study area (location, administration and transportation)

2.2. *Data sources*

Provincial and prefecture-level geographic basemap data (2015) were obtained from the Resource and Environmental Science Data Center of the Chinese Academy of Sciences, and the county-level basemap data were obtained from the website of the former National Bureau of Surveying and Mapping, and some administrative areas had been edited and updated according to the latest standard map (2018, downloaded from the Hubei Natural Resources Department). The data for land area, population, regional GDP, and total retail sales of social consumer goods in prefecture-level cities and counties (districts, municipalities) in Hubei Province (2018, part of 2017) were obtained from the "Hubei Statistical Yearbook 2019," and the 2018 annual statistical yearbooks for prefecture-level cities; no relevant data for 2019 was collected. Baidu migration index data (January 17, 2020, to February 1, 2020) was obtained from Smarteye Map (Baidu). The confirmed NCP case data for Hubei Province and its 12 prefecture-level cities, one state, three directly managed cities, and one prefecture-level forest district were obtained from the official website of the Hubei Provincial Health Committee. Except for the NCP epidemic data for the counties (district, city) under the jurisdiction of Wuhan, Ezhou, and Jingzhou, the data for the counties (district, city) under the jurisdiction of the other nine prefecture-level cities in Hubei Province can be found online, and the epidemic data were obtained from the official websites of the municipal governments or the websites of the affiliated health committees. The data for confirmed NCP cases at the provincial level and the data for discharges and deaths related to NCP in prefecture-level cities in Hubei Province were compiled and provided by Meishuclass of a news agency *The Paper*. After the relevant data were sorted and cleaned, ArcGIS was used to establish a national (with a focus on Hubei Province) NCP epidemic geospatial database.

2.3. *Research methods*

This study first used spatial autocorrelation technology to test whether the number of cases in the statistical areas at the provincial, prefecture and county levels had significant global or local spatial autocorrelation characteristics, which was achieved using ArcGIS 10.7 spatial statistical modeling. Second, SPSS 22 was used to carry out Pearson statistical correlation analysis to detect correlations between confirmed NCP cases at the prefecture level and county level and other indicators such as population, economy and environment to explore the influencing factors.

Spatial autocorrelation is an exploratory spatial statistical analysis method that has been widely used to study the spatial distribution of populations [23-24], regional economic patterns [25], epidemic situations [17-19] and urban thermal environments [26] and can be divided into two methods: global spatial autocorrelation and local spatial autocorrelation. In this study, global spatial autocorrelation is used to measure the spatial characteristics of the number of cases in the entire region and to analyze the overall spatial correlation and spatial difference within the region. The global Moran's *I* index is often used to measure the global spatial correlation [25]. The value range of Moran's *I* index is [-1,1]. If Moran's *I* index is statistically positive, there is a positive correlation in the spatial distribution, and there is a spatial clustering effect; otherwise, a negative spatial correlation exists. If Moran's *I* index is zero, there is a random distribution. A statistical *Z*-score and p-value need to be combined to determine statistical significance. Local spatial autocorrelation is commonly characterized by local Moran's *I* and Getis-Ord G_i^*. Local Moran's *I* is the decomposition of the global Moran's *I* into various regional units, also known as local indicators of spatial association (LISA) [27]. In this study, ArcGIS clustering and anomaly analysis (Anselin local Moran's *I*, ALMI) [27] were used to detect the local spatial autocorrelation characteristics of NCP cases and identify the regions with high/low clustering significance. Getis-Ord G_i^* was also used to identify the spatial association between hotspots and cold spots with statistical significance [28] and to detect whether the number of cases in each space indicate high-value aggregation or low-value aggregation.

3. Results

3.1. NCP epidemic spatial statistics

This section analyzes the time series variation characteristics of global spatial autocorrelation and local spatial autocorrelation (clustering and anomaly, Getis-Ord G_i^* hotspots) of the NCP epidemic in Hubei Province from three scales: provincial, prefecture and county.

3.1.1 Spatial autocorrelation analysis of provincial NCP outbreaks nationwide

The results of the global spatial autocorrelation analysis indicated that from January 30, 2020, to February 18, 2020, there were no significant global spatial autocorrelations in the cumulative confirmed NCP cases in various provinces (municipalities and autonomous regions) nationwide. ($p>0.05$, $Z<1.96$).

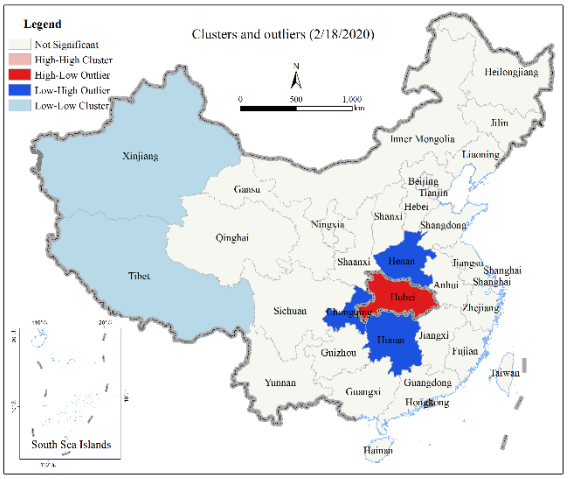

Figure 2. Cluster and outlier analysis results of provincial cumulative confirmed NCP cases nationwide

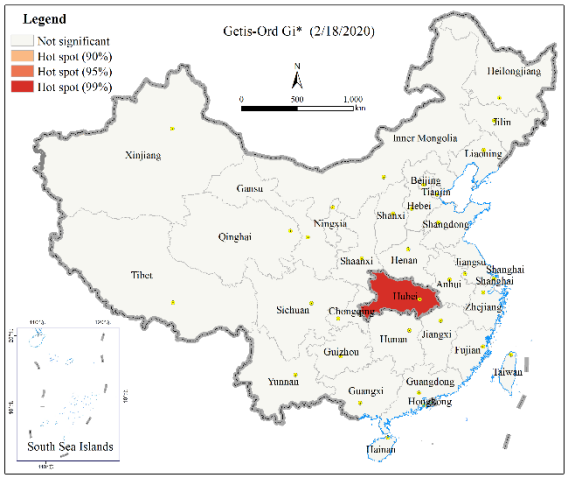

Figure 3. Hotspots (Getis-Ord G_i^*) of provincial cumulative confirmed NCP cases nationwide

Anselin local Moran's *I* (ALMI) analysis results (Figure 2, other dates omitted) revealed that from January 19, 2020, to February 18, 2020, Hubei Province was a significant high-low abnormal area of cumulative confirmed NCP cases nationwide (95% confidence) and that some surrounding provinces were significant low-high anomaly regions. Inverse distance-based hotspot (Getis-Ord G_i^*) analysis results (Figure 3, other dates omitted) also showed that from January 19, 2020, to February 18, 2020, Hubei Province was the only extremely significant hotspot area (99% confidence) for cumulative confirmed NCP cases at the provincial level nationwide. The results of the local spatial autocorrelation analysis showed that Hubei Province had the most substantial NCP epidemic statistics at the provincial level.

3.1.2. Spatial autocorrelation analysis of prefecture-level NCP outbreaks

It can be seen from Figure 4 that from January 19, 2020 to January 22, 2020, there was no significant global spatial autocorrelation in prefecture-level cumulative NCP confirmed cases nationwide ($p>0.05$, $Z<1.96$), which may be related to the limited national reports of confirmed NCP cases from prefecture-level cities during this period. On January 23, 2020, there was a significant global spatial autocorrelation ($p<0.05$, $Z>1.96$). From January 24, 2020, to February 18, 2020, there was an extremely significant global spatial autocorrelation ($p<0.01$, $Z>2.58$); this indicated that the nationwide prefecture-level NCP cases had a very significant aggregation distribution feature. The Moran *I* index in Figure 4 displayed two phases: an initial increase and then decrease around January 30, 2020, indicating that January 30, 2020, was a turning point in the global spatial autocorrelation (Moran *I* index) of cumulative confirmed NCP cases at the prefecture level nationwide.

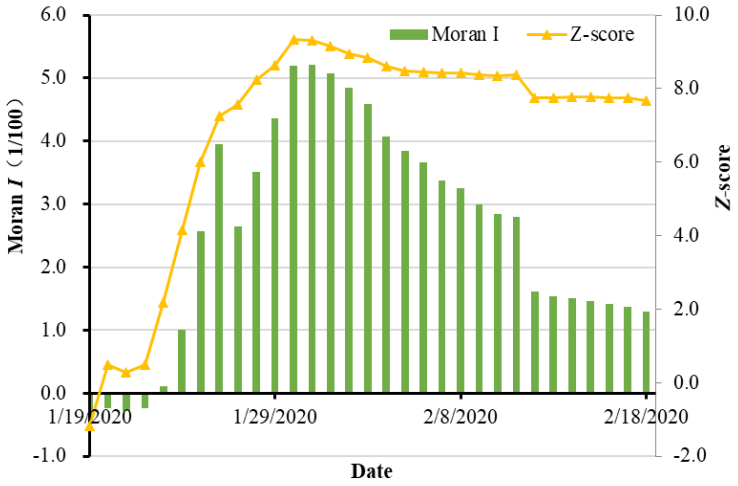

Figure 4. Global spatial autocorrelation analysis results of cumulative confirmed NCP cases at the prefecture level nationwide

The ALMI results revealed that from January 19, 2020, to February 18, 2020, only the prefecture-level cumulative confirmed NCP cases in Hubei Province had significant high-high cluster areas and that some areas around Hubei Province prefecture-level cities had low-high anomaly areas. Figure 5 shows that on February 18, 2020, 13 prefecture-level cities, namely, Wuhan, Xiaogan, Huanggang, Ezhou, Huangshi, Xianning, Suizhou, Xiangyang, Jingmen, Yichang, Jingzhou, Tianmen and Xiantao in Hubei Province, were all significant high-high cluster areas for cumulative confirmed NCP cases. Xinyang in Henan Province, Changsha in Hunan Province, and Nanchang in Jiangxi Province were also significant high-high cluster areas for cumulative confirmed NCP cases. In contrast, some areas closely adjacent to these prefecture-level cities were significant low-high abnormal areas.

Inverse distance-based hotspot (Getis-Ord G_i^*) analysis results (Figure 6, other dates omitted) indicated that in Hubei Province, from January 19 to February 18, 2020, Wuhan was an extremely significant hotspot area for cumulative confirmed NCP cases at the prefecture level (99% confidence, Figure 6a-6d). On January 24, 2020, Huanggang City became a significant hotspot area (95% confidence) due to the increase in cumulative confirmed NCP cases; it became an extremely significant hotspot area (99% confidence) on January 25, and it returned to a significant hotspot area (95% confidence, Figure 6c) on February 8. Due to a sharp increase in the number of confirmed NCP cases in Wuhan after February 12, Huanggang City became a nonsignificant area. On January 28, 2020, the cumulative increase in confirmed NCP cases in Xiaogan City made the city a significant hotspot area (95% confidence, Figure 6b); it became an extremely significant

hotspot area on January 29 (99% confidence), and it further became a significant hotspot area (95% confidence) on February 11. After February 12, Xiaogan became a nonsignificant area due to the sharp increase in the number of confirmed NCP cases in Wuhan. Although other prefecture-level cities in Hubei Province also had a large number of cumulative confirmed NCP cases relative to other domestic cities, they were not significant hotspots due to the excessively high cumulative number of confirmed NCP cases in Wuhan. This suggests that Wuhan was prefecture-level city with the most intense NCP epidemic in Hubei Province and the country and that Huanggang City and Xiaogan City were also prefecture-level cities with the most intense NCP epidemic in Hubei Province and in the country for a period of time.

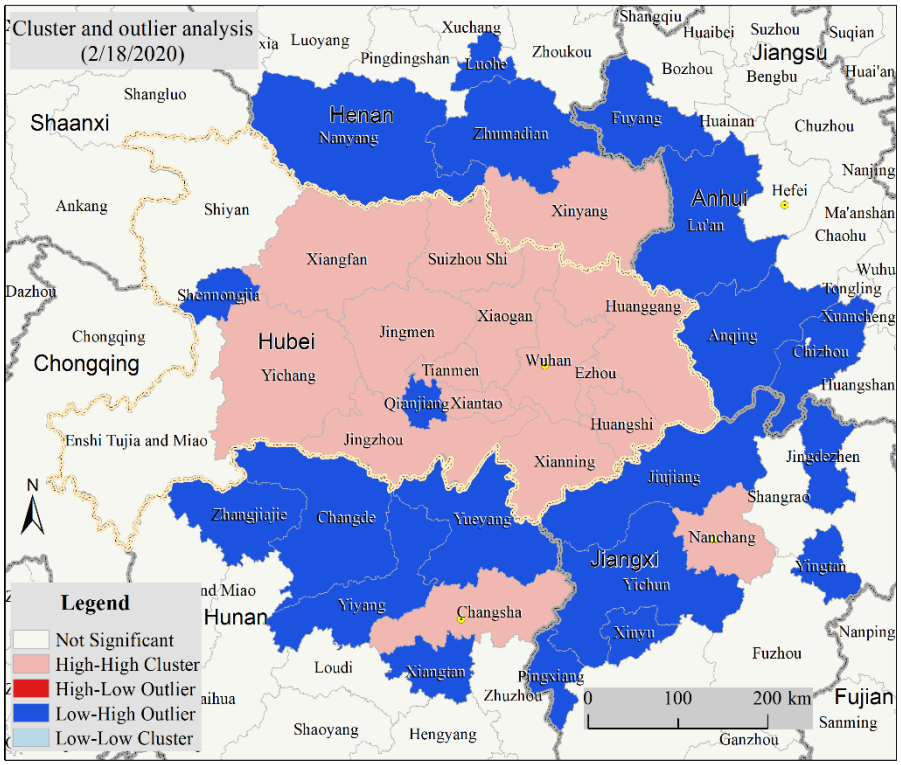

Figure 5. Cluster and outlier analysis results of cumulative confirmed NCP cases at the prefecture level in Hubei Province and surrounding areas

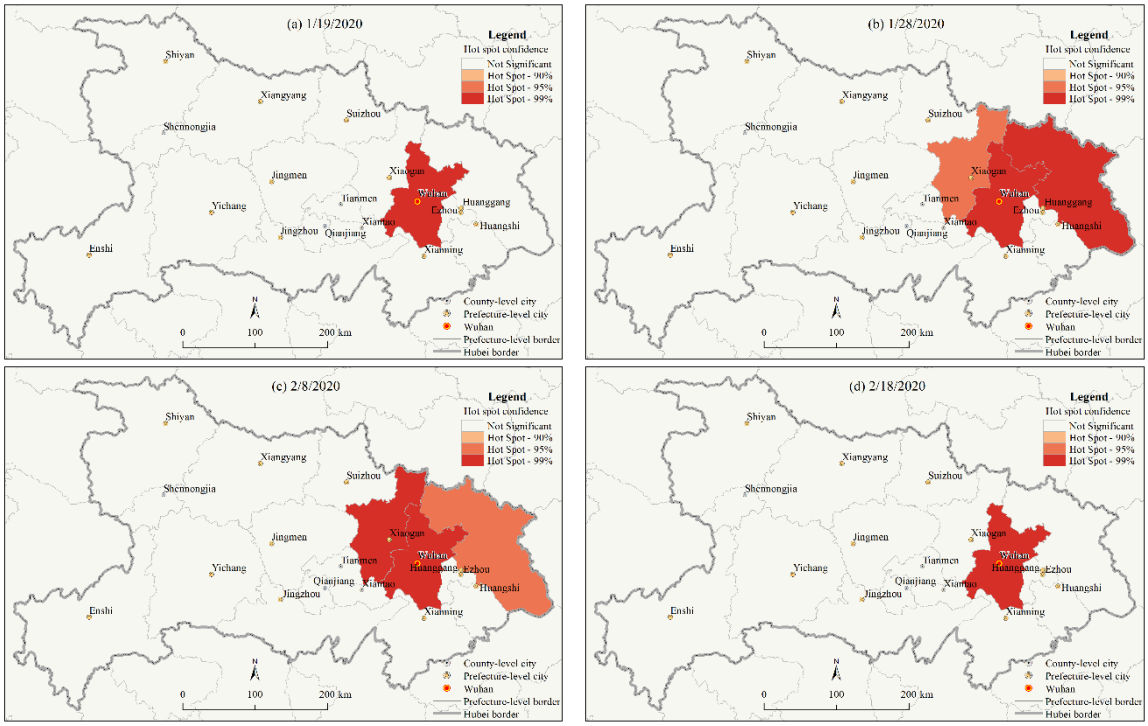

Figure 6. Hotspot (Getis-Ord G_i^*) analysis results for cumulative confirmed NCP cases in Hubei Province

3.1.3. Spatial autocorrelation analysis of county-level NCP outbreaks in Hubei Province

Because county-level NCP epidemic data were temporarily unavailable, the county-level NCP epidemic data for the most serious outbreaks in Hubei Province were collected. However, the county-level NCP epidemic data for Wuhan, Ezhou and Jingzhou of Hubei Province were not available; therefore, only their prefecture-level NCP epidemic data were collected. To eliminate excessive statistical data for Wuhan, Jingzhou and Ezhou, the cumulative confirmed NCP cases were weighted by the resident county populations (cities, districts) to perform spatial autocorrelation analysis. We selected county-level NCP data for Hubei Province from January 30, 2020, to February 18, 2020, for analysis.

The results of the global spatial autocorrelation analysis (Figure 7) illustrated that from January 30, 2020, to February 18, 2020, the cumulative confirmed NCP cases in Hubei Province at the county level had an extremely significant spatial autocorrelation ($p<0.01$, $Z>2.58$) and that the global Moran's *I* index and Z-score increased since January 31, indicating that the county-level NCP epidemic in Hubei Province had very significant aggregation and distribution characteristics and that the spatial autocorrelation was increasingly intense. By comparing Figure 7 and Figure 4, it can be found that the overall spatial autocorrelation of the county-level NCP epidemic in Hubei Province was significantly higher than that of the prefecture-level NCP epidemic and that the trends from January 30, 2020, to February 18, 2020, were also different. This indicated that the spatial scale had a significant effect on the global spatial autocorrelation of the NCP epidemic.

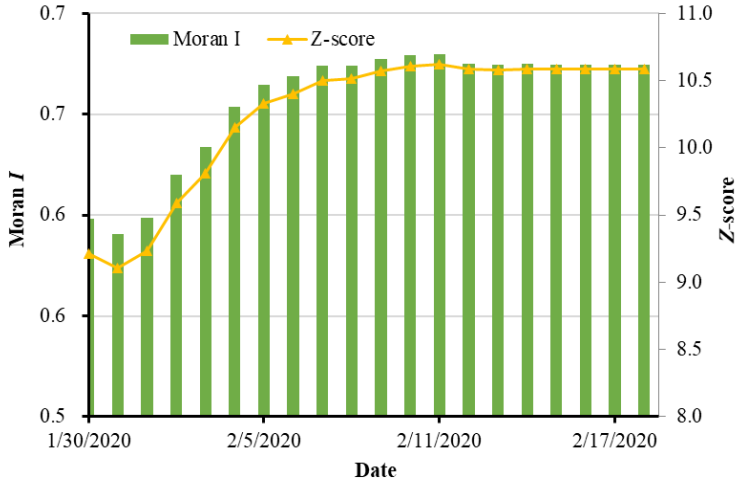

Figure 7. Global spatial autocorrelation analysis results for cumulative confirmed NCP cases at the county level in Hubei Province

The ALMI results (Figure 8, other times omitted) demonstrated that from January 30, 2020, to February 18, 2020, there were significant clusters and abnormal areas regarding the cumulative confirmed NCP cases at the county level in Hubei ($p<0.05$); Wuhan's 12 districts and neighboring Xiaonan District and Hanchuan City (Xiaogan City) were significant high-high cluster regions (Figure 8). The Wuhan Hannan District and its neighborhood Hong'an County (Huanggang City), Huarong District, and Liangzihu District (Ezhou City) were significant low-high anomaly regions, and most county-level administrative regions in southwest Hubei were significant low-low cluster regions. This indicated that the county-level NCP epidemic in Wuhan was serious during this period. Notably, Xiling District in Yichang City showed a significant high-low anomaly, from January 30, 2020, to February 9, 2020 (Figure 8), indicating that the cumulative number of confirmed NCP cases in this county was significantly higher than its surroundings. In neighboring counties and districts, the NCP epidemic was relatively serious compared to the surrounding areas; from February 10, 2020, to February 18, 2020, there were no abnormal features, which indicated that the NCP epidemic in the region had been relatively alleviated during this period.

The hotspot analysis (Getis-Ord G_i^*) results based on inverse distance (Figure 9) showed that from January 30, 2020, to February 18, 2020, the extremely significant hotspot areas (99% confidence) and significant hotspot areas (95% confidence) for the cumulative confirmed county-level NCP cases in Hubei Province were the core urban areas in Wuhan, Huanghua District and Jiangxia District, as well as the neighboring Huangzhou District in Huangzhou. From the perspective of time series development, there were also some changes in the hotspot areas of cumulative confirmed NCP cases at the county level in Wuhan, which will not be described here. In short, the results of the local spatial autocorrelation analysis revealed that

the districts under the jurisdiction of Wuhan were the regions with the most serious county-level NPC cases in Hubei Province. Although the number of confirmed NCP cases in other counties in Hubei Province was relatively high, they was not considered hotspots because the numbers of confirmed NCP cases among those districts in Wuhan were too high.

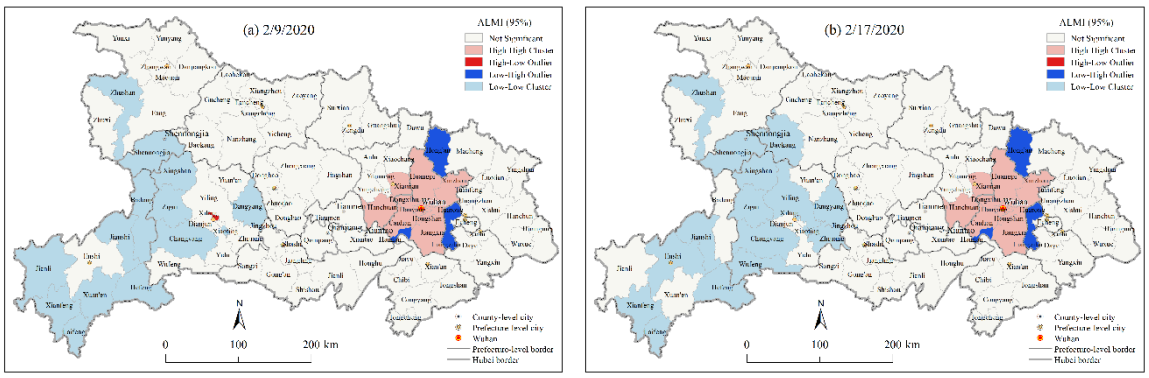

Figure 8. ALMI results of cumulative confirmed NCP cases at the county level in Hubei Province

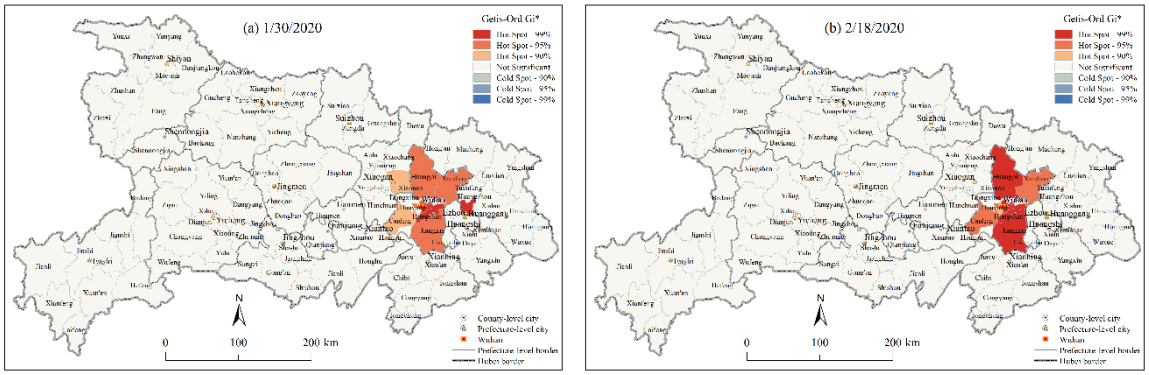

Figure 9. Hotspot analysis (Getis-Ord G_i^*) of cumulative confirmed NCP cases at the county level in Hubei Province.

3.2 The influencing factors of the NCP epidemic

The NCP epidemic in Hubei Province was serious, with concentrated outbreaks and a closed spatial correlation. To explore the causes of the spatiotemporal changes in the NCP epidemic and the influencing factors, we analyzed the following three aspects at the prefecture level and county level for the NCP epidemic in Hubei Province: nature (land area), society (registered population, permanent population, population density, and Baidu migration index) and economy (regional GDP and total retail sales of consumer goods).

3.2.1. Influencing factors of prefecture-level NCP outbreaks in Hubei Province

NCP (COVID-19) is a new infectious disease, and its transmission mechanism is still unclear [29]. Analysis from the aspects of nature, society, and the environment may help provide an understanding of the influencing factors of the NCP epidemic and the transmission mechanism. The Pearson correlation analysis results (Table 1, data omitted for other dates) demonstrated that from January 30, 2020, to February 18, 2020 (20 days), the cumulative confirmed NCP cases at the prefecture level in Hubei Province had an extremely significant positive correlation with the five indicators (population density, permanent population, Baidu migration index, regional GDP, and total retail sales of consumer goods ($p<0.01$)) and an extremely significant positive correlation with the registered population from January 30, 2020, to February 13, 2020 ($p<0.01$). From February 14, 2020, to February 18, 2020, the correlation was significant and positive ($p<0.05$), while the correlation with land area was not significant. To eliminate the possible instability of the data, the mean cumulative confirmed NCP cases at the prefecture level in Hubei Province from January 30, 2020, to February 18, 2020 (20 days), was used for correlation analyses with the above seven indicators. Table 1 shows that this average value had an extremely significant and strong positive correlation with population density, Baidu migration index, regional GDP, and total retail sales of consumer goods ($p<0.01$) an extremely significant

positive correlation with the resident population and the registered population ($p<0.01$). In contrast, the correlation with land area was not significant.

Judging from the correlation coefficients for the cumulative confirmed NCP cases at the prefecture level in Hubei Province (Table 1), the population density, Baidu migration index, regional GDP, and total retail sales of consumer goods were all greater than 0.8, an extreme correlation level, indicating that population flow and social and economic development had significant impacts on the spread of the NCP epidemic; the registered population and resident population trends were consistent (picture omitted); that is, from January 30, 2020, to February 2, 2020, there was an upward trend, and then it declined until February 18. Before February 9, 2020, the correlation coefficient for the resident population was greater than 0.8, an extreme correlation level, and then it ranged from 0.6 to 0.8, a strong correlation level. The correlation coefficient for the registered population was between 0.6 and 0.8 within the 20 days, indicating a strong correlation. This showed that the number of people was also an influencing factor that cannot be ignored in the spread of the NCP epidemic. The more mobile people are and the greater the population density is after an outbreak, the greater the risk of infection with novel coronavirus.

Table 1. Correlation results for cumulative confirmed NCP cases at the prefecture-city level in Hubei Province

Factors	cumulative confirmed NCP cases						
	1/30/2020	2/2/2020	2/6/2020	2/10/2020	2/14/2020	2/18/2020	Mean#
Land area	-0.007	0.004	-0.032	-0.049	-0.065	-0.070	-0.056
Population density	0.813**	0.807**	0.819**	0.823**	0.818**	0.819**	0.820**
Registered population	0.696**	0.709**	0.671**	0.643**	0.605*	0.598*	0.623**
Resident population	0.837**	0.846**	0.819**	0.799**	0.770**	0.765**	0.785**
Baidu migration index	0.916**	0.900**	0.912**	0.918**	0.942**	0.940**	0.931**
Regional GDP	0.949**	0.951**	0.950**	0.952**	0.951**	0.950**	0.952**
The total retail sales of social consumer goods	0.970**	0.972**	0.970**	0.970**	0.966**	0.965**	0.969**

** indicates that the correlation is significant when the confidence (double test) is 0.01; * indicates that the correlation is significant when the confidence (double test) is 0.05; # Refers to the mean cumulative confirmed NCP cases for 20 days from January 30, 2020, to February 18, 2020.

Figure 10 shows that the Baidu migration index for Wuhan to Xiaogan City and Huanggang City in Hubei Province was significantly higher than that for other cities, which was consistent with the current statistical results for the intense NCP epidemic in Xiaogan City and Huanggang City compared to that experienced in other cities in Hubei Province. Wuhan, as the capital city of Hubei Province and the core city of China's central city cluster, has a strong influence on surrounding cities [30], and its connection through traffic [31] and economic ties [32] with Xiaogan, Huanggang, Jingzhou and Xianning are also higher than with other cities. Consequently, there was numerous personnel and commercial contact between Wuhan and these cities, increasing the chance of transmission and the spread of novel coronavirus.

The Pearson correlation analysis results (table abbreviated) showed that from January 30, 2020, to February 18, 2020 (20 d), the correlations between the number of newly confirmed cases in prefecture-level cities in Hubei Province and population density, permanent population, regional GDP, and total retail sales of consumer goods were positive and extremely significant ($p<0.01$), that from January 30, 2020 to February 3, 2020, the correlation between the number of newly confirmed cases in prefecture-level cities in Hubei Province and the registered population was positive and extremely significant ($p<0.01$); from February 3, 2020, to February 18, 2020, the correlation was positive and extremely strong ($p<0.05$). From January 30, 2020, to February 6, 2020, the correlations between the number of newly confirmed cases in prefecture-level cities in Hubei Province and the registered population was positive and extremely significant ($p<0.01$), becoming a positive moderately significant correlation ($p<0.05$) from February 7, 2020, to February 18, 2020. It had an extremely significant strong positive correlation with the Baidu migration index from January 30, 2020 to February 11, 2020 ($p<0.01$), and the related relationship was unstable from February 12, 2020, to February 18, 2020, which may be due to newly confirmed cases including excessive clinically confirmed cases; the correlation with land area was not significant.

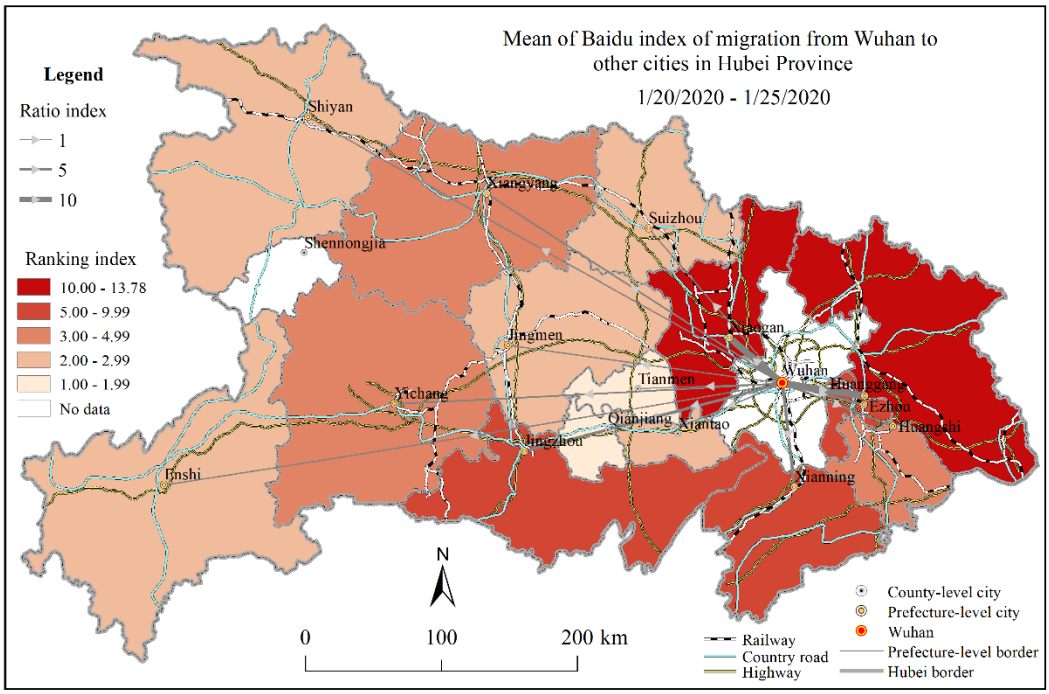

Figure 10. Thematic map of the Baidu migration index from Wuhan to other cities in Hubei Province

From January 30, 2020 to February 18, 2020(20 d), the correlations between the average number of newly confirmed NCP cases at the prefecture level in Hubei Province and population density, Baidu migration index, regional GDP, and total retail sales of social consumer goods were positive and extremely significant ($p<0.01$). The correlation between the average number of newly confirmed NCP cases at the prefecture level in Hubei Province and the resident population was positive and extremely significant, and the correlation between the average number of newly confirmed NCP cases at the prefecture level in Hubei Province and the registered population was positive and moderately significant ($p<0.05$); the correlation with land area was not significant.

3.2.2. Influencing factors of county-level NCP outbreaks in Hubei Province

The Pearson correlation analysis results (Table 2, data omitted so as to include other data) indicated that from January 30, 2020 to February 18, 2020, the cumulative confirmed NCP cases at the county level in Hubei Province was extremely significantly and positively correlated with the registered population, resident population, regional GDP, and total retail sales of consumer goods ($p<0.01$) and significantly and negatively correlated with land area ($p<0.05$) but had no significant correlation with population density. From the cumulative confirmed NCP cases at the county level (Table 2) and the correlation coefficients for various indicators, the cumulative confirmed NCP cases at the county level in Hubei Province was weakly correlated with land area and regional GDP and weakly correlated with the registered population from January 30, 2020, to February 3, 2020; however, it was moderately correlated from February 4, 2020, to February 18, 2020, and moderately associated with the resident population and the total retail sales of social consumer goods. As with the cumulative confirmed NCP cases at the prefecture level, the trend for the correlation between the cumulative confirmed NCP cases at the county level and the registered population was basically the same as that for the permanent population.

Table 2 shows that the mean cumulative confirmed NCP cases at the county level in Hubei Province, from January 30, 2020, to February 18, 2020 (20 d), had an extremely significant positive correlation with registered population, resident population, regional GDP, and total retail sales of consumer goods ($p<0.01$), had a significant negative correlation with land area ($p<0.05$), had no significant correlation with population density, had a weak correlation with regional GDP, and had moderate significant correlations with registered population, resident population and social consumer goods.

Correlation analysis of the newly confirmed NCP cases at the county level with the above five indicators (Table omitted) did not establish a stable correlation. From January 30, 2020, to February 18, 2020 (20 d), the average daily new confirmed NCP cases at the county level and registration population, resident population, regional GDP, and total retail sales of consumer goods had a positive and extremely moderate correlation

(p<0.01). Additionally, the average daily new confirmed NCP cases at the county level had a significant weak negative correlation with land area (p<0.05) and no significant correlation with population density.

Comparing Table 2 and Table 1, it can be found that the correlation coefficients between the cumulative confirmed NCP cases at the prefecture level and county level in Hubei Province and the resident population were larger than that for the registered population, indicating that the resident population may have had a more critical impact on the spread of NCP. The correlation coefficient between the cumulative confirmed NCP cases at the prefecture level and county level in Hubei Province and the total retail sales of social consumer goods was higher than that for the regional GDP, which indicated that the flow of social consumer goods had a more critical impact on the spread of the NCP epidemic than did the regional GDP. The correlation coefficients for the cumulative confirmed NCP cases at the prefecture level in Hubei Province and various indicators were significantly higher than those for the corresponding indicators at the county level, and the correlation with population density and land area also changed. This indicates that the influencing factors of NCP outbreaks at different scales were not exactly the same.

Table 2. Correlation analysis results for cumulative confirmed NCP cases at the county level in Hubei Province

Indicator	Cumulative confirmed NCP cases						
	1/30/2020	2/2/2020	2/6/2020	2/10/2020	2/14/2020	2/18/2020	Mean#
Area	-0.241*	-0.274*	-0.259*	-0.253*	-0.246*	-0.245*	-0.256*
Population Density	0.121	0.167	0.156	0.166	0.159	0.147	0.160
Registered Population	0.329**	0.373**	0.445**	0.470**	0.486**	0.488**	0.467**
Resident Population	0.401**	0.448**	0.521**	0.541**	0.549**	0.552**	0.537**
Regional GDP	0.317**	0.312**	0.344**	0.357**	0.366**	0.377**	0.365**
The total retail sales of social consumer goods	0.481**	0.525**	0.541**	0.569**	0.586**	0.592**	0.578**

** indicates that the correlation is significant when the confidence (double test) is 0.01; * indicates that the correlation is significant when the confidence (double test) is 0.05. # Refers to the average of cumulative confirmed NCP cases for 20 days from January 30, 2020, to February 18, 2020.

4. Discussions

Due to the urgent time and limited energy, our study only investigated the geostatistical characteristics and influencing factors of the NCP epidemic in Hubei Province, China at the early increasing stage of its spreading situation from January 30, 2020, to February 18, 2020. Although a further understanding of the NCP epidemic was reached spatiotemporally, this study has a few of limitations at present.

First, the mechanism controlling the geostatistical characteristics and the NCP spread on different scales has not be resolved. It is an important and complex issue. Our results showed that cumulative confirmed NCP cases at the provincial, prefecture and county levels had significant local spatial autocorrelation and clustering characteristics, but there was no global spatial autocorrelation at the provincial level. The global spatial autocorrelation of cumulative confirmed NCP cases at the county level in Hubei Province was significantly higher than that at the prefecture level, and the trends for the rapid spread of the NCP epidemic from January 30, 2020, to February 18, 2020, were also different. As the worldwide spread of the COVID-19 is growing rapidly, new, big and detail data could be collected and used to analyze geostatistical characteristics of the COVID-19 epidemic and to explain its controlling mechanism further from global to local scales.

Second, the correlation analysis of the influencing factors of the COVID-19 in Hubei Province is only based on the relationship between different index data, but not their cause and effect. The spread of the NCP epidemic is not only related to the natural environment, population, and social and economic activities but also closely related to the prevention and control of the epidemic. Information regarding the NCP epidemic has not been publicly disclosed for long, and the correct early warning and pre-control measures were not adopted in a timely manner. The best opportunity for "early detection, early report, early isolation, and early treatment" was missed for some reason, resulting in a rapid and concentrated outbreak. The migration of a large number of people before the Chinese traditional Spring Festival also accelerated the spread of the epidemic. The study explored the impact of the NCP epidemic and the correlation between cumulative

confirmed NCP cases and seven indicators closely related to the natural environment, population, and social and economic activities, such as land area, registered population, resident population, population density, Baidu migration index, regional GDP, and total retail sales of consumer goods. Some knowledge has been obtained, but it is not comprehensive or sufficiently in-depth. The NCP epidemic correlation with more indexes of population, economy and environment deserves further consideration.

Third, limited by the availability of data, the analysis of the spatiotemporal characteristics of the NCP epidemic at the county scale in this study is not comprehensive. Openness, transparency, and sharing of epidemic data are essential for epidemic prevention and control and scientific research. In future studies, we will continue to update relevant data and analyze the impact mechanisms of the NCP epidemic and social, economic, and environmental indicators at provincial, prefecture and county levels (even townships and streets) in a more comprehensive and in-depth manner. We will carry out NCP epidemic spatial-temporal simulations based on GTWR [33-34] and geographical detectors [35] to further study the spatial distribution pattern and the spread trend for NCP and provide more accurate and reliable decision-making references for scientific prevention and control of the COVID-19 epidemic.

5. Conclusion

Based on publicly available epidemic data, this study used GIS spatial analysis methods to systematically analyze the spatial autocorrelation of NCP outbreaks from the provincial, prefectural and county levels and explored the influencing factors using Pearson correlation analysis. The results showed that cumulative confirmed NCP cases had significant local clustering characteristics, namely, regional outbreak characteristics, and had significant correlation with some social, economic and environmental indexes on different scales. This finding indicates that during the outbreak and spread of the NCP epidemic, it was necessary to focus on treating infectious cases in the most seriously affected areas, carrying out classified social management and providing medical support to achieve scientific prevention and control and precise policy implementation. Our finding also indicates that the epidemic prevention and control measures which restrict population movements and control economic exchange were correct and necessary under the spreading situation of the epidemic and had great significance in preventing further spread of the epidemic.

Declarations

Ethics approval and consent to participate

Not applicable. This study did not require ethics committee approval because no individual data was used. Only open and official data published online was used.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request. The following will be available online at https://github.com/IRESSS/COVID-19, Dataset S1: COVID-19Hubei.xlsx, and Dataset S2: Updates0219_NC.csv.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Conceptualization, Y.X. and Y.W.; methodology, Y.X. and F.C.; investigation, Y.X. and M.Z.; data curation, Y.X. and M.Z.; writing—original draft preparation, Y.X. and M.Z.; writing—review and editing, Y.X., Y.W., F,C. and M.Z.; visualization, Y.X.; funding acquisition, Y.X. All authors read and approved the final manuscript.

Funding

This research was funded by the Guangdong Natural Science Foundation, grant number 2017A030307040 and the Intangible Cultural Heritage Research Base Project of Guangdong Province, grant number 17KYKT13.

Acknowledgments

The authors would like to thank the American Journal Experts for translating and editing our manuscript for free of charge. The authors are grateful to the Meishuclass of *The Paper* for providing the NCP data and the official website of relevant government departments for providing the NCP data, statistical data and geographic basemap data. The authors would like to thank the editors and reviewers for their valuable comments.

References

1. Guan W, Ni Z, Hu Y, et al. Clinical characteristics of 2019 novel coronavirus infection in China. *MedRxiv*, 2020, (in review): 2020.02.06.20020974.
2. Wang C, Horby PW, Hayden FG, et al. A novel coronavirus outbreak of global health concern. *The Lancet,* 2020, in publication.
3. Biscayart C, Angeleri P, Lloveras S, et al. The next big threat to global health? 2019 novel coronavirus (2019-nCoV): What advice can we give to travellers? – Interim recommendations January 2020, from the Latin- American society for Travel Medicine (SLAMVI). *Travel Med Infect Di*, 2020: 101567.
4. Chan JF, Yuan S, Kok K, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *The Lancet*, 2020, in publication.
5. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*, 2020, in publication.
6. Chen J. Pathogenicity and transmissibility of 2019-nCoV—A quick overview and comparison with other emerging viruses. *Microbes Infect*, 2020, in publication.
7. Wu JT, Leung K, Leung G M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet*, 2020, in publication.
8. Tang JW, Tambyah PA, Hui DS C. Emergence of a novel coronavirus causing respiratory illness from Wuhan, China. *J Infection*, 2020, in publication.
9. Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. *New Engl J Med*, 2020, in publication.
10. Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *The Lancet*, 2020, in publication.
11. [11] Zhou P, Yang X, Wang X, et al. A pneumonia outbreak associated with a novel coronavirus of probable bat origin. *Nature*, 2020: in publication.
12. Lu R, Zhao X, Li J, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet,* 2020, in publication.
13. Xu X, Chen P, Wang J, et al. evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. *Science China Life Sciences*, 2020, in publication.
14. Liao X, Wang B, Kang Y. Novel coronavirus infection during the 2019–2020 epidemic: preparing intensive care units—the experience in Sichuan Province, *China. Intens Care Med*, 2020, in publication.
15. Cao Z, Zeng D, Zheng X, et al. Spatio-temporal evolution of Beijing 2003 SARS epidemic. *Science China Earth Science*, 2010, 40 (6): 776-788.
16. Hu B, Gong J, Sun J, et al. Exploring the epidemic transmission network of SARS in-out flow in mainland China. *China Science Bulletin*, 2013, 58 (5): 452-464.
17. Wu J, Wang J, Meng B, et al. Spatial association analysis on epidemic of SARS in Beijing, 2003. *Journal of Zhejiang University* (Ag ric. & Life Sci.), 2005, 31 (1): 97- 101.
18. Fan X, Ying L. An exploratory spatial data analysis of SARS epidemic in China. *Advances in Earth Science*, 2005, 20 (3): 282-291.
19. Cao Z, Wang J, Gao Y, et al. Risk factors and autocorrelation characteristics on severe acute respiratory syndrome in Guangzhou. *Acta Geographica Sinica*, 2008, 63 (9): 981-993.
20. Cao C, Chen W, Zheng S, et al. Analysis of Spatiotemporal Characteristics of Pandemic SARS Spread in Mainland China. *Biomed Res Int*, 2016, 2016: 1-12.
21. Hu B, Gong J, Zhou J, et al. Spatial-temporal characteristics of epidemic spread in-out flow--Using SARS epidemic in Beijing as a case study. Science *China Earth Science*, 2013, 43: 1499-1517 .
22. Wang R, Zhang X, Irwin DM, et al. Emergence of SARS-like Coronavirus poses new challenge in China. *J Infection*, 2020, in publication.
23. Zhao M, Wang D. The spatial sprawl and driving mechanism of the floating population in Beijing metropolitan areas. *Scientia Geographica Sinica*, 2019, 39 (11): 1729-1738.
24. Wang R, Huang X, Xue D, et al. Spatio-temporal change of facaulty members of higher education institute and its influential factors in China in 2005-2015. *Scientia Geographica Sinica*, 2019, 39 (8): 1199- 1207.
25. Zhou Y, Li N, Wu W, et al. evolution of spatial-temporal pattern of county economic development in China during 1982-2010. *Progress in Geography*, 2014, 33 (1): 102-113.
26. Ge R, Wang J, Zhang L, et al. Impacts of urbanization on the urban thermal environment in Beijing. *Acta Ecologica Sinica*, 2016, 36 (19): 6040-6049.
27. Anselin L. Local Indicators of Spatial Association—LISA. *Geogr Anal*, 1995, 27 (2): 93-115.
28. Ord K, Getis A. Local spatial autocorrelation statistics: Distributional issues and an application. *Geogr Anal*, 2010, 27: 286-306.

29. Liu T, Hu J, Xiao J, et al. Time-varying transmission dynamics of Novel Coronavirus Pneumonia in China. *BioRxiv*, 2020, (in review): 2020.01.25.919787.
30. Yu YY, Tong Y, Hu S, et al. Measurement of Intercity Interaction among Wuhan urban clusters. *Urban Issues*, 2017, (1): 44-52.
31. Huang J, Li J, Zhou H, et al. Dynamic development of Wuhan metropolitan area based on urban connection. *Planner*, 2017, 33 (1): 85-92.
32. Wei S, Chen M. Study on Urban Economic Relations and Regional Economic Development in Hubei Province. *Statistics & Decision*, 2018, (2): 127-130.
33. Huang B, Wu B, Barry M. Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *Int J Geogr Inf Sci*, 2010, 24 (3): 383-401.
34. Wang H, Zhang B, Liu Y, et al. Multi-dimensional analysis of urban expansion patterns and their driving forces based on the center of gravity-GTWR model: A case study of the Beijing-Tianjin-Hebei urban agglomeration. *Acta Geographica Sinica*, 2018, 73 (6): 1076-1092.
35. Wang J, Xu C. Geodetector: Principle and prospective. *Acta Geographica Sinica*, 2017, 72 (1): 116-134.