



# “Distance-Driven” Versus “Density-Driven”: Understanding the Role of “Source-Case” Distance and Gathering Places in the Localized Spatial Clustering of COVID-19—A Case Study of the Xinfadi Market, Beijing (China)

DOI:

[10.1029/2021GH000458](https://doi.org/10.1029/2021GH000458)

## Document Version

Final published version

[Link to publication record in Manchester Research Explorer](#)

## Citation for published version (APA):

Zhang, S., Yang, Z., Wang, M., & Zhang, B. (2021). “Distance-Driven” Versus “Density-Driven”: Understanding the Role of “Source-Case” Distance and Gathering Places in the Localized Spatial Clustering of COVID-19—A Case Study of the Xinfadi Market, Beijing (China). *GeoHealth*, 5(8), Article e2021GH000458. <https://doi.org/10.1029/2021GH000458>

## Published in:

GeoHealth

## Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

## General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

## Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact [openresearch@manchester.ac.uk](mailto:openresearch@manchester.ac.uk) providing relevant details, so we can investigate your claim.



**Special Section:**

The COVID-19 pandemic:  
linking health, society and  
environment

**Key Points:**

- The spatiotemporal characteristics of localized spatial clustering of COVID-19 were quantitatively analyzed by multimethod
- The improved quantitative research methods were applied in the modeling of the spatial clustering factors at the micro scale
- The trend from "distance-driven" to "density-driven" of the localized spatial clustering of COVID-19 was predicted

**Correspondence to:**

B. Zhang,  
[blzhangsd@163.com](mailto:blzhangsd@163.com)

**Citation:**

Zhang, S., Yang, Z., Wang, M., & Zhang, B. (2021). "Distance-driven" versus "density-driven": Understanding the role of "source-case" distance and gathering places in the localized spatial clustering of COVID-19—A case study of the Xinfadi market, Beijing (China). *GeoHealth*, 5, e2021GH000458. <https://doi.org/10.1029/2021GH000458>

Received 16 MAY 2021

Accepted 24 JUL 2021

**Author Contributions:**

**Conceptualization:** Sui Zhang  
**Data curation:** Sui Zhang, Zhao Yang  
**Formal analysis:** Sui Zhang  
**Funding acquisition:** Baolei Zhang  
**Methodology:** Sui Zhang  
**Project Administration:** Baolei Zhang  
**Software:** Sui Zhang  
**Supervision:** Baolei Zhang

© 2021. The Authors. *GeoHealth* published by Wiley Periodicals LLC on behalf of American Geophysical Union. This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs License](#), which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

# "Distance-Driven" Versus "Density-Driven": Understanding the Role of "Source-Case" Distance and Gathering Places in the Localized Spatial Clustering of COVID-19—A Case Study of the Xinfadi Market, Beijing (China)

Sui Zhang<sup>1</sup> , Zhao Yang<sup>1</sup>, Minghao Wang<sup>1</sup>, and Baolei Zhang<sup>1</sup> 

<sup>1</sup>School of Geography and Environment, Shandong Normal University, Jinan, China

**Abstract** The frequent occurrence of local COVID-19 today gives a strong necessity to better understand the effects of "source-case" distance and gathering places, which are often considered to be the key factors of the localized spatial clustering of an epidemic. In this study, the localized spatial clustering of COVID-19 cases, which originated in the Xinfadi market in Beijing from June–July 2020, was investigated by exploring the spatiotemporal characteristics of the clustering using descriptive statistics, point pattern analysis, and spatial autocorrelation calculation approaches. Spatial lag zero-inflated negative binomial regression model and geographically weighted Poisson regression with spatial effects were also introduced to explore the factors which influenced the clustering of COVID-19 cases at the micro spatial scale. It was found that the local epidemic can be significantly divided into two stages which are asymmetric in time. A significant spatial spillover effect of COVID-19 was identified in both global and local modeling estimation. The dominant role of the "source-case" distance effect, which was reflected in both global and local scales, was revealed. Relatively, the role of gathering places is not significant at the initial stage of the epidemic, but the upward trend of the significance of some places is obvious. The trend from "distance-driven" to "density-driven" of the localized spatial clustering of COVID-19 was predicted. The effectiveness of blocking the transformation trend will be a key issue for the global response to the local COVID-19.

**Plain Language Summary** The frequent occurrence of local COVID-19 is becoming a global public health problem today. Most previous studies have focused on the influence of the "density" of socio-economic and demographic factors on the spatial clustering of COVID-19 (named "density-driven" in this study) while ignoring the effect of the spatial distance between the source of the local COVID-19 and the cases (named "distance-driven" in this studies). Therefore, this manuscript took the local COVID-19 cases which originated from the Xinfadi market from June–July 2020 as the study case. To obtain more accurate results, two improved quantitative research methods are adopted. It was found that the roles which the "source-case" distance played were far more decisive than the scholars thought before. We also found that the local epidemic had a significant trend of changing from the "distance-driven" diffusion mode to the "density-driven" diffusion mode, and the critical point between these two driving modes will be one of the keys to prevent and control the clustering.

## 1. Introduction

Since the first confirmed cases of COVID-19 were discovered in Wuhan, China in December 2019, the COVID-19 pandemic, which is highly infectious and has a high mortality rate, has spread throughout the world (F. Wu et al., 2020). The pandemic has brought great challenges in terms of disease treatment, prevention and control, economic development, social stability, environmental governance, and other aspects (Hamidi & Zandiatashbar, 2021; Mahato et al., 2020; Zhao et al., 2020). As of February 2021, more than 100 million people worldwide have been infected with the new coronavirus (Organization, 2021).

Following the outbreak of the COVID-19 epidemic, researchers have carried out extensive studies on the mode of epidemic spread (Franch-Pardo et al., 2020; Sohrabi et al., 2020). In summarizing previous studies,

**Visualization:** Sui Zhang, Minghao Wang

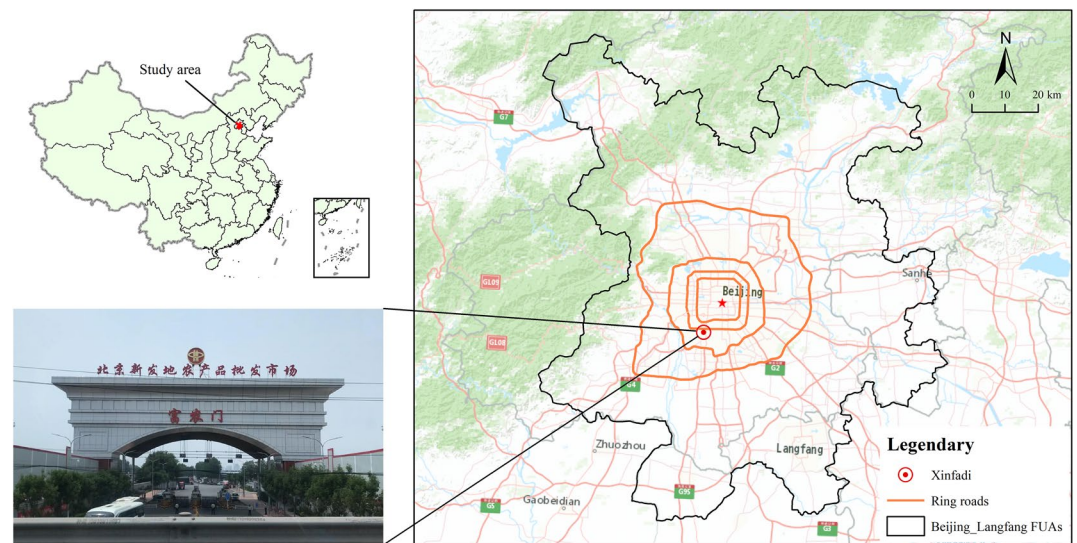
**Writing – original draft:** Sui Zhang

**Writing – review & editing:** Sui Zhang, Baolei Zhang

the transmission of COVID-19 at a local level may be described as involving “case clustering + mobile diffusion” (Q. Li et al., 2020). Infected people may then act to form a secondary point source whereby the outbreak spreads via community transmission, which then can result in a cascade of further outbreaks through the population at large, hence, causing an exponential increase in infections in densely populated cities (Copiello & Grillenzoni, 2020; Kraemer et al., 2020; Zhou et al., 2020). Given that most of the epidemics in the world today are “density-driven,” it may be readily noted that in most countries and regions, the centers of gravity of the epidemic are highly matched with the centers of “density” of various socio-economic factors, such as population density, the intensity of economic activities, density of population movements, and density of gathering places (Kadi & Khelifaoui, 2020; Martins et al., 2020; Yu et al., 2021). Studies, although, have found that perhaps due to China’s strong and effective COVID-19 prevention and control measures in the early part of the outbreak nationwide, the epidemic exhibited a “counter-intuitive” relationship concerning population density (Byass, 2020). However, as more studies based on different scales and periods were conducted and where different methods were adopted, it has been shown that the social-demographic “density” factors are the dominant factors that control the subsequent development of COVID-19 (Sun et al., 2020; J. Zhang et al., 2020). Thus, it is obvious that areas with high “densities” of these factors will significantly impact the direction of transmission of the epidemic, making the final center of gravity of the epidemic shift toward those areas with high “density.”

However, the spread of infectious diseases in the spatial dimension is typically not simply related to the socio-economic, demographic, environmental, and location factors called “density-driven” factors in this paper. Distances from the inflected cases to the epidemic sources (“source-case” distance) also play an important role in determining the direction of COVID-19 transmission and the degree of spatial clustering, topics that need more scientific investigation. However, due to the rapid and severe globalization of COVID-19, the sources for most of the areas affected by the epidemic have often been difficult to identify and define because these so-called “sources” are often mobile or derived from other areas (Chan et al., 2020; D. Wang et al., 2020). Therefore, much of the literature has avoided discussing the impact of the epidemic sources and the “source-case” distance on the development of the epidemic, thus knowledge regarding the impact of the sources of COVID-19 as well as its relationship with those “distance-driven” factors on the spatial distribution and transmission modes of the epidemic remain somewhat vague (Han et al., 2021). In addition, because COVID-19 is highly contagious among people, the “spatial spillover effect” also can not be ignored in spatial epidemiology, especially COVID-19 research (Franch-Pardo et al., 2020; Kraemer et al., 2020). The spatial spillover effect is proposed based on Tobler’s First Law, spatial regression models incorporate the “spatial lag term” of the dependent variable distributed in the spatial matrix according to a certain weight into the model and its coefficient can be calculated and used to characterize the role of spatial proximity and interaction intensity in epidemiological spatial diffusion, which can complement the index system of spatial epidemiological modeling and make the modeling results more accurate and reasonable (Getis & Ord, 1992; Getis & Aldstadt, 2004; Tobler, 1970). However, different from a large number of studies with clear results on spatial spillover effect at the macro scale, the micro-scale evidence of spatial effect in COVID-19 has not been demonstrated quantitatively, which undeniably deserves more attention (Mollalo et al., 2020; Sannigrahi et al., 2020).

Therefore, the relative neglect of such geographical “source” and spatial spillover effect associated research on COVID-19 infection reflects the fact that most regional governments are reference lacking such that they invariably have to adopt mandatory global lockdown or large-scale investigational and quarantine measures, which makes it difficult to control the socio-economic and resource costs while operating under conditions that seek to limit and control the spread of the epidemic (Elvidge et al., 2020; Sarkodie & Owusu, 2020). Based on this, it is evident that the identification and causal analysis of areas where case clustering (especially at the individual or community-scale) occurs are of great significance for the prevention and control of an epidemic with the costs as lower as possible in a region (Dungan et al., 2002). However, existing research on the epidemic status of COVID-19 has focused mainly on the epidemiological characteristics, the analysis of the case characteristics at the provincial and the municipal scale, and an evaluation of prevention and control policies, and much less attention has been given to the spatiotemporal clustering characteristics at the community level (Byass, 2020; Chen et al., 2021). Although some researchers have sought to reveal, from a geographical perspective, the pattern of the spread of COVID-19 at the micro-scale level, such work has focused mainly on the spatiotemporal evolution characteristics of the epidemic and its



**Figure 1.** The Beijing\_Langfang functional urban areas (FUAs) and the location of the Xinfadi market.

impact on the economy, society, and the environment, whilst ignoring the factors and their spatiotemporal heterogeneity responsible for driving the spread of the epidemic at the micro-scale, which played a huge role in the cost assessment of the disaster but did not try to reduce the cost at the source of the problems (the “reckless” global lockdown measures due to the lack of understanding of the characteristics and dynamics of localized spatial clustering of the COVID-19; X.-Y. Wang et al., 2021).

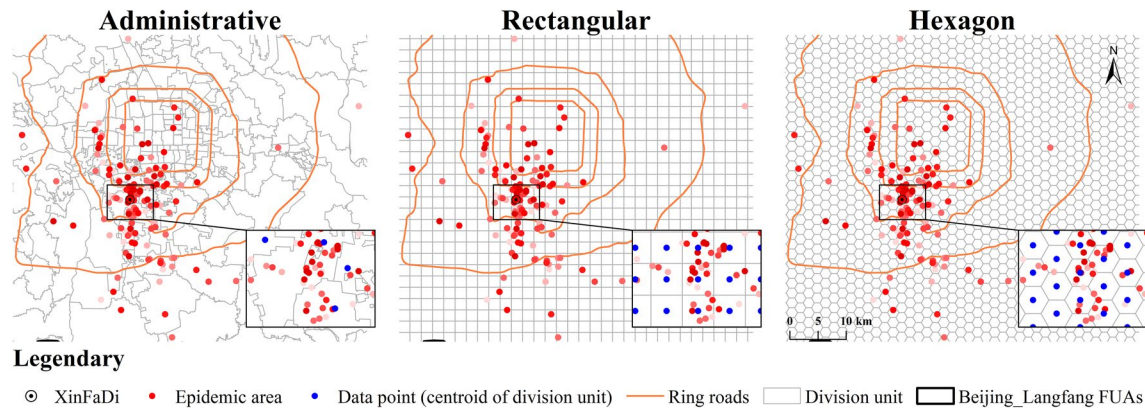
The present research was concerned with the investigation of the localized spatial clustering of COVID-19 cases, which originated from the Xinfadi market in Beijing from June to July 2020. The main objectives of research included the following: (a) to explore the spatiotemporal characteristics of localized spatial clustering of COVID-19 by descriptive statistics, point pattern analysis, and spatial autocorrelation analysis methods, (b) to introduce the spatial lag zero-inflated negative binomial regression model (SL-ZINB) and the geographically weighted Poisson regression with spatial lag approach (GWPR-SL) to explore the spatial spillover effects and the influence of the “source-case” distance and the places where people gather such that localized spatial clustering COVID-19 at the micro-scale occurred, and (c) to discuss the roles of “source-case” distance (distance-driven) and the places where people gather (density-driven) in terms of how they affect the development of the epidemic and the temporal variation trends.

## 2. Data Collection and Methodology

### 2.1. Study Area

Beijing (115° 25′–117° 30′E, 39° 28′–41° 05′N), as the capital city, is the center of politics, economics, and culture in China. According to a report issued by the Center for Disease Control and prevention of China, a local scale COVID-19 cluster was observed on June 11, 2020, after a 56-days absence of new confirmed cases. As of 5th July, there were 337 confirmed COVID-19 cases. All cases were directly or indirectly linked to the Xinfadi market, which was considered to be the single source of the outbreak (P. Liu et al., 2020; Pang et al., 2020; X.-Y. Wang et al., 2021; Y. Zhang et al., 2020). The Xinfadi market, a fixed spatial location for the source of a localized spatial clustering of COVID-19, makes this case different from the “mobile source” cases mentioned above hence as a consequence this case has high representativeness and high research value. The Xinfadi market, located in the Fengtai District of Beijing, covers an area of 112 hm<sup>2</sup>. It is the largest wholesale market of agricultural products in China and Asia (Figure 1). There are 5,526 fixed stalls, more than 8,000 fixed points (individuals) customers, a daily average traffic flow of more than 30,000 vehicles, and a passenger flow of more than 50,000 people. To explore the spread of COVID-19, which is controlled mainly by population mobility, the functional urban areas (FUAs) delineation results for the Beijing\_Langfang FUAs using the massive Didi ride-hailing records (Ma & Long, 2020) were used. Generally, there are





**Figure 2.** The spatial distribution of the confirmed cases and the schematic diagram of nearest-neighbor interactions under various study unit division methods: Deep red points represent early confirmed epidemic areas, while light red points represent later confirmed epidemic areas.

three methods to divide spatial study units: administrative regions, rectangular grids, and hexagon grids. Based on the suggestion of Kwan (2018), this paper divides the study unit to the minimum as much as possible (Kwan, 2018). To make the generation of spatial adjacency lag term in the following spatial modeling more scientific, we abandoned the administrative division and rectangular division with great differences between spatial lag terms for these differences will make the spatial lag terms with similar or even the same results have very different properties, which is unfavorable to the rationality of subsequent analysis (Figure 2). The Beijing\_Langfang FUAs were divided into 4,200 hexagon grids with a length of 1 km which conforms to the size of most human travel ranges during the pandemic and based on a referral to the idea of implementing a community grid epidemic prevention and control strategy (Birch et al., 2007; Z. Li & Gao, 2020; Ling & Wen, 2020).

## 2.2. Data Collection and Selection of Variables

The data used in this study include the data for COVID-19 cases and Point of Interest (POI) data. The data for COVID-19 cases were obtained from the Beijing Municipal Health Commission ([wjw.beijing.gov.cn/English](http://wjw.beijing.gov.cn/English)) and the Chinese Center for Disease Control and Prevention ([www.chinacdc.cn](http://www.chinacdc.cn)). The information on COVID-19 cases included age, gender, work status, and address of all cases from 11th June to 5th July 2020. The POI data were obtained from POIbase ([www.poibase.com](http://www.poibase.com)) and deviations were corrected manually. The catering places, residential areas, shopping places, public service facilities, and health-care facilities were selected as the places where people gather. All these variables have variance inflation factor (VIF) values of less than 7.5, which can be accepted by models with strong interpretability. The coordinate system for the above data was CGCS2000 with a Gauss Kruger projection and a 3-degree zone projection (40 zones).

## 2.3. Methodology

### 2.3.1. Statistical Analysis

The exponential modified Gaussian function (EMG), which is usually used in the peak fitting calculation of chromatographic peaks, was used to identify the temporal trends and the turning point of the local COVID-19 outbreak (Grushka, 1972). The EMG has the asymmetry which the standard Gauss function does not have for its convolution of the standard Gaussian and exponential function. In addition, sufficient studies have confirmed that due to the lockdown measures taken by China, the local COVID-19 epidemic often showed significant asymmetry temporally, thus the temporal characteristics of the epidemic can be reflected more accurately by this specific method (Ji et al., 2020; K. Wang et al., 2020). The days before the turning point are called the spread duration (SD), and the days after the turning point are called the decay duration (DD). Given that the value corresponding to the turning point of the fitting curve is usually not an integer, considering the similarity of the case characteristics near the turning point, an extra whole day of cases was taken for both two durations in the statistical analysis. The chi-square test ( $\chi^2$ ) was used to make

descriptive statistics on a series of demographics and the Xinfadi market-associated characteristics of cases between the SD and the DD.

### 2.3.2. Point Pattern Analysis

The standard deviational ellipse (SDE) has long been a general GIS tool to describe the geographical distribution of features. It describes the trend in the geographical distribution by summarizing the dispersion and the directivity of the observed samples (Lefever, 1926). The SDE can make a basic judgment on the direction of the spread and the characteristics of the COVID-19 cases and provide a theoretical basis for the determination of the key direction of epidemic prevention and control for the spatial dimension (S. Liu et al., 2020). Besides, Ripley's K function is also an effective tool to measure the distribution characteristics of point elements. It is a local second-order point pattern analysis approach that is, used to evaluate the distribution characteristics (clustered, random, or dispersed) of spatial point features on multiple spatial scales. It is suitable for quantitative analysis and comparison of the distribution patterns of COVID-19 cases (Boots & Getis, 1988; Getis, 1984; Wiegand & Moloney, 2004). After the L(d) transformation, the formula is as follows:

$$L(d) = \sqrt{\frac{A \sum_{i=1}^n \sum_{j=1, j \neq i}^n k_{i,j}}{\pi n(n-1)}} \quad (1)$$

where  $A$  is the total area of the features;  $d$  is the search radius;  $n$  is the total number of confirmed cases;  $k_{i,j}$  are binary piecewise functions added to set the distance threshold.

### 2.3.3. Spatial Autocorrelation Analysis

The Global Moran's I index of global spatial autocorrelation analysis approach was used to analyze the global spatial distribution of the local COVID-19 cases (Getis & Ord, 1992; Griffith, 1987a, 1987b), which can be calculated as follows:

$$I_{Global} = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{i,j}} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{i,j} (x_i - \bar{X})(x_j - \bar{X})}{\sum_{i=1}^n (x_i - \bar{X})^2} \quad (2)$$

where  $w_{i,j}$  is the spatial weights matrix between  $i$  and  $j$  (in this paper, we build a  $k = 1$  queen spatial adjacency matrix based on a hexagonal grid; this will not be explained here). To better understand the spatial difference of the spatial autocorrelation in the study area, we used the local spatial autocorrelation tool, the Anselin Local Moran's I (LISA) index, to identify spatial clusters and outliers (Anselin, 2010), and to analyze and visualize the layout of the local spatial autocorrelation of COVID-19:

$$I_{Local} = \frac{n(x_i - \bar{X})}{\sum_{i=1}^n (x_i - \bar{X})^2} \sum_{j=1}^n w_{ij} (x_j - \bar{X}) \quad (3)$$

The grids of the study area were divided into five categories based on Anselin local Moran's I analysis: high-high (H-H) agglomeration, low-low (L-L) agglomeration, high-low (H-L) outlier, low-high (L-H) outlier, and not significant. The above spatial analysis and visualization were based on the use of Geoda and ArcGIS 10.7 software.

### 2.3.4. Spatial Lag Zero-Inflated Negative Binomial Regression Model

To explore the spatial patterns and impact factors of the population flow-oriented events such as in the COVID-19 pandemic, selection of the data processing approaches used in counting are normally relevant to the accuracy of the research, especially at the micro-scale where results might easily be affected by data processing errors. The Gravity Model, a mathematical model widely used in the study of human activities, has unique advantages in the study of an "origin-destination" type of question (Stewart, 1948). However, Flowerdew and Aitkin (1982) believed that the regression results are sensitive to the zero value flow (Flowerdew & Aitkin, 1982). Considering the limitation above and the non-negative integer characteristics

**Table 1**  
Statistical Description of the Variables

Variables	Obs	Sum	Mean	Std. Dev.	Min	Max
Number of cases	4,200	337	0.0802381	1.566799	0	72
Catering	4,200	8,534	2.031905	8.375729	0	155
Residences	4,200	28,485	6.782143	17.90986	0	174
Shopping	4,200	32,818	7.81381	30.50214	0	597
Public services	4,200	26,599	6.333095	23.26272	0	430
Health-care	4,200	11,046	2.63	9.120022	0	163

of population flow, the Poisson model, and the negative binomial model (NB), as well as the zero-inflated model, have been considered by the academic community to replace the traditional gravity model (Boyle & Flowerdew, 1993; Flowerdew & Boyle, 1995; Lambert, 1992). The estimation process for the zero-inflated model consists of a Logit regression and a Poisson regression or a negative binomial regression so that it can better solve the processing problem of excess zero values (Burger et al., 2009; Lambert, 1992). Taking the zero-inflated negative binomial regression model (ZINB) with a wider application value as an example:

$$\Pr(n_{ij} = 0) = \psi_{ij} + (1 - \psi_{ij}) \left( \frac{\alpha^{-1}}{\alpha^{-1} + \lambda_{ij}} \right)^{\alpha^{-1}} \quad (4)$$

$$\Pr(n_{ij} \neq 0) = (1 - \psi_{ij}) \frac{\Gamma(n_{ij} + \alpha^{-1})}{n_{ij}! \Gamma(\alpha^{-1})} \left( \frac{\alpha^{-1}}{\alpha^{-1} + \lambda_{ij}} \right)^{\alpha^{-1}} \left( \frac{\lambda_{ij}}{\alpha^{-1} + \lambda_{ij}} \right)^k \quad (5)$$

where  $n_{ij}$  is the interaction intensity or tendency between “origin  $i$ ” and “destination  $j$ ”;  $\psi_{ij}$  is the proportion of observations with a strictly zero count ( $0 < \psi_{ij} < 1$ ), which is determined by the logit part of the model;  $\alpha$  is an overdispersion coefficient; the bigger the value of  $\alpha$ , the higher the overdispersion level. When  $\alpha = 0$ , that is, when there is no overdispersion problem, the negative binomial regression model will degenerate into a Poisson regression model; the conditional mean  $\lambda_{ij}$  is related to the exponential function of the regression explanatory variable:

$$\lambda_{ij} = \exp \left( \beta_0 + \sum_{q=1} \beta_{1q} \ln P_{iq} + \sum_{l=1} \beta_{2l} \ln P_{jl} - \beta_3 \ln d_{ij} \right) \quad (6)$$

where  $\beta_{1q}$  is the coefficient of the  $q$ th variable of the “origins” ( $P_{iq}$ );  $\beta_{2l}$  is the coefficient of the  $l$ th variable of the “destinations” ( $P_{jl}$ );  $\beta_3$  is an impedance factor reflecting the distance decay in euclidean distance  $d_{ij}$ ;  $\beta_0$  is the constant term. Considering the frequent occurrence of a zero value in the explanatory variables for the gathering places, natural logarithm operations cannot be performed. To ensure successful calculation by the model and the reliability of the results, the data are translated to a certain extent, and the amount of translation is usually 1 (Table 1).

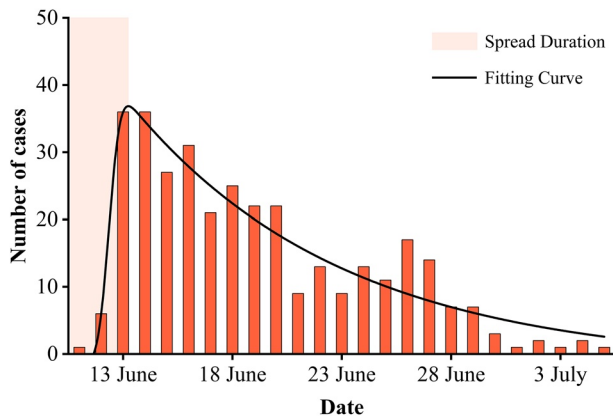
Besides, Considering the spatial autocorrelation and spatial spillover effect of an epidemic with clustering distribution, it is precisely this problem that needs special consideration and discussion at the micro-scale research level in the case of COVID-19 (Gatrell & Bailey, 1996). By incorporating a spatial lag term of the dependent variable into the negative binomial regression part of the model, the spatial lag zero-inflated negative binomial regression model is built to further refine the modeling results for the spatial spillover effect, and formula 6 is transformed into (Cai et al., 2016; Metulini et al., 2018; Middela & Ramadurai, 2020):

$$\lambda_{ij} = \exp \left[ \beta_0 + \sum_{q=1} \beta_{1q} \ln P_{iq} + \sum_{l=1} \beta_{2l} \ln P_{jl} + \delta w_{i,j} (\ln \lambda_{ij}) - \beta_3 \ln d_{ij} \right] \quad (7)$$

where  $w_{i,j}(\ln \lambda_{ij})$  is the spatial lag term of the dependent variable;  $\delta$  is the coefficient of the spatial spillover effect, and when  $\delta = 0$ , the models degenerate into the non-spatial zero-inflated count regression. To save the number of repeated calculations,  $\sum_{q=1} \beta_{1q} \ln P_{iq}$  can usually be omitted in the application of the “single-origin” case.

### 2.3.5. Geographically Weighted Poisson Regression With Spatial Lag

The spatial perspective of health issues usually has two remarkable properties: spatial heterogeneity and spatial autocorrelation, and the spatial autocorrelation is usually measured by the spatial spillover effect mentioned above (Griffith, 1987a; Kwan, 2021). In addition to the spatial spillover effect, some local spatial modeling methods, such as geographically weighted regression (GWR) and its extended model, can



**Figure 3.** Daily increase of COVID-19 cases in Beijing\_Langfang functional urban areas (FUAA) from June 11 to July 5.

effectively detect the spatial heterogeneity of the determinants of spatial epidemics (Feuillet et al., 2015; Nakaya et al., 2005). However, little attention was paid to the intersection of these two properties, that is, the spatial heterogeneity of the spatial spillover effect, which is also crucial to spatial epidemiology (Brunsdon et al., 1998). Therefore, the spatial lag term is incorporated into the GWPR to construct the GWPR-SL model, which can provide the spatially varying coefficient of the spatial spillover effect and help researchers to obtain a deeper understanding of this phenomenon. The model is described as follows:

$$\lambda_{ij} = \exp \left[ \begin{aligned} &\beta_0(u_k, v_k) + \sum_{q=1} \beta_{1q}(u_k, v_k) \ln P_{iq} + \sum_{l=1} \beta_{2l}(u_k, v_k) \ln P_{jl} \\ &+ \delta(u_k, v_k) w_{i,j} (\ln \lambda_{ij}) - \beta_3(u_k, v_k) \ln d_{ij} \end{aligned} \right] \quad (8)$$

where  $(u_k, v_k)$  is the spatial coordinates of hexagon  $k$ ; When  $\delta = 0$ , the models degenerate into the general GWPR model. Due to the endogenous problem caused by the strong relationship between the dependent variable and its spatial lag term, a specific two-stage estimation approach, in

which the first stage is linear regression and the second stage is Poisson regression, is set due to the continuous variable property of the spatial lag term of the dependent variable and the discrete variable property of the dependent variable itself, is used in this model: instrumental variable  $P_z$  is used to construct a proxy variable in the estimation process for model endogeneity. We followed the suggestion by previous studies that for a matrix  $M$  containing  $\ln P_{ij}$ ,  $P_z = w_{i,j} M$ , which is the standard candidate used in spatial two-stage estimation to construct the instrument variable. (Geniaux & Martinetti, 2018; Shoff et al., 2014). The spatial econometric modeling and visualization were based on the Stata 15.1, R and ArcGIS 10.7 software.

### 3. Results

#### 3.1. Statistical Analysis

The daily increase and subsequent decrease of the number of COVID-19 cases may be visualized via a temporal analysis (Figure 3). According to the curve fitting by the EMG model, the turning point of the epidemic was calibrated between June 13 and June 14. At 3:00 a.m. on June 13, Fengtai District announced that the Xinfadi market was temporarily closed for comprehensive sanitation and environmental disinfection, which cut off the local epidemic transmission path and controlled the epidemic (Pang et al., 2020). This is in good agreement with the turning point obtained by the EMG model. Due to the strict and effective intervention measures taken by the regional government in the early stages of the epidemic such as large-scale SARS-CoV-2 nucleic acid detection and quarantining of close contact individuals, the duration of spread was short-lived and the rate of detection of infected individuals was efficient; these time frames contrasted with the duration of the decay of the infection which was of relatively long duration (Kucharski et al., 2020; Wilder-Smith & Freedman, 2020).

The overall characteristics of the 337 confirmed cases and the associations between the two development phases are outlined in Table 2. There were significant differences in the three variables in terms of whether (a) the cases were visitors to the Xinfadi market, (b) the cases were employees of the market, or (c) the cases were close contacts of individuals infected during the different periods of the epidemic. Regarding the duration of spread and the occurrence of cases, the infected individuals were most likely to be visitors to the market, while over the decay period, the characteristics of the cases tended to suggest that individuals had more close contacts with people away from the market rather than having contact with infected cases directly at the Xinfadi market. Most of the close contacts may be categorized as second or third-generation family members of the infected persons from SARS-CoV-2, hence the trend characteristics have indirectly demonstrated the hierarchy of the COVID-19 transmission. This difference in the sources of the cases results in an essential heterogeneity for the two-time durations of the epidemic, which will also be reflected in the follow-up analysis and discussion. It is worth noting that the age variability associated with COVID transmission that has been widely confirmed to affect the probability of COVID-19 infection is not significant for the change in the duration of this localized spatial clustering epidemic; it was found that there



**Table 2**  
*General Characteristics of Confirmed Cases in the Different Durations of COVID-19*

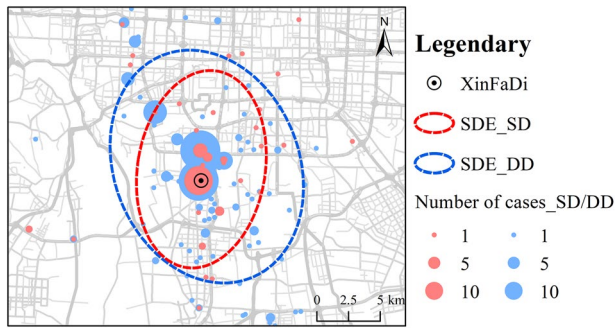
Variables	Duration of spread( <i>n</i> = 79)	Duration of decay( <i>n</i> = 330)	Total	P-value
<b>Gender</b>				
Female	30 (46.15)	146 (45.20)	176 (45.36)	0.888
Male	35 (53.85)	177 (54.80)	212 (54.64)	
<b>Age</b>				
<15	0 (0.00)	5 (1.82)	5 (1.47)	0.202
15–64	65 (100.00)	262 (95.27)	327 (96.18)	
≥65	0 (0.00)	8 (2.91)	8 (2.35)	
<b>Native place</b>				
No	59 (85.51)	184 (82.88)	243 (83.51)	0.608
Yes	10 (14.49)	38 (17.12)	48 (16.49)	
<b>Visitor to XFD</b>				
No	2 (2.99)	34 (14.23)	36 (11.76)	0.012*
Yes	65 (97.01)	205 (85.77)	270 (88.24)	
<b>Employee of XFD</b>				
No	11 (18.33)	65 (30.95)	76 (28.15)	0.055
Yes	49 (81.67)	145 (69.05)	194 (71.85)	
<b>Contact with infected cases</b>				
No	39 (49.37)	39 (15.06)	78 (23.08)	0.000**
Yes	40 (50.63)	220 (84.94)	260 (76.92)	
<b>Family members of infected cases</b>				
No	67 (94.37)	254 (94.42)	321 (94.41)	0.786
Yes	4 (5.63)	15 (5.58)	19 (5.59)	
<b>Work</b>				
No	2 (3.17)	17(7.91)	19(6.83)	0.305
Yes	61 (96.83)	198(92.09)	259(93.17)	
<b>Clinical type</b>				
Mild	28 (43.08)	80 (27.12)	108 (30.00)	0.055
Common	36 (55.38)	210 (71.19)	246 (68.33)	
Severe	0 (0.00)	3 (1.02)	3 (0.83)	
Asymptomatic	1 (1.54)	2 (0.68)	3 (0.83)	

\**p* < 0.05.  
\*\**p* < 0.01.

were only a few infected elderly people and children, who are usually judged as susceptible groups, and this may be related to the relatively low level and range of activities of these cohorts (Clark et al., 2020; Davies et al., 2020; Dong et al., 2020).

### 3.2. Point Pattern Analysis

For the duration of spread and decay, the spatial heterogeneity of COVID-19 is illustrated in Figure 4. Regarding the duration of spread, the SDE of COVID-19 presented in a northeast-southwest direction, while for the duration of decay, the SDE of COVID-19 exhibited a northwest-southeast distribution. In addition, the clustering characteristics for the two distributions were also significantly different as expressed by the eccentricity whereby the SDE for the duration of decay was less than that for the duration of the spread, even though the difference was not as significant as it was for the directionality. Similarly, the calculation

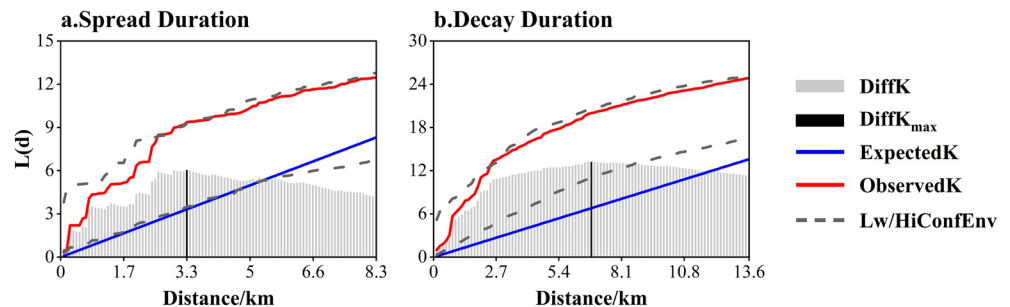


**Figure 4.** The standard deviational ellipse of COVID-19 cases in the two durations.

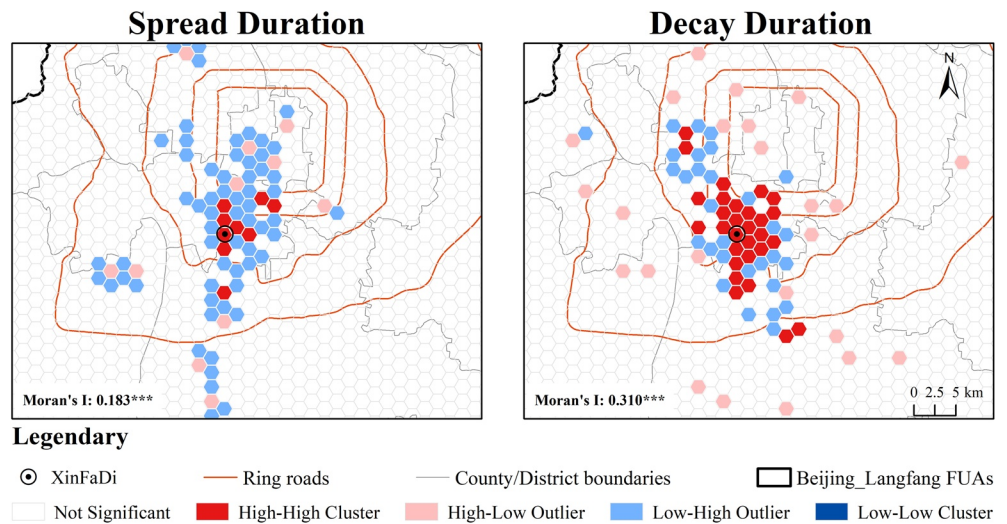
of Ripley's K function also confirmed a change of agglomeration. We not only calculated the K function after  $L(d)$  transformation but also visualized the difference between the observed value and the expected value (DiffK) to observe the change of the maximum value (Figure 5). The comparison of the maximum value of DiffK for the two durations showed that the radius of clustering obtained by standard circle measure in space was about 3 km for the duration of spread, while the maximum value reached 6 km for the duration of decay. It is also worth noting that in the case of the clustering characteristics for the two-stage epidemic, the results for quantitative analysis indicated that the clustering intensity of COVID-19 was as high as 6 and 12, respectively. However, the characteristics of the clustering distributions are still not significantly higher than their confidence interval. These unconventional results in point pattern analysis reflect the importance and the uncertainty of the "source-case" distance in the study of the localized spatial clustering of COVID-19 from a relatively one-sided perspective.

### 3.3. Spatial Autocorrelation Analysis

The global spatial autocorrelation analysis of COVID-19 cases in the two durations showed that COVID-19 had a very significant spatial positive autocorrelation, which provided a theoretical basis for introducing into the econometric model the spatial lag term of the dependent variable. The results of local spatial autocorrelation were also calculated and visualized (Figure 6). There were much more L-H outlier grids and fewer H-H cluster grids in the calculation result of the spread duration than the other one, and they present a narrow spatial layout from north to south. Because of the rapid emergency measures taken by Beijing Municipal Government, the vast majority of cases in the early stage of the epidemic are clustered near the Xinfadi market and only a few cases are scattered in areas far away from the source of the local epidemic. Therefore, the spatial diffusion scale was temporarily small and generally limited within the community, and the outbreak areas presented isolated and sporadic distribution, resulting in more L-H and H-L outlier grids. In the decay duration of the epidemic, the numbers for the H-H cluster, the H-L outlier, and the L-H outlier grids were relatively large, and they were concentrated in the areas near the Xinfadi market and even within the Sixth Ring Road, which showed that the Xinfadi market and its surrounding areas were the areas where the outbreak was concentrated, and the number of isolated infected areas decreased with distance from the Xinfadi market. Meanwhile, the visualization results for the local spatial autocorrelation revealed other spatial morphological characteristics of the epidemic in that clustering of the COVID-19 cases did not follow a standard distance decay path following the geometric space but had clear directionality, which suggested that location factors played an important role in the spread of the epidemic.



**Figure 5.** The Ripley's K function analysis results of COVID-19 cases in the two durations.



**Figure 6.** The global Moran's index and the Anselin Local Moran's I map of COVID-19 cases in two durations.

### 3.4. Cross-Sectional Modeling Result

As the proportion of non-zero data of the dependent variable of the regression model is less than 2%, it can be seen that the log-likelihood test and the information criterion test results of the gravity model under the OLS framework are deceptive and biased. Although the Vuong test has been widely used as a necessary test in research and application in the zero-inflated model for a long time, Wilson (2015) concluded that the Vuong test is not suitable for testing the zero-inflated model. Therefore, the advantages and disadvantages of the zero-inflated model based on the Akaike Information Criterion (AIC) (Vuong, 1989; William, 1994; Wilson, 2015) were examined in this study. Through the deduction of the Poisson regression and the negative binomial regression to zero inflation negative binomial regression, the estimates of the model based on the AIC are significantly reduced. Given the significant spatial autocorrelation of the dependent variable and the significant results of the (robust) Lagrange multiplier test, the spatial interaction of the dependent variable is incorporated into the model in this study. The SL-ZINB model with better model fitness is then constructed to estimate and discuss the distance effect, the spatial interaction, and the location factors which influence the COVID-19 distribution.

The results of estimating the impact factors on the cross-sectional data of the COVID-19 distribution are presented in Table 3. Undoubtedly, the significance of the distance variable for the source of the COVID-19 in this epidemic is most prominently reflected in the model, and it played an extremely significant negative role in the spread of COVID-19. Moreover, the spatial spillover effect of COVID-19 at the micro-scale not only exists but also maintains a high coefficient at a high significance level, which means that the spatial spillover effect of COVID-19 at the micro-scale cannot be easily dismissed. The variables for the places where people gather also have a strong statistical explanatory power. The increase in the density of the locations of the public service facilities and healthcare facilities can significantly promote the development of the epidemic. It is worth stating that restaurants and catering facilities, and residential areas, have shown atypical effects for the outbreak, the extremely significant negative effects being contrary to what one would logically expect. In addition, even if all the cases included in this study had a fairly large association with the Xinfadi market and were accounted for by the residents' addresses, no significant effect was found for the variables associated with shopping places and residential areas variables in this outbreak.

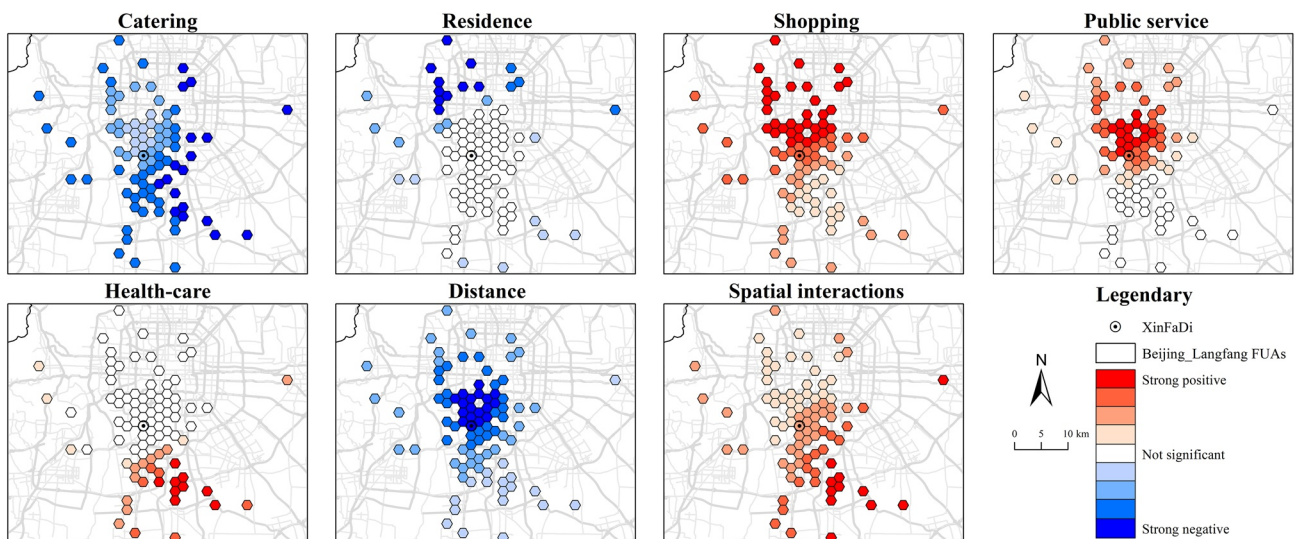
The strong heterogeneity in the spatial distribution of the local effects of the "source-case" distance and the places where people gather is visible (Figure 7). Consistent with the results of the global model, the spatial spillover effect of the local COVID-19 epidemic showed a significant promotion effect globally (Table 4). On this basis, the GWPR-SL model provides the spatial variability of the coefficient of this effect: the spatial spillover effect in the southeast of the outbreak area is stronger than that in the northwest. In other words, the southeast of the outbreak area should be the key area of spatial interaction prevention for infected

**Table 3**  
*Estimation Results for the Five Regression Approaches*

Variable	OLS	Poisson	NB	ZINB	SL-ZINB
ln (Catering)	0.0068	-0.2416**	-0.3302	-0.9630***	-1.0076***
ln (Residences)	-0.0176***	0.1689*	0.2810*	0.0159	0.0170
ln (Shopping)	-0.0025	0.0801	0.0638	0.4882*	0.3051
ln (Public services)	0.0145***	0.5078***	0.0767	0.4406*	0.5321**
ln (Health-care)	0.0067	0.0698	0.4297**	0.4298	0.5799**
ln (Distance)	-0.1266***	-2.2192***	-2.6344***	-1.3323***	-0.6685**
$W_{ij} * \ln (\text{Cases})$					1.0463**
Constant	1.3644***	17.5435***	21.6494***	10.3913***	4.0046
Log likelihood	1467.80	-475.11	-368.94	-348.69	-345.57
Alpha			2.53	1.11	1.11
AIC	-2921.60	964.22	753.88	727.37	723.15

\* $p < 0.1$ .  
\*\* $p < 0.05$ .  
\*\*\* $p < 0.01$ .

people in this local epidemic. The spatial effects of places where people gather and distance variables on the epidemic can be divided into two forms: global significance and local significance. First of all, the effect of “source-case” distance on the inside-out decay with the Xinfadi market as the core is obvious, but the decay effect of distance is more intense in the south of the outbreak area. Meanwhile, catering places and shopping places had strong positive effects on the occurrence of the north. In addition, residence areas, public service facilities, and health-care facilities had a significant local impact on the development of the epidemic: the significant effect of public service facilities have a relatively wide distribution, mainly clustered in the vicinity of the northern outbreak, while residence areas and health-care facilities have inhibited and promoted the development of the epidemic in the northern and southern areas far from the source respectively. It is precisely because of the significant directional characteristics of the local effects that the basic pattern of the north-south distribution of the epidemic has been shaped. In general, the effect of “source-case” distance is centered on the source, and the significant effect area of the places is mainly distributed around the periphery of the outbreak area.



**Figure 7.** Spatial heterogeneity of the effect of factors on the occurrences of COVID-19.



**Table 4**  
*Estimate Summaries of Geographically 17 Weighted Poisson Regression With Spatial Effects Coefficients of the Variables*

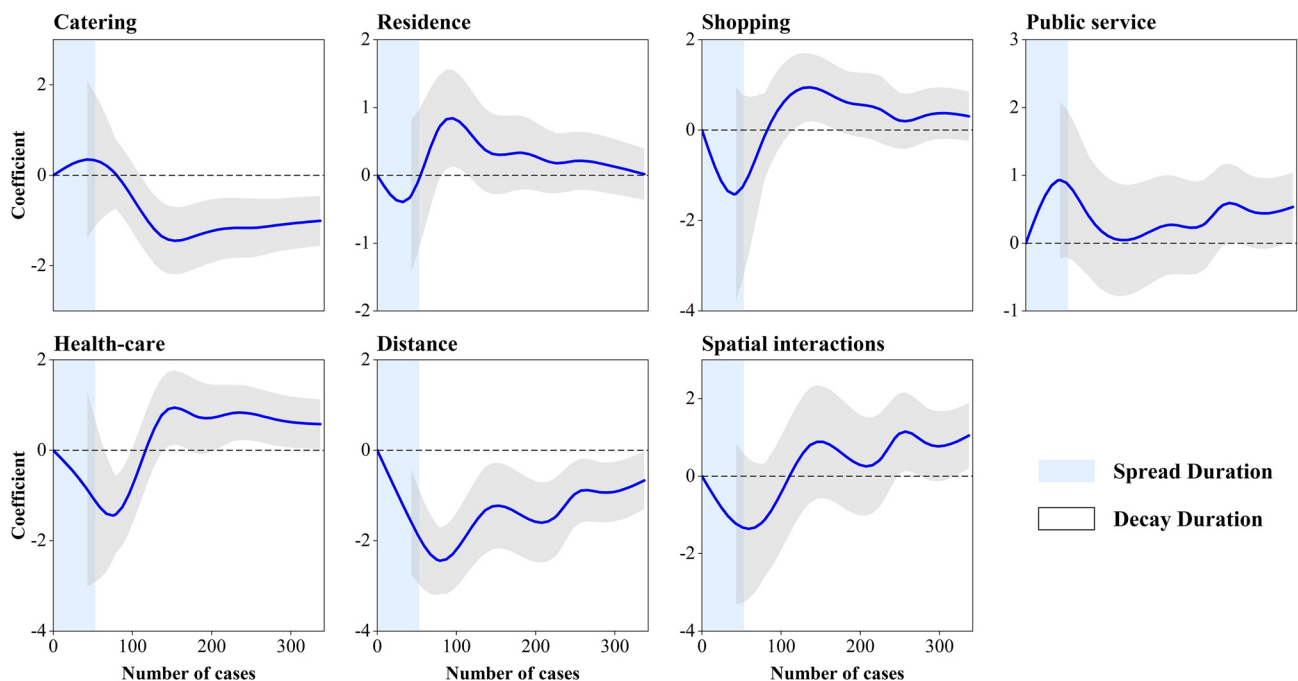
Variables	Minimum	Lower quartile	Median	Upper quartile	Maximum
ln (Catering)	-0.64	-0.60	-0.58	-0.54	-0.44
ln (Residences)	-0.31	-0.24	-0.20	-0.17	-0.14
ln (Shopping)	0.33	0.40	0.51	0.59	0.64
ln (Public services)	0.03	0.22	0.49	0.66	0.92
ln (Hospitals)	-0.15	0.06	0.18	0.29	0.33
ln (Distance)	-1.52	-1.21	-1.02	-0.91	-0.72
$W_{ij} * \ln (\text{Cases})$	0.31	0.33	0.35	0.38	0.43
Constant	6.32	7.58	8.17	9.18	10.75
Kernel-type	Adaptive Gaussian				
AIC	189.46				

## 4. Discussion

### 4.1. Longitudinal Characteristics of Modeling

To have a deeper understanding of the relative relationship between the "source-case" distance and the place where people gather in COVID-19 and its dynamics, the results of the cross-section model are analyzed in time series. Because of the recommendation by statisticians that the number of samples used for statistical analysis should be not less than 30 to assure confidence, and to ensure the integrity of the case statistics and the rationale for the data that were acquired on the same date, the local COVID-19 epidemic was divided into several  $n > 30$ -time series analysis units (as described below) according to the number of new cases.

The negative number for the regression coefficient of the Euclidean distance variable in the gravity model is usually called the distance decay coefficient (Nekola & White, 1999). The calibration of the distance decay coefficient has not only a theoretical basis as a function in the application code, but it can also serve as an intuitive response in terms of how a dependent variable is affected by the spatial distance (Figure 8). In



**Figure 8.** Temporal variation of the effect of factors on the occurrences of COVID-19.

this epidemic, the distance decay coefficient increased sharply at first and then decreased, and the trend of the curve showed great differences in the two durations of the epidemic: the distance decay coefficient increased for the duration of the spread and decreased for the duration of the decay. When the epidemic approached the turning point, the distance decay coefficient also reached its maximum value of about 2.5, while for the overall epidemic, the distance decay coefficient was calibrated at 0.668. (Note that the distance decay coefficient in the field of urban studies is often considered to be between 1 and 2). The positive result for the duration of the spread showed the effectiveness of the measures taken by the Beijing government in the early stages of the epidemic. Comparisons of the results for the overall results and some related distance decay coefficient calibrations such as service facilities accessibility or population mobility did show certain differences (Luo & Qi, 2009; McGrail & Humphreys, 2009). Therefore, although the “source-case” distance variable in the localized spatial clustering COVID-19 was deemed significant, its damping effect was lower than that of conventional academic studies concerning the distance decay of population flow at the city level (Brockmann et al., 2006). In other words, the sprawl of the epidemic in space was faster than the population flows and more difficult to be constrained in the long-term by the Euclidean distance.

Besides, the spatial lag variables associated with COVID-19 also played a significant role in promoting the transmission and spread of COVID-19. The positive impact of the spatial interactions on the epidemic increased during the duration of the decay. In contrast, the model estimates that the spatial spillover effects during the duration of the spread were negative and insignificant. The turning point of the outbreak is also the turning point of the spatial spillover effect. This may have been due to the large number of cases occurring in a relatively small number of spatial units in the market such that rapid clusters of cases would have readily broken out in a short period and the spatial spillover effect across communities and grids began to play a significant positive role after the epidemic spread greatly in the spatial dimension (Z. Wu et al., 2020). Because of the limitation of the size of the space unit in this research, we did not try to explore further the microscopic spatial dimension of this phenomenon (Kwan, 2012). Nevertheless, the impact of the spatial lag variables on the development of the epidemic was not negligible.

Moreover, both the direction and the effect of the intensity of the variables associated with where people gather changed significantly before and after the turning point of the epidemic. In general, regarding the period of transmission, the effects of the places where people gathered often showed a sharp rise or decline in the trend of the duration of the spread, and this trend often changed abruptly around the turning point. Regarding the duration of the decay, the variability in the impact of these factors, in general, tended to be relatively modest and quite stable. Thus, the spread of the epidemic was somewhat inhibited over a significant distance, which may be a reflection of the fact that commercial catering activities were completely locked down by the Beijing government, and it also might relate to certain relationships associated with the urban spatial structure in Beijing (Lau et al., 2020; Tian et al., 2020). A more thorough understanding of this point needs to be followed up on in subsequent research.

Among the factors associated with the five locations (The catering places, residential areas, shopping places, public service facilities, and health-care facilities), the most positive effect on the epidemic arose from the locations of the public services facilities. This positive effect was not only reflected in the duration of the decay of the epidemic but also reflected in the duration of the spread of the epidemic, the difference being that the positive effect was strong but not significant in the early stage of the epidemic, while the positive effect was relatively weaker but more significant in the later stages of the epidemic. The number of residential areas, shopping places, and healthcare facilities played a negative role for the duration of the spread of the epidemic, but a relatively positive role in the overall COVID-19 epidemic. Although the impact of residential areas and shopping places on the overall development of the epidemic was not significant, the trend was clear. An interesting contrast was observed between the positive effect of hospital density on the local occurrence of COVID-19 and some research results on the impact of the distribution and accessibility of health facilities and medical resources locally; that is, medical resources have multiple effects on COVID-19, and this will undoubtedly stimulate more detailed quantitative and in-depth discussion in urban and emergency planning (Y. Wu et al., 2020).

The quantitative results presented above have been used to discuss the spatiotemporal characteristics of COVID-19 which originated at the Xinfadi market in Beijing from June to July 2020, as well as the impact of variables such as the “source-case” distance, the spatial interactions, and the places where people gather. In

terms of the statistical analysis, irrespective of whether, descriptive statistics, point pattern analysis, spatial autocorrelation analysis, or the temporal statistical results of the econometric model were used, all the characteristics of the variables before and after the turning point of this epidemic exhibited strong heterogeneity as described in the previous section. The results are very revealing for the geographers who often focus on the spatial dimension and tend to ignore the temporal characteristics of the development of the epidemic and the great heterogeneity in its spatial distribution and the impact of factors on the temporal evolution, which often leads to relatively one-sided research results (Rashed et al., 2020). Such a situation is not conducive to implementing rapidly effective prevention and control measures.

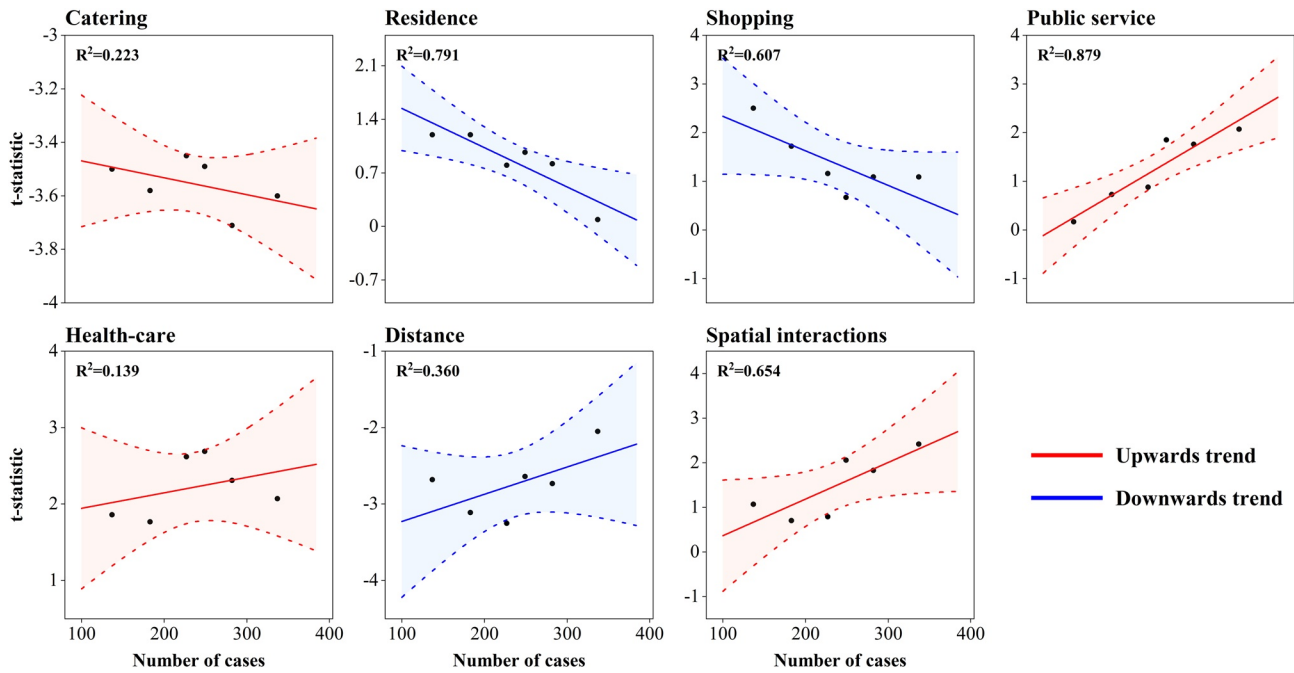
For the spatial dimension, based on the analysis of the spatial autocorrelation and zero-inflated model modeling results, there is sufficient evidence to demonstrate that the spatial spillover effect of the epidemic is quite significant, which undoubtedly is related to the basic nature of human-to-human transmission of COVID-19 (Q. Li et al., 2020). However, although the size of the grid is reduced as much as possible based on the actual situation and the hexagon grid is selected based on the equal spatial distance between neighbor grids to reduce the effect from the MAUP still exists. To address this, a series of research topics should be proposed, namely, more attention should be paid to the scale and morphology of research units (Dungan et al., 2002; Kwan, 2012). The estimated results of the factors associated with the epidemic will undoubtedly be greatly affected by the change of the significant spatial spillover effect caused by the scale and morphological changes of the study units (Y. Wang & Di, 2020). Therefore, in future research, more scientific methods and evaluation criteria are needed to detect the sensitivity of this issue in micro-scale epidemic research, to ensure the robustness and comparability of the results of epidemic-related research.

#### 4.2. Diffusion Mode Recognition: “Distance-Driven” Versus “Density-Driven”

From a statistical standpoint, it has been learned that the spread of the local outbreak at the Xinfadi market was mainly “distance-driven” rather than “density-driven” with high mathematical significance, which means the effect of the “source-case” distance on the development of the epidemic was far greater than the effect of places where people gather. This feature allows us to highlight in essence the importance of the role that “source-case” distance plays in the localized spatial clustering of the COVID-19 infection as being the driving force of the epidemic as distinct from other mobile-source or multi-source outbreaks. Different from previous “density-driven” studies, the results of the present study have analyzed a complete “distance-driven” epidemic and calibrated the distance decay effect of the spread of the epidemic.

In a similar study by Han et al. (2021), the distance decay coefficient was not discussed in quantitative terms in detail. Instead of using the gravity model, the correlation coefficient and the classical OLS model regression were used based on the raster data converted from vector data. This processing approach not only does not carry out certain mathematical transformations on the data according to their distribution to make it suitable for OLS regression but also “equates” the “source-case” distance factor with the place factor unintentionally. At the same time, due to a lack of emphasis on the particularity (localized spatial clustering) of the epidemic, a certain degree of biased estimation occurs, although the contribution made in their study in the analysis of the spatial heterogeneity of the impact factors is instructive (Han et al., 2021). Based on the limitations of their works, the present study was revised and the methodology was expanded to better reflect the decisive role of the “source-case” distance factor in the development of the localized spatial clustering of COVID-19. This remarkable “distance-driven” feature is most likely related to the rapid and strict actions taken by Beijing’s government in response to the COVID-19 outbreak.

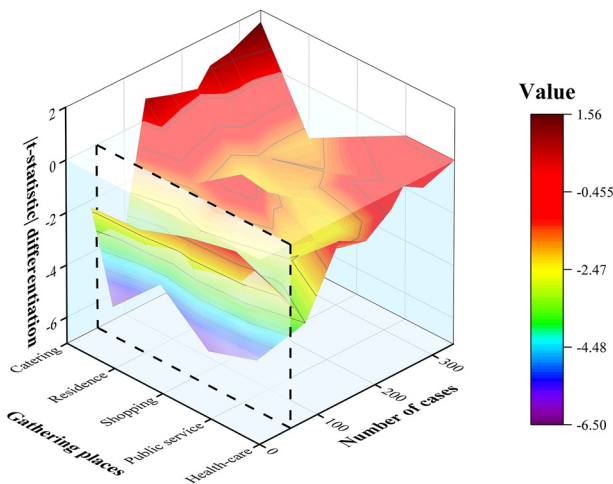
However, to assess the value obtained from modeling, although the “source-case” distance factor assumes a dominant position for the factors associated with places where people gather, close examination of the significance of the impact factors of the six test points, which on the day after the turning point of the epidemic, in the case of the “pure” decay duration, there was a significant downward trend of the “source-case” distance factor and gradually strong intervention by some of the gathering place factors (The catering places, public service facilities and health-care facilities in this study) on the spatial transmission mode of the epidemic (Figure 9). Moreover, Figure 10 illustrates the time-series change of the differentiation between the absolute value of the regression t-statistic value of the “source-case” distance and the gathering place variables. The impact of these gathering places, which are regarded as the “density-driven” factors on COVID-19 in present studies, is narrowing the gap with “distance-driven” factors on an overall level, and the



**Figure 9.** Temporal variations of the significance of the factors in the decay duration of COVID-19: The red line indicates that the variable is in a trend of significant enhancement; the blue line indicates that the variable is in a trend of significant weakening.

significance of some “density-driven” factors has even surpassed the real-time “source-case” distance factor (catering places in the present study). Thus, we may speculate whether this phenomenon is a precursor sign of the change from “distance-driven” to “density-driven” of the localized spatial clustering COVID-19, and this conjecture has yet to be confirmed in follow-up epidemiology studies where there will be many more cases to consider. In effect, the distance decay coefficient, its rate of change, and the degree of the switch from “distance-driven” to “density-driven” of the localized clustering of COVID-19 might serve as a reference tool for examining how well a regional government performs for implementing its prevention and

control strategy for dealing with a localized spatial clustering epidemic. Based on the above-mentioned differences between the localized spatial clustering of COVID-19 and the global large-scale COVID-19, the idea of increasing the distance decay effect and “cut off” the trend from “distance-driven” to “density-driven” of the epidemic as quickly as possible can not only provide a brand-new reference for solving the epidemic at a local level but also distinguish this prevention and control strategy from those adopted in large-scale global epidemics which are quite different. Such an approach should prove beneficial to the classification and scientific prevention and control of different types of epidemics. If the transition or critical point between a “distance-driven” and a “density-driven” epidemic can be identified quantitatively, it will undoubtedly play a very important and enlightening role in the prevention and control of not only the localized spatial clustering of COVID-19 but also the pandemic.



**Figure 10.** Temporal trends of the differentiation of the absolute value of z-score between the gathering places factors (“density-driven” factor) and the “source-case” distance (“distance-driven” factor): The real-time t-statistic of “source-case” distance is set to 0; the polygon of the black dotted box represents the turning point of the COVID-19.

### 4.3. Limitations and Uncertainties

Tracing back to the source, the open-source case data independently counted by the Chinese Center for Disease Control and Prevention and the Beijing Municipal Health Commission in this study is not the most detailed and accurate. The deviation caused by the error caused by approximate processing in such microscopic research is still worth discussing.



Moreover, as mentioned above, we have not discussed modifiable areal unit problem (MAUP) too much in this study, although this problem may cause deviation to the results of spatial statistics. In addition, all the analyses in this study are based on the confirmed data of cases, without considering other attributes of cases (such as hospitalization time and symptom onset time). Combined with the asymmetry of COVID-19 data in time, the analysis results may become relatively deceptive due to the tough prevention and control measures in Beijing. And finally, the scale of the local COVID-19 in Beijing is relatively small, so the results need to be further tested by larger local epidemics in future studies.

## 5. Conclusions

The frequent occurrence of local clusters of COVID-19 in China has highlighted the directions needed for the prevention and control of COVID-19 in the post-pandemic era for the rest of the world. It is, however, still necessary to conduct in-depth research on the epidemic for the spatial distribution characteristics of the pandemic. In this study, the spatiotemporal characteristics of the localized spatial clustering of COVID-19 and the impact and its variation for the “source-case” distance aspect and the places where people gather were explored at the micro-scale. The SL-ZINB and the GWPR-SL, first applied in public health and disease geography, were used and it was found that COVID-19 had a global significant spatial spillover effect at the micro-scale. The “source-case” distance was shown to play a decisive role in the spreading characteristics of COVID-19 and the resultant spatial distribution, which previously in research have been seriously underestimated. Factors associated with places where people gather also played a significant role in modeling the spatial distribution of the epidemic in terms of the time series, but these effects were relatively weak and insignificant concerning the distance factor. Finally, we find that the trend from “distance-driven” to “density-driven” is relatively significant, and the critical point between these two driving modes will be one of the keys to solving the clustering of the local COVID-19, which is also instructive for the prevention and control of the pandemic.

## Conflict of Interest

The authors declare no conflict of interest relevant to this study.

## Data Availability Statement

The functional urban areas data are published by the Beijing City Lab ([www.beijingcitylab.com](http://www.beijingcitylab.com)). COVID-19 case data are reported daily by the Beijing Municipal Health Commission ([wjw.beijing.gov.cn/English](http://wjw.beijing.gov.cn/English)) and the Chinese Center for Disease Control and Prevention ([www.chinacdc.cn](http://www.chinacdc.cn)). The Point of Interest (POI) data are available at POIbase ([www.poiibase.com](http://www.poiibase.com)). This study is categorized as low-risk, as only aggregated, publicly available confirmed cases data were analyzed.

## Acknowledgments

This research was funded by the National Social Science Foundation of China (No. 18BJY086). The authors wish to gratefully acknowledge the anonymous reviewers who helped to improve this paper through their thorough review.

## References

- Anselin, L. (2010). Local indicators of spatial association-LISA. *Geographical Analysis*, 27, 93–115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- Birch, C. P. D., Oom, S. P., & Beecham, J. A. (2007). Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. *Ecological Modelling*, 206, 347–359. <https://doi.org/10.1016/j.ecolmodel.2007.03.041>
- Boots, B. N., & Getis, A. (1988). *Point pattern analysis*. Regional Research Institute, West Virginia University.
- Boyle, P. J., & Flowerdew, R. (1993). Modelling sparse interaction matrices: Interward migration in Hereford and Worcester, and the under-dispersion problem. *Environment & Planning A*, 25(1), 201–209. <https://doi.org/10.1068/a251201>
- Brockmann, D., Hufnagel, L., & Geisel, T. (2006). The scaling laws of human travel. *Nature*, 439, 462–465. <https://doi.org/10.1038/nature04292>
- Brunsdon, C., Fotheringham, A. S., & Charlton, M. (1998). Spatial nonstationarity and autoregressive models. *Environment and Planning a-Economy and Space*, 30, 957–973. <https://doi.org/10.1068/a300957>
- Burger, M., Van Oort, F., & Linders, G.-J. (2009). On the Specification of the Gravity Model of Trade: Zeros, Excess Zeros and Zero-inflated Estimation. *Spatial Economic Analysis*, 4, 167–190. <https://doi.org/10.1080/17421770902834327>
- Byass, P. (2020). Eco-epidemiological assessment of the COVID-19 epidemic in China, January-February 2020. *Global Health Action*, 13(1), 1760490. <https://doi.org/10.1080/16549716.2020.1760490>
- Cai, Q., Lee, J., Eluru, N., & Abdel-Aty, M. (2016). Macro-level pedestrian and bicycle crash analysis: Incorporating spatial spillover effects in dual state count models. *Accident Analysis & Prevention*, 93, 14–22. <https://doi.org/10.1016/j.aap.2016.04.018>

- Chan, J. F.-W., Yuan, S., Kok, K.-H., To, K. K.-W., Chu, H., Yang, J., et al. (2020). A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: A study of a family cluster. *Lancet*, 395, 514–523. [https://doi.org/10.1016/S0140-6736\(20\)30154-9](https://doi.org/10.1016/S0140-6736(20)30154-9)
- Chen, Y., Li, Q., Karimian, H., Chen, X., & Li, X. (2021). Spatio-temporal distribution characteristics and influencing factors of COVID-19 in China. *Scientific Reports*, 11, 3717. <https://doi.org/10.1038/s41598-021-83166-4>
- Clark, A., Jit, M., Warren-Gash, C., Guthrie, B., Wang, H. H. X., Mercer, S. W., et al. (2020). Global, regional, and national estimates of the population at increased risk of severe COVID-19 due to underlying health conditions in 2020: A modelling study. *Lancet Global Health*, 8, 1003–1017. [https://doi.org/10.1016/S2214-109X\(20\)30264-3](https://doi.org/10.1016/S2214-109X(20)30264-3)
- Copiello, S., & Grillenzoni, C. (2020). The spread of 2019-nCoV in China was primarily driven by population density. Comment on "Association between short-term exposure to air pollution and COVID-19 infection: Evidence from China" by Zhu et al. *The Science of the Total Environment*, 744. <https://doi.org/10.1016/j.scitotenv.2020.141028>
- Davies, N. G., P. Klepac, Y. Liu, K. Prem, M. Jit, R. M. Eggo, et al. 2020: Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nature Medicine*, 26, 1205–1211. <https://doi.org/10.1038/s41591-020-0962-9>
- Dong, Y., Mo, X., Hu, Y., Qi, X., Jiang, F., Jiang, Z., & Tong, S. (2020). Epidemiology of COVID-19 among children in China. *Pediatrics*, 145, e20200702. <https://doi.org/10.1542/peds.2020-0702>
- Dungan, J. L., Perry, J. N., Dale, M. R. T., Legendre, P., Citron-Pousty, S., Fortin, M.-J., et al. (2002). A balanced view of scale in spatial statistical analysis. *Ecography*, 25, 626–640. <https://doi.org/10.1034/j.1600-0587.2002.250510.x>
- Elvidge, C. D., Ghosh, T., Hsu, F.-C., Zhizhin, M., & Bazilian, M. (2020). The dimming of lights in China during the COVID-19 pandemic. *Remote Sensing*, 12, 2851. <https://doi.org/10.3390/rs12172851>
- Feuillet, T., Charreire, H., Menai, M., Salze, P., Simon, C., Dugas, J., et al. (2015). Spatial heterogeneity of the relationships between environmental characteristics and active commuting: Towards a locally varying social ecological model. *International Journal of Health Geographics*, 14(1). <https://doi.org/10.1186/s12942-015-0002-z>
- Flowerdew, R., & Aitkin, M. (1982). A method of fitting the gravity model based on the Poisson distribution. *Journal of Regional Science*, 22, 191–202. <https://doi.org/10.1111/j.1467-9787.1982.tb00744.x>
- Flowerdew, R., & Boyle, P. J. (1995). Migration models incorporating interdependence of movers. *Environment & Planning A*, 27(1), 493–502. <https://doi.org/10.1068/a271493>
- Franch-Pardo, I., Napoletano, B. M., Rosete-Verges, F., & Billa, L. (2020). Spatial analysis and GIS in the study of COVID-19. A review, *The Science of the Total Environment*, 739. <https://doi.org/10.1016/j.scitotenv.2020.140033>
- Gatrell, A. C., & Bailey, T. C. (1996). Interactive spatial data analysis in medical geography. *Social Science & Medicine* (1982), 42, 843–855. [https://doi.org/10.1016/0277-9536\(95\)00183-2](https://doi.org/10.1016/0277-9536(95)00183-2)
- Geniaux, G., & Martinetti, D. (2018). A new method for dealing simultaneously with spatial autocorrelation and spatial heterogeneity in regression models. *Regional Science and Urban Economics*, 72, 74–85. <https://doi.org/10.1016/j.regsciurbeco.2017.04.001>
- Getis, A. (1984). Interaction modeling using second-order analysis. *Environment and Planning A: Economy and Space*, 16, 173–183. <https://doi.org/10.1068/a160173>
- Getis, A., & Aldstadt, J. (2004). Constructing the spatial weights matrix using a local statistic. *Geographical Analysis*, 36, 90–104. <https://doi.org/10.1111/j.1538-4632.2004.tb01127.x>
- Getis, A., & Ord, J. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24, 189–206.
- Griffith, D. A. (1987a). *Spatial autocorrelation: A primer resource publications in geography*. Association of American Geographers.
- Griffith, D. A. (1987b). *Spatial autocorrelation. A primer*. Association of American Geographers.
- Grushka, E. (1972). Characterization of exponentially modified Gaussian peaks in chromatography. *Analytical Chemistry*, 44, 1733–1738. <https://doi.org/10.1021/ac60319a011>
- Hamidi, S., & Zandiatashbar, A. (2021). Compact development and adherence to stay-at-home order during the COVID-19 pandemic: A longitudinal investigation in the United States. *Landscape and Urban Planning*, 205, 103952. <https://doi.org/10.1016/j.landurbplan.2020.103952>
- Han, Y., Yang, L., Jia, K., Li, J., Feng, S., Chen, W., et al. (2021). Spatial distribution characteristics of the COVID-19 pandemic in Beijing and its relationship with environmental factors. *The Science of the Total Environment*, 761, 144257. <https://doi.org/10.1016/j.scitotenv.2020.144257>
- Ji, T., Chen, H.-L., Xu, J., Wu, L.-N., Li, J.-J., Chen, K., & Qin, G. (2020). Lockdown contained the spread of 2019 novel coronavirus disease in Huangshi City, China: Early epidemiological findings. *Clinical Infectious Diseases*, 71, 1454–1460. <https://doi.org/10.1093/cid/ciaa390>
- Kadi, N., & Khelifaoui, M. (2020). Population density, a factor in the spread of COVID-19 in Algeria: Statistic study. *Bulletin of the National Research Centre*, 44, 138. <https://doi.org/10.1186/s42269-020-00393-x>
- Kraemer, M. U. G., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D. M., et al. (2020). The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*, 368, 493, 497. <https://doi.org/10.1126/science.abb4218>
- Kucharski, A. J., Klepac, P., Conlan, A. J. K., Kissler, S. M., Tang, M. L., Fry, H., et al. (2020). Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: A mathematical modelling study. *The Lancet Infectious Diseases*, 20, 1151–1160. [https://doi.org/10.1016/S1473-3099\(20\)30457-6](https://doi.org/10.1016/S1473-3099(20)30457-6)
- Kwan, M.-P. (2012). The uncertain geographic context problem. *Annals of the Association of American Geographers*, 102, 958–968. <https://doi.org/10.1080/00045608.2012.687349>
- Kwan, M.-P. (2018). The limits of the neighborhood effect: Contextual uncertainties in geographic, environmental health, and social science research. *Annals of the American Association of Geographers*, 108, 1482–1490. <https://doi.org/10.1080/24694452.2018.1453777>
- Kwan, M.-P. (2021). The stationarity bias in research on the environmental determinants of health. *Health & Place*, 70, 102609. <https://doi.org/10.1016/j.healthplace.2021.102609>
- Lambert, D. (1992). Zero-inflated Poisson Regression, with an application to defects in manufacturing. *Technometrics*, 34, 1–14. <https://doi.org/10.2307/1269547>
- Lau, H., Khosrawipour, V., Kocbach, P., Mikolajczyk, A., Schubert, J., Bania, J., & Khosrawipour, T. (2020). The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China. *Journal of Travel Medicine*, 27, taaa037. <https://doi.org/10.1093/jtm/taaa037>
- Lefever, D. W. (1926). Measuring geographic concentration by means of the standard deviational ellipse. *American Journal of Sociology*, 32, 88–94. <https://doi.org/10.1086/214027>
- Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., et al. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *New England Journal of Medicine*, 382, 1199–1207. <https://doi.org/10.1056/nejmoa2001316>
- Li, Z., & Gao, G. F. (2020). Strengthening public health at the community-level in China. *Lancet Public Health*, 5, E629–E630. [https://doi.org/10.1016/S2468-2667\(20\)30266-8](https://doi.org/10.1016/S2468-2667(20)30266-8)

- Ling, C., & Wen, X. (2020). Community grid management is an important measure to contain the spread of novel coronavirus pneumonia (COVID-19). *Epidemiology and Infection*, *148*, e167. <https://doi.org/10.1017/s0950268820001739>
- Liu, P., Yang, M., Zhao, X., Guo, Y., Wang, L., Zhang, J., et al. (2020a). Cold-chain transportation in the frozen food industry may have caused a recurrence of COVID-19 cases in destination: Successful isolation of SARS-CoV-2 virus from the imported frozen cod package surface. *Biosafety and Health*, *2*, 199–201. <https://doi.org/10.1016/j.bsheal.2020.11.003>
- Liu, S., Qin, Y., Xie, Z., & Zhang, J. (2020b). The spatio-temporal characteristics and influencing factors of covid-19 spread in Shenzhen, China-An analysis based on 417 cases. *International Journal of Environmental Research and Public Health*, *17*, 7450. <https://doi.org/10.3390/ijerph17207450>
- Luo, W., & Qi, Y. (2009). An enhanced two-step floating catchment area (E2SFCA) method for measuring spatial accessibility to primary care physicians. *Health & Place*, *15*, 1100–1107. <https://doi.org/10.1016/j.healthplace.2009.06.002>
- Ma, S., & Long, Y. (2020). Functional urban area delineations of cities on the Chinese mainland using massive Didi ride-hailing records. *Cities*, *97*, 102532. <https://doi.org/10.1016/j.cities.2019.102532>
- Mahato, S., Pal, S., & Ghosh, K. G. (2020). Effect of lockdown amid COVID-19 pandemic on air quality of the megacity Delhi, India. *The Science of the Total Environment*, *730*, 139086. <https://doi.org/10.1016/j.scitotenv.2020.139086>
- Martins, L. D., da Silva, I., Batista, W. V., Andrade, M. d. F., de Freitas, E. D., & Martins, J. A. (2020). How socio-economic and atmospheric variables impact COVID-19 and influenza outbreaks in tropical and subtropical regions of Brazil. *Environmental Research*, *191*, 110184. <https://doi.org/10.1016/j.envres.2020.110184>
- McGrail, M. R., & Humphreys, J. S. (2009). Measuring spatial accessibility to primary care in rural areas: Improving the effectiveness of the two-step floating catchment area method. *Applied Geography*, *29*, 533–541. <https://doi.org/10.1016/j.apgeog.2008.12.003>
- Metulini, R., Patuelli, R., & Griffith, D. (2018). A Spatial-Filtering Zero-Inflated Approach to the Estimation of the Gravity Model of Trade. *Econometrics*, *6*, 9. <https://doi.org/10.3390/econometrics6010009>
- Middela, M. S., & Ramadurai, G. (2020). Incorporating spatial interactions in zero-inflated negative binomial models for freight trip generation. *Transportation*. <https://doi.org/10.1007/s11116-020-10132-w>
- Mollalo, A., Vahedi, B., & Rivera, K. M. (2020). GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *The Science of the Total Environment*, *728*, 138884. <https://doi.org/10.1016/j.scitotenv.2020.138884>
- Nakaya, T., Fotheringham, A. S., Brunson, C., & Charlton, M. (2005). Geographically weighted Poisson regression for disease association mapping. *Statistics in Medicine*, *24*, 2695–2717. <https://doi.org/10.1002/sim.2129>
- Nekola, J. C., & White, P. S. (1999). The distance decay of similarity in biogeography and ecology. *Journal of Biogeography*, *26*, 867–878. <https://doi.org/10.1046/j.1365-2699.1999.00305.x>
- Organization, W. H. (2021). *COVID-19 Weekly Update - 9 March 2021*.
- Pang, X., Ren, L., Wu, S., Ma, W., Yang, J., Di, L., et al. (2020). Cold-chain food contamination as the possible origin of COVID-19 resurgence in Beijing. *National Science Review*, *7*, 1861–1864. <https://doi.org/10.1093/nsr/nwaa264>
- Rashed, E. A., Kodera, S., Gomez-Tames, J., & Hirata, A. (2020). Influence of absolute humidity, temperature and population density on COVID-19 spread and decay durations: Multi-prefecture study in Japan. *International Journal of Environmental Research and Public Health*, *17*, 5354. <https://doi.org/10.3390/ijerph17155354>
- Sannigrahi, S., Pilla, F., Basu, B., Basu, A. S., & Molter, A. (2020). Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach. *Sustainable Cities and Society*, *62*, 102418. <https://doi.org/10.1016/j.scs.2020.102418>
- Sarkodie, S. A., & Owusu, P. A. (2020). Global assessment of environment, health and economic impact of the novel coronavirus (COVID-19). *Environment, Development and Sustainability*, *23*, 5005–5015. <https://doi.org/10.1007/s10668-020-00801-2>
- Shoff, C., Chen, V. Y.-J., & Yang, T.-C. (2014). When homogeneity meets heterogeneity: The geographically weighted regression with spatial lag approach to prenatal care utilisation. *Geospatial Health*, *8*, 557–568. <https://doi.org/10.4081/gh.2014.45>
- Sohrabi, C., Alsafi, Z., O'Neill, N., Khan, M., Kerwan, A., Al-Jabir, A., et al. (2020). World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). *International Journal of Surgery*, *76*, 71–76. <https://doi.org/10.1016/j.ijssu.2020.02.034>
- Stewart, J. Q. (1948). Demographic gravitation: Evidence and applications. *Sociometry*, *11*, 31–58. <https://doi.org/10.2307/2785468>
- Sun, Z., Zhang, H., Yang, Y., Wan, H., & Wang, Y. (2020). Impacts of geographic factors and population density on the COVID-19 spreading under the lockdown policies of China. *The Science of the Total Environment*, *746*, 141347. <https://doi.org/10.1016/j.scitotenv.2020.141347>
- Tian, H., Liu, Y., Li, Y., Wu, C.-H., Chen, B., Kraemer, M. U. G., et al. (2020). An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science*, *368*, 638–642. <https://doi.org/10.1126/science.abb6105>
- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, *46*, 234–240. <https://doi.org/10.2307/143141>
- Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, *57*, 307. <https://doi.org/10.2307/1912557>
- Wang, D., Hu, B., Hu, C., Zhu, F., Liu, X., Zhang, J., et al. (2020a). Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *Jama-Journal of the American Medical Association*, *323*, 1061–1069. <https://doi.org/10.1001/jama.2020.1585>
- Wang, K., Lu, Z., au, fnm, Wang, X., Li, H., Li, H., et al. (2020b). Current trends and future prediction of novel coronavirus disease (COVID-19) epidemic in China: A dynamical modeling analysis. *Mathematical Biosciences and Engineering*, *17*, 3052–3061. <https://doi.org/10.3934/mbe.2020173>
- Wang, X.-Y., Zhang, Y.-Q., & Cai, L.-W. (2021). Spatiotemporal characteristics of the COVID-19 resurgence in the metropolitan wholesale market of Beijing, China. *Journal of Travel Medicine*, *28*, taab008. <https://doi.org/10.1093/jtm/taab008>
- Wang, Y., & Di, Q. (2020). Modifiable areal unit problem and environmental factors of COVID-19 outbreak. *The Science of the Total Environment*, *740*, 139984. <https://doi.org/10.1016/j.scitotenv.2020.139984>
- Wiegand, T., & Moloney, K. A. (2004). Rings, circles, and null-models for point pattern analysis in ecology. *Oikos*, *104*, 209–229. <https://doi.org/10.1111/j.0030-1299.2004.12497.x>
- Wilder-Smith, A., & Freedman, D. O. (2020). Isolation, quarantine, social distancing and community containment: Pivotal role for old-style public health measures in the novel coronavirus (2019-nCoV) outbreak. *Journal of Travel Medicine*, *27*, taaa020. <https://doi.org/10.1093/jtm/taaa020>
- William, H. G. (1994). Accounting for excess zeros and sample selection in Poisson and negative binomial regression models.
- Wilson, P. (2015). The misuse of the Vuong test for non-nested models to test for zero-inflation. *Economics Letters*, *127*, 51–53. <https://doi.org/10.1016/j.econlet.2014.12.029>

- Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Song, Z.-G., et al. (2020a). A new coronavirus associated with human respiratory disease in China. *Nature*, 579, 265, 269. <https://doi.org/10.1038/s41586-020-2008-3>
- Wu, Y., Yan, X., Zhao, S., Wang, J., Ran, J., Dong, D., et al. (2020b). Association of time to diagnosis with socioeconomic position and geographical accessibility to healthcare among symptomatic COVID-19 patients: A retrospective study in Hong Kong. *Health & Place*, 66, 102465. <https://doi.org/10.1016/j.healthplace.2020.102465>
- Wu, Z., Wang, Q., Zhao, J., Yang, P., McGoogan, J. M., Feng, Z., & Huang, C. (2020c). Time course of a second outbreak of COVID-19 in Beijing, China, June–July 2020. *Jama—Journal of the American Medical Association*, 324, 1458–1459. <https://doi.org/10.1001/jama.2020.15894>
- Yu, X., Wong, M. S., Kwan, M. P., Nichol, J. E., Zhu, R., Heo, J., et al. (2021). COVID-19 infection and mortality: Association with PM2.5 concentration and population density—an exploratory study. *ISPRS International Journal of Geo-Information*, 10. <https://doi.org/10.3390/ijgi10030123>
- Zhang, J., Litvinova, M., Liang, Y., Wang, Y., Wang, W., Zhao, S., et al. (2020a). Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science*, 368, 1481–1486. <https://doi.org/10.1126/science.abb8001>
- Zhang, Y., Pan, Y., Zhao, X., Shi, W., Chen, Z., Zhang, S., et al. (2020b). Genomic characterization of SARS-CoV-2 identified in a reemerging COVID-19 outbreak in Beijing's Xinfadi market in 2020. *Biosafety and Health*, 2, 202–205. <https://doi.org/10.1016/j.bsheal.2020.08.006>
- Zhao, C., Liu, Z., & Ding, Y. (2020). How COVID-induced Uncertainty Influences Chinese Firms' OFDI Binary Margins. *Emerging Markets Finance and Trade*, 56, 3613–3625. <https://doi.org/10.1080/1540496x.2020.1855139>
- Zhou, C., Tao, P., Yunyan, D. U., Jun, X., Jiaoe, W., Guoyi, Z., et al. (2020). Bigdata analysis on COVID-19 epidemic and suggestions on regional prevention and control policy. *Bulletin of Chinese Academy of Sciences*, 35, 200–203. <https://doi.org/10.16418/j.issn.1000-3045.20200209001>