Exploring Urban Spatial Features of COVID-19 Transmission in Wuhan Based on Social Media Data

Zhenghong Peng [1], **Ru Wang** [1], **Lingbo Liu** [2] and **Hao Wu** [1,*]

1   Department of Graphics and Digital Technology, School of Urban Design, Wuhan University,
    Wuhan 430072, China; pengzhenghong@whu.edu.cn (Z.P.); wang_ru@whu.edu.cn (R.W.)
2   Department of Urban Planning, School of Urban Design, Wuhan University, Wuhan 430072, China;
    lingbo.liu@whu.edu.cn
*   Correspondence: wh79@whu.edu.cn; Tel.: +86-27-6877-3062

Abstract: During the early stage of the COVID-19 outbreak in Wuhan, there was a short run of medical resources, and Sina Weibo, a social media platform in China, built a channel for novel coronavirus pneumonia patients to seek help. Based on the geo-tagging Sina Weibo data from February 3rd to 12th, 2020, this paper analyzes the spatiotemporal distribution of COVID-19 cases in the main urban area of Wuhan and explores the urban spatial features of COVID-19 transmission in Wuhan. The results show that the elderly population accounts for more than half of the total number of Weibo help seekers, and a close correlation between them has also been found in terms of spatial distribution features, which confirms that the elderly population is the group of high-risk and high-prevalence in the COVID-19 outbreak, needing more attention of public health and epidemic prevention policies. On the other hand, the early transmission of COVID-19 in Wuhan could be divide into three phrases: Scattered infection, community spread, and full-scale outbreak. This paper can help to understand the spatial transmission of COVID-19 in Wuhan, so as to propose an effective public health preventive strategy for urban space optimization.

Keywords: COVID-19; social media data; sina weibo; spatiotemporal characteristics

---

1. Introduction

In January 2020, the novel coronavirus pneumonia, COVID-19, broke out in Wuhan, the capital of Hubei province in China and the development of the epidemic has been a rising worldwide concern [1]. In order to slow down and block the spread of the virus, Wuhan announced a shutdown on January 23rd, 2020 to suspend the city's public transportation and imposed an unprecedented restriction on personal mobility. The effect of contact limitation gradually appeared after one month, which proves that for cities with high population density, severe restrictions on population movement can play a positive role in suppressing the spread of infectious diseases [2].

The existing researches include the spatial-temporal dynamics study of COVID-19 on the country level [3–5], with the deficiency of the detail reveal of the characteristics in the early stage in the urban space, which is critical for the development of prevention and control work in the city during the epidemic. The outbreak of COVID-19 verifies that epidemic prevention in public health is an essential part of urban planning and governance [6,7]. It is important to understand the heterogeneity of urban space in terms of social space and geographic space in the city [8], to identify susceptible people [9,10] and disease-prone spaces, and to grasp the dynamic spread of infectious diseases in urban spaces [11]. Thus, preventive control nodes can deploy corresponding measures for specific groups and spaces [12,13].

With the rapid development of information and communications technology (ICT), big data has been widely used in the field of public health [14,15] in recent years. Social media data such as Twitter and Weibo data were used to study the public attention [16,17], to predict epidemic outbreak [18–20], and to make research on human sentiments [21,22] during the COVID-19 epidemic. Compared with traditional cases data, there are unique advantages such as the wider population sample coverage, the easier accessibility, and the accurate geographic information. Such big data in epidemic studies can improve the understanding of the epidemic in time and space and will play an important role in formulating targeted urban prevention and control strategies.

In the early days of the lockdown, Wuhan, as the epidemic center, faced great challenges. The panic caused by COVID-19 outbreak stimulated a large number of people to enter the hospital, leading to a short-term collapse of medical system. Such a squeeze on medical resources also prevented many diagnosed and suspected people of having COVID-19 from receiving timely treatment. More than 1000 families in Wuhan posted help information on Sina Weibo, a Chinese social media platform, seeking immediate medical treatments. Most information was posted from February fourth to eighth, and then decreased rapidly, when the policy that "Guarantee that all suspected and confirmed cases should be collected and cured" was released on February fifth and a series of measures were taken by the government to quickly supplement medical resources, including increase the number of medical beds and medical staff supporting Wuhan (Figure 1).
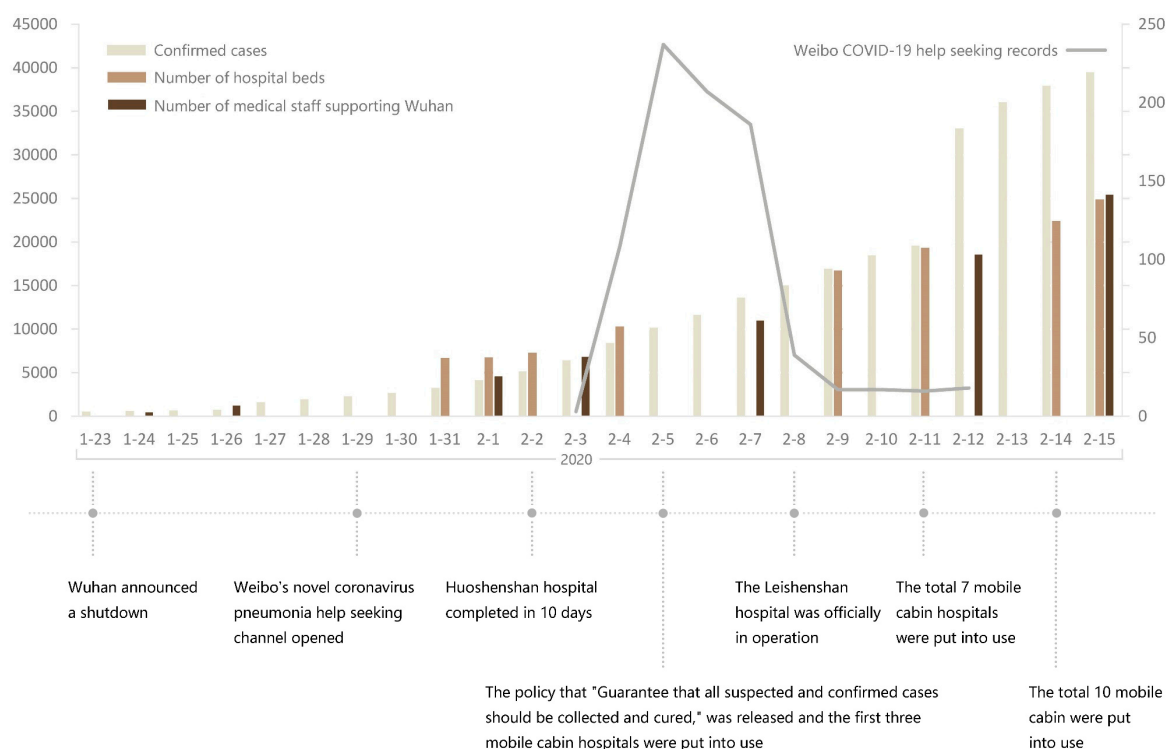
Figure 1. Time-series change histogram in the early stage of the epidemic.

The geotagged Weibo data can reflect the spatial distribution of COVID-19 infectors and provided an accessible sample data reflecting the spatiotemporal characteristics of the epidemic development. On the other hand, mobile phone data with age tag may help to characterize the spatial distribution of the elderly population in the study area. This article aims to analyze the spatial distribution of COVID-19 transmission based on geographic information system (GIS) visualization [23,24] with Weibo COVID-19 help-seekers data from February 3rd to February 12th, to explore the correlation between the spatial characteristics of the epidemic distribution and aged population, and to analyze the spatiotemporal features of epidemic spread. The analysis of the detailed characteristics of the epidemic in urban space can help cities that may have or have had outbreak to better understand the

mechanism of disease transmission and it is of great significance to improve citizens' awareness of protection and provide certain policy references for government departments.

2. Study Area

Wuhan, as the capital of Hubei Province, (Figure 2a) was the earliest COVID-19 outbreak area in China. There was a short medical run period between the lockdown on January 23rd, 2020 and the increase in the number of medical beds and medical staff supporting Wuhan, leading to 99% of Weibo help seeking patients being in Wuhan (Figure 3). Wuhan is located in the east of Jianghan Plain and the middle reaches of the Yangtze River. The Yangtze River and its largest tributary, the Han River, run across the center of the city, dividing the central urban area of Wuhan into three regions of Wuchang, Hankou, and Hanyang, standing across the river (Figure 2b). Wuhan's main urban area (MUA) is the main concentration of urban function area, overlapped with the administration boundary of sub-districts.
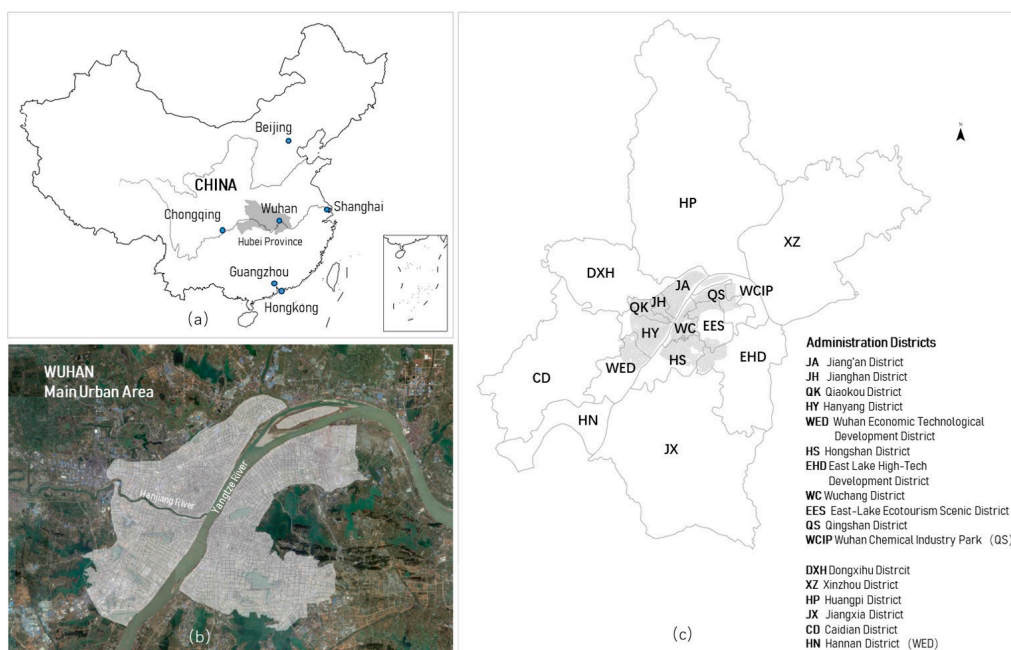
Figure 2. Map of the study area in Wuhan, China: (**a**) the geographic location of the Wuhan, China; (**b**) the main urban area of Wuhan (MUA); and (**c**) administrative districts of Wuhan.
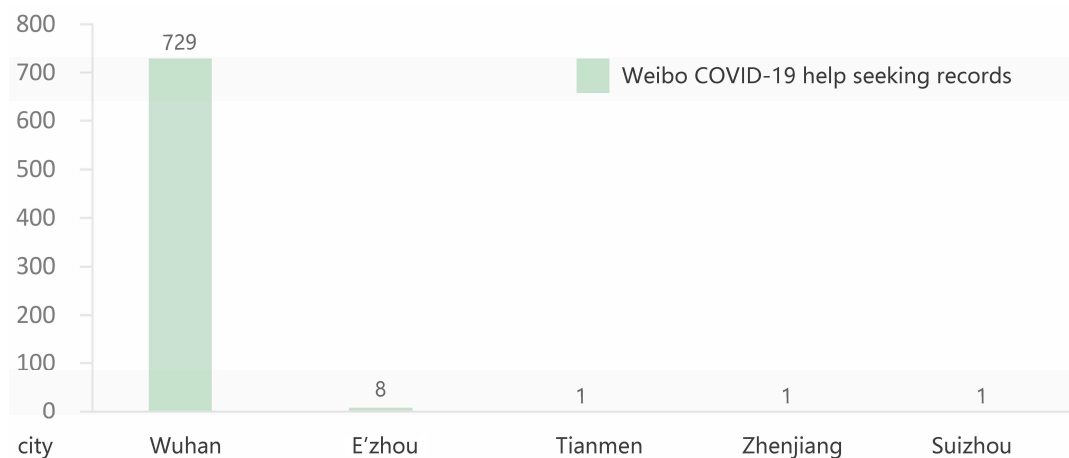
Figure 3. The distribution of Weibo COVID-19 help seeking records.

Wuhan contains seventeen sub-districts, which concludes Jiang'an (JA), Jianghan (JH), Qiaokou (QK), Hanyang (HY), Wuhan Economic Technological Development District (WED), Hongshan (HS), East Lake High-Tech Development District (EHD), Wuchang (WC), East-Lake Ecotourism Scenic District (EES), Qingshan (QS), Wuhan Chemical Industry Park (WCIP), Dongxihu (DXH), Xinzhou (XZ), Huangpi (HP), Jiangxia (JX), Caidian (CD), and Hannan (HN) (Figure 2c). Wherein, WCIP and HN usually statistically belong to QS and WED, respectively.

According to the epidemic data of June 11th, 2020, about 76.327% of the total cumulative cases in Wuhan were in the MUA (Table 1), which is also the relatively concentrated area for early Weibo help seekers.

Table 1. Case statistics of Wuhan. (As of June 11th, 2020 source: Wuhan Municipal Commission of Health).

Districts	Cumulative Confirmed Case	Proportion
Jiang'an (MUA)	6563	13.037%
Jianghan(MUA)	5242	10.413%
Qiaokou(MUA)	6854	13.616%
Hanyang(MUA)	4691	9.319%
Wuchang(MUA)	7551	15.000%
Qingshan(MUA)	2804	5.570%
Hongshan(MUA)	4718	9.372%
Dongxihu	2478	4.923%
Caidian	1424	2.829%
Jiangxia	860	1.708%
Huangpi	2117	4.205%
Xinzhou	1071	2.128%
East Lake Ecotourism Scenic District	483	0.959%
East Lake High-Tech Development District	2173	4.317%
Wuhan Economic Technological Development District	1088	2.161%
Other places	223	0.443%
Total	50340	100.00%

3. Materials and Methods

3.1. Data and Preprocessing

3.1.1. Weibo Data

Sina Weibo is one of the most influential social media platforms in China. It had 486 million monthly active users by the end of June, 2019 [25]. In the early stage of the COVD-19 epidemic, due to the short-term collapse of medical system, Weibo opened the novel coronavirus pneumonia help seeking channel [26] to help patients who could not get timely treatment. The help seeking records were mainly from February 3rd to February 12th, 2020. The study collected about 1200 Weibo records under the topic of "novel coronavirus pneumonia help-seeking" and considered valid information including name, age, home address, time of illness, and number of people infected (Figure 4). As such, 740 records of valid information were finally obtained after data cleaning, wherein 729 records were in Wuhan (Figure 3).
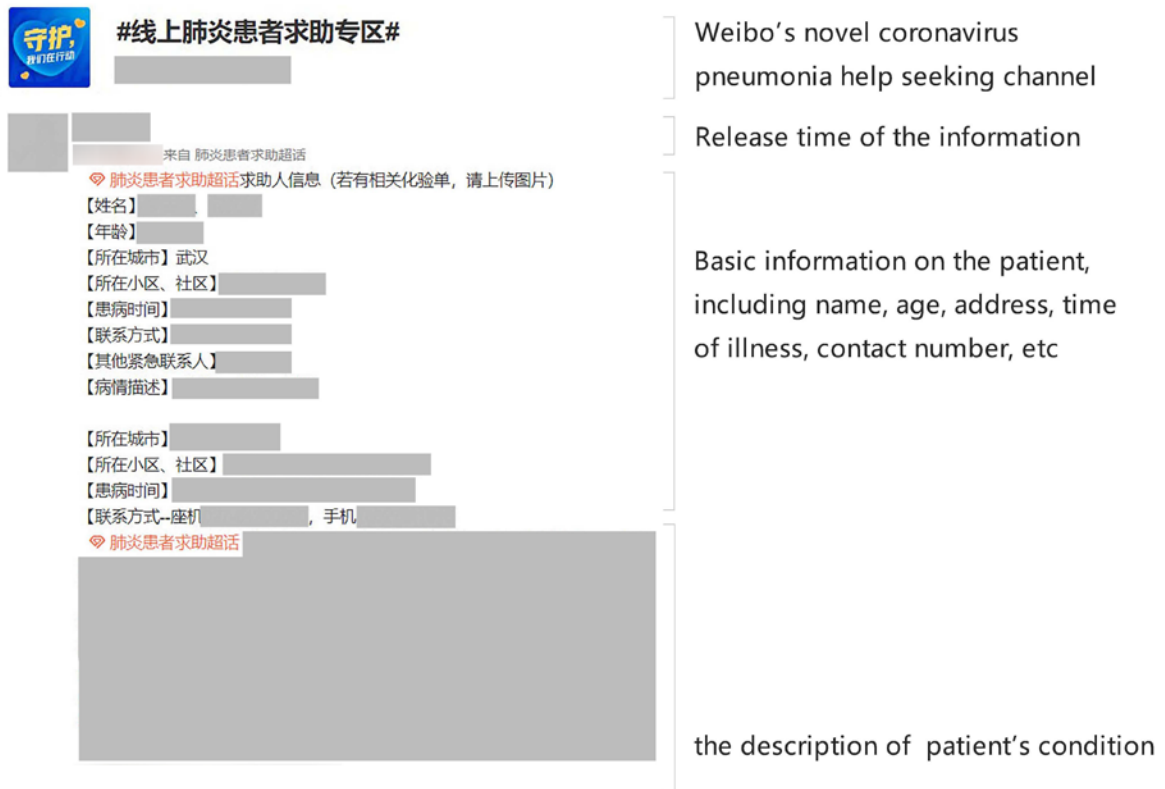
Figure 4. An example of Weibo help-seeking information.

The spatial distribution of help seekers is obtained by geocoding (Figure 5). The data shows that a large number of help seekers were concentrated in the main urban area of Wuhan, and a small number of records were outside.
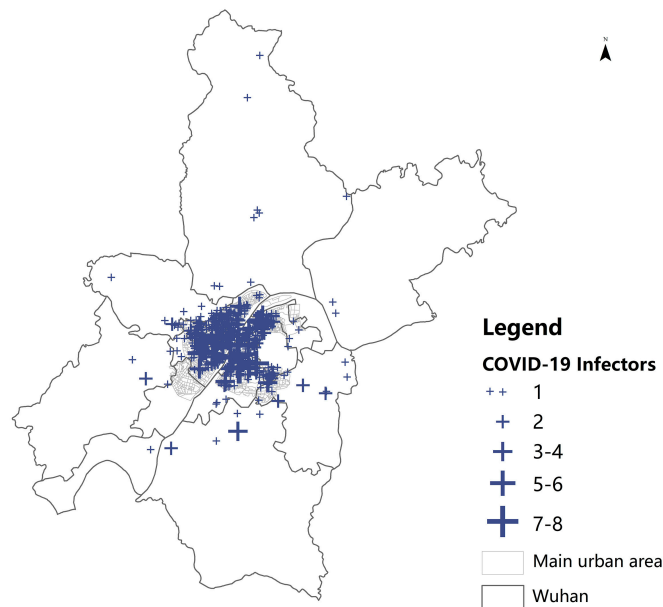
Figure 5. Spatial distribution of Wuhan Weibo help-seeking data.

3.1.2. Mobile Phone Data

Wuhan city's March 2017 call detail record (CDR) data with age tags was used in this study and the spatial distribution of base stations is shown in Figure 6. The steps were as follows to obtain the spatial distribution of the mobile phone population and the elderly population.

- We matched the mobile phone number and the user ID to eliminate all private information, and then removed the invalid and noise data.
- We counted the base stations with the highest call frequency of users, matching the base station code and the user ID, and summarized the number of users that the base station served.
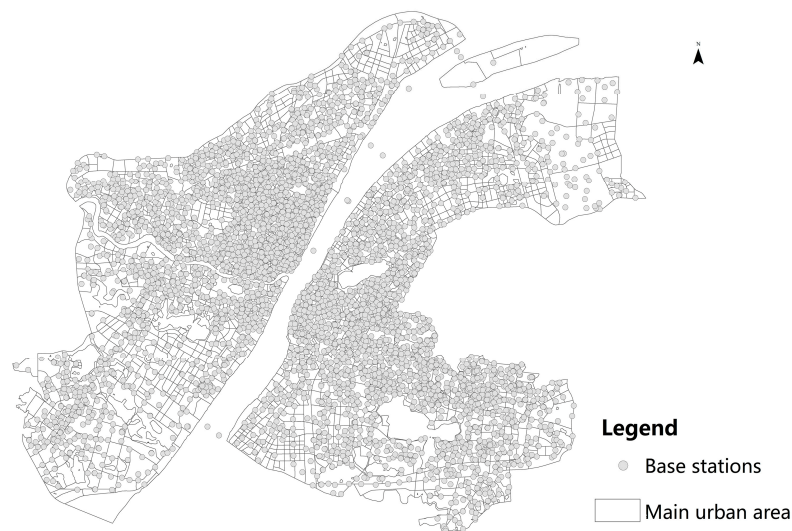
Figure 6. Spatial distribution of base stations in the main urban area of Wuhan.

3.2. Research Methods

3.2.1. Kernel Density Analysis

The Kernel density method was applied since the Weibo data could be seen as a sample data from total COVID-19 infectors. Kernel density analysis is capable to calculate the unit density of the measured values of points and line elements within a specified neighborhood, intuitively reflecting the distribution of discrete measured values in the continuous area. The result is the smooth surface with a large median value and a small peripheral value. The grid value is the unit density, which is reduced to 0 at the boundary of the neighborhood. Kernel density analysis can be used for service facility accessibility [27], crime prediction [28], business analysis, etc. Its function expression is as follows:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \tag{1}$$

where K is the kernel (a non-negative function), $h > 0$ is a smoothing parameter called the bandwidth, x_i is the sample point.

3.2.2. Ordinary Least Square Regression

Based on Kernel density method, the average value of each space unit is calculated. The regression models, using the interpolation of infected people from Weibo data, the interpolation of population, and the elderly population generated by mobile phone data in community units in the main urban area were then constructed. The covariant explanatory variables were checked by the variance inflation

factor (VIF) parameter, and the explanatory variables passing the P value of 1% significance level test were obtained. The function expression are as follows:

$$Y = \beta_m + \beta_1 x_1 + \varepsilon_1 \tag{2}$$

$$Y = \beta_n + \beta_2 x_2 + \varepsilon_2 \tag{3}$$

where Y is the interpolation of Weibo infected people in the community unit, x_1 is the interpolation of population derived from mobile phone data in the community unit, x_2 is the interpolation of the elderly population derived from mobile phone data in the community unit, β_m and β_n are intercepts, β_1 and β_2 are regression coefficients of factors, and ε_1 and ε_2 are random errors.

4. Results

4.1. Preliminary Analysis on COVID-19 Cases of Weibo Data

4.1.1. Demographic Statistics

There are 691 records containing age tags, wherein the maximum age was 95 and the minimum was 11-months. Among them, the proportion of infectors between 30 and 69 was 69.47%, which was lower than the corresponding proportion of 77.2% in the whole sample of Wuhan till February 11th, 2020 [29], and the proportion of infectors over 60 was 57.02%, which was higher than the corresponding proportion of 44.1% in the whole sample of Wuhan, showing that the susceptible population of novel coronavirus was mainly the middle-aged and the elderly groups. Moreover, the elderly who suffer from basic diseases are more likely to develop into critical patients, making the elderly patients become the group that accounted for more than half of Weibo help seekers (Figure 7).
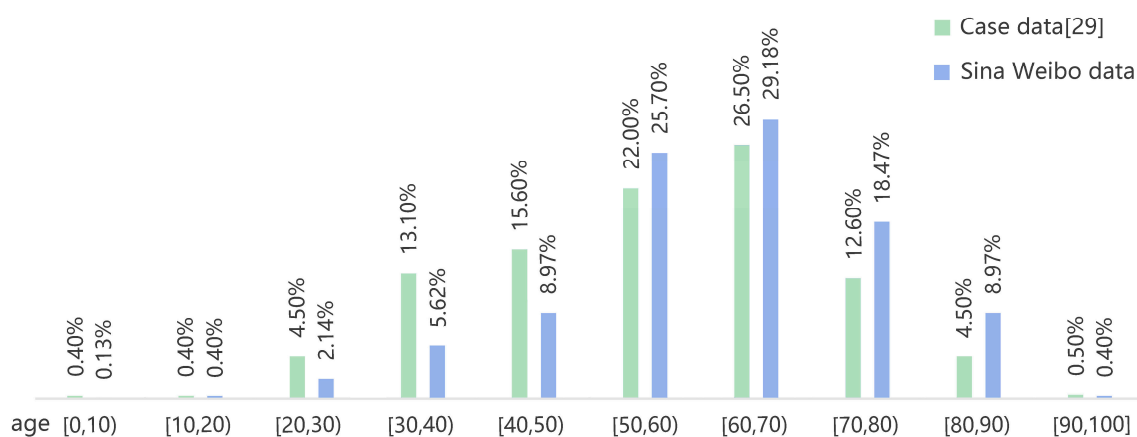
Figure 7. Histogram of patient age distribution of case data and Weibo data.

4.1.2. Spatial Distribution

According to the statistics of the number of household infectors reported in each single record, most records reported one to two infections, with the highest number of eight, reflecting the severity of clustered infections in families (Figure 8). The spatial distribution of COVID-19 cases of Weibo data in MUA showed relatively concentrated regional patterns in Hankou, northern Hanyang, and Wuchang (Figure 9).
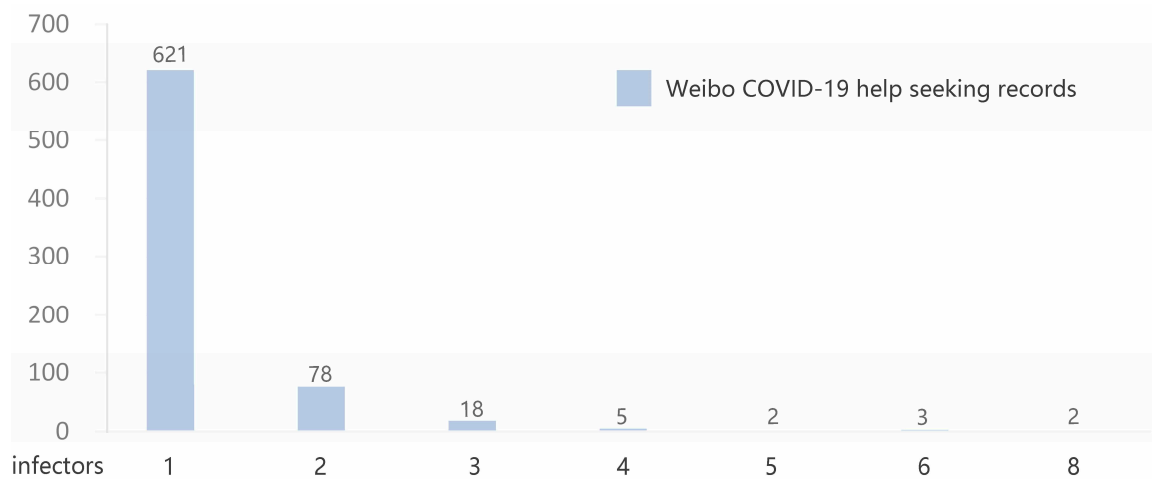
Figure 8. The number of infectors reported in each of the Weibo records.
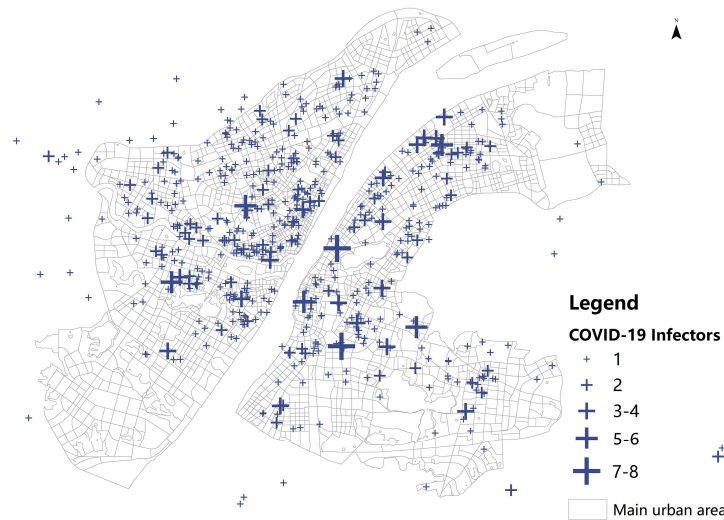
Figure 9. Spatial distribution of Weibo help seekers in the main urban area of Wuhan.

4.1.3. Time Series Statistics

Based on the number of COVID-19 infection of Weibo records and total infector reported, the corresponding change curve by onset time from December 20th, 2019 to February 10th, 2020 was obtained (Figure 10). It can be clearly seen that as time goes on, the absolute value of the difference between the two increased gradually, reaching the maximum when the epidemic was in full outbreak, around January 23rd, 2020, and then a general downward trend. It reflected that during the period of rapid growth of the infectors from December 22nd, 2020, the isolation control measures such as traffic restrictions and the improvement of medical facilities reduced the number of household infections effectively.
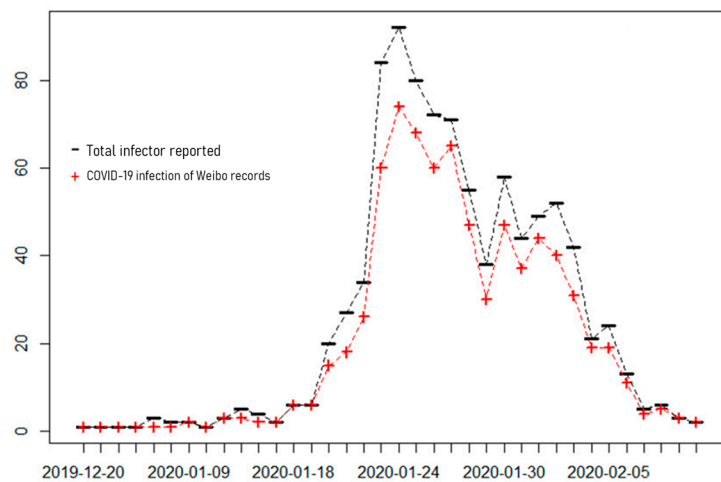
Figure 10. Timing chart of COVID-19 infection of Weibo records and total infector reported

4.2. Spatial Correlation Between COVID-19 Cases and Population Density

4.2.1. Kernel Density Analysis

The spatial features of COVID-19 cases and population density were visualized and compared, with the weight of the number of infections, population, and the elderly population, respectively, according to the Kernel density method. As is shown in Figure 11, COVID-19 cases of Weibo data were mainly concentrated in the central area along the Yangtze river, which is similar to the spatial distribution of population density, especially of the elderly population (Figure 11).
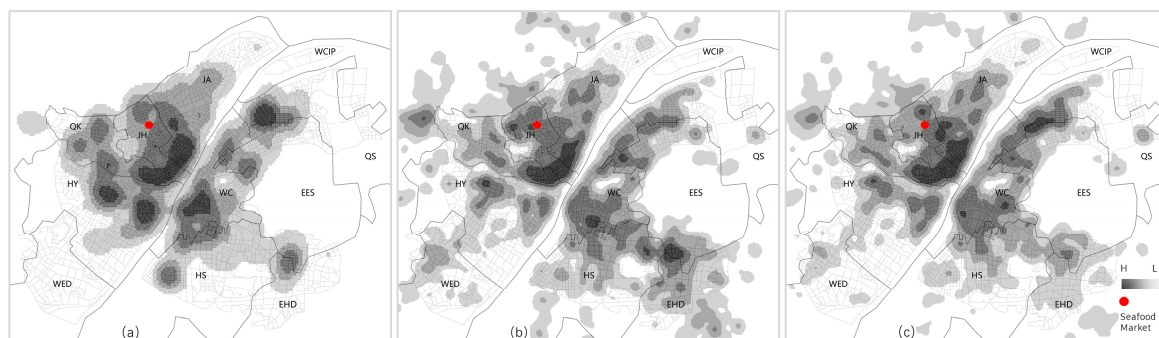
Figure 11. Kernel density analysis: (**a**) COVID-19 cases by Weibo data; (**b**) population density; and (**c**) the elderly population density generated by mobile phone data.

Interestingly, the Huanan Seafood Market, regarded as the origin point of the outbreak, had not been the geographic center of epidemic outbreak. This may be explained by the medium population density and low density of elderly population.

4.2.2. Ordinary Least Squares (OLS) Regression

According to the urban arterial roads, traffic analysis zone (TAZ) [30], the spatial statistical unit for urban population commuting characteristics analysis, was divided by the traffic management department. Since there was collinearity between the population and the elderly population in the study, the OLS model was established based on the TAZ level, using the interpolation of help seekers from Weibo data, with the interpolation of the population and the elderly population generated by mobile phone data, respectively. The results further verify that the distribution of the elderly population was significantly related to the infectors, with higher adjusted R-Squared of 0.7038 (Table 2).

Table 2. OLS diagnostics.

Variable	Coefficient	Stand Error	*p*-Value	Intercept	R-Squared	Adjusted R-Squared
The elderly	39.8216	0.5408	0.0000	7212.6697	0.7039	0.7038
population	3.2416	0.0576	0.0000	7265.7562	0.5811	0.5809

4.3. Spatiotemporal Features of COVID-19 Transmission

This paper further explored the early spatiotemporal characteristics of the epidemic transmission based on the infection time of COVID-19 cases provided by Weibo data. Wherein, the hot spots can be regarded as the initial pathogen transmission, and the areas with the highest density levels in each period can be regarded as the areas with the fastest transmission of infection. The COVID-19 transmission map of Weibo data shows a clear process of three stages: Scattered infection, community spread, and full-scale outbreak.

In the COVID-19 data of Weibo, only 3 infectors had been reported before 2020, and 25 infectors were from January 1st to 18th, 2020. The earliest infected spots already covered all the outbreak areas except hotspots in Hongshan district (Figure 12a), which began to appear in the second period (Figure 12b). The result shows that before the lockdown of Wuhan on January 23rd, cases mainly existed in the Jiang'an, Jianghan, Qiaokou, Hanyang, Wuchang, Hongshan, and Qingshan districts in the early stage (Figure 12).
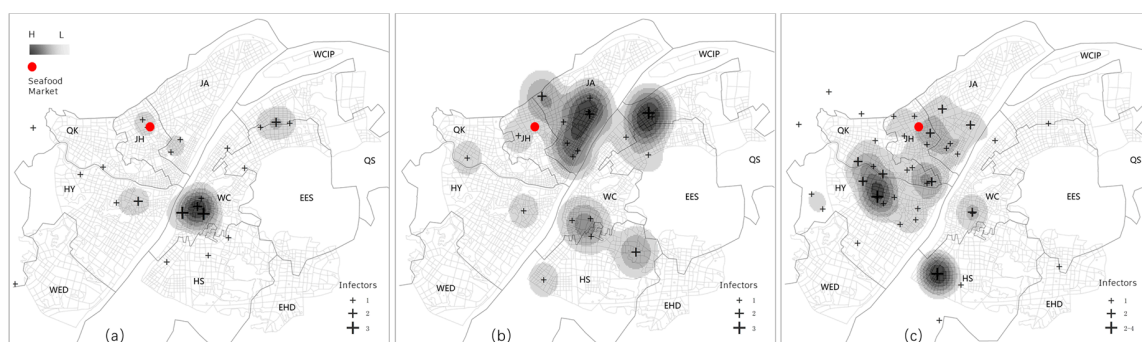
Figure 12. Spatial distribution of help seekers from December 20th, 2019 to January 22nd, 2020: (**a**) before January 18th, 2020; (**b**) from January 19th to 20th, 2020; and (**c**) from January 21st to 22nd, 2020.

The epidemic peak appeared around January 23rd, 2020 (Figure 13a). There had been multiple outbreak centers, and the higher ones were the hotspots in Qiaokou, Jiang'an, Wuchang, Hongshan, and Qingshan. It was found that the epidemic outbreak areas were all high-density residential areas, representing the entering stage of community transmission (Figure 13).
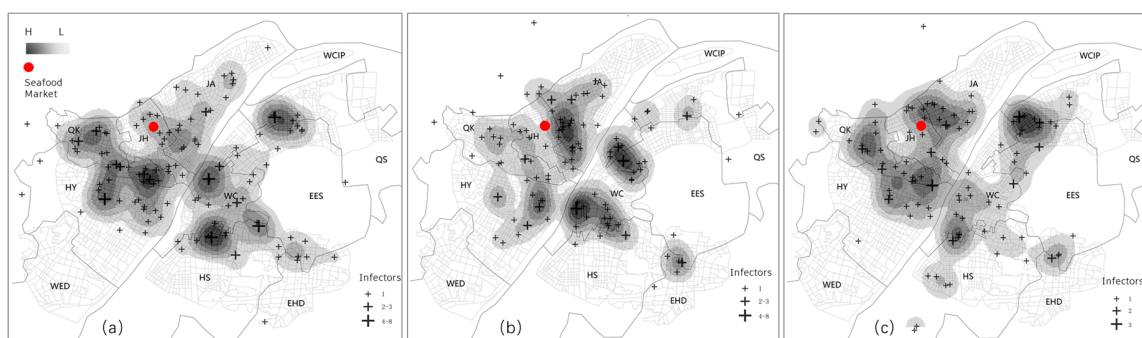
Figure 13. Spatial distribution of help seekers from January 23rd to 28th, 2020: (**a**) from January 23rd to 24th, 2020; (**b**) from January 25th to 26th, 2020; and (**c**) from January 27th to 28th, 2020.

From January 29th, 2020 to February 3rd, 2020, it was the sub peak in several periods (Figure 14a–c). The distribution of cases was in an average trend in all regions, with the hotspots concentrated in Jiang'an, Jianghan, Wuchang, and Qingshan districts, being still basically within the initial range. The number of the core density regions decreased while the range of transmission further expanded from February 4th to February 7th, 2020 (Figure 14d,e). By February 10th, 2020, the number of patients had been trending to single digit, and the spatial distribution characteristics were no longer typical in terms of sampling coverage (Figure 14f).
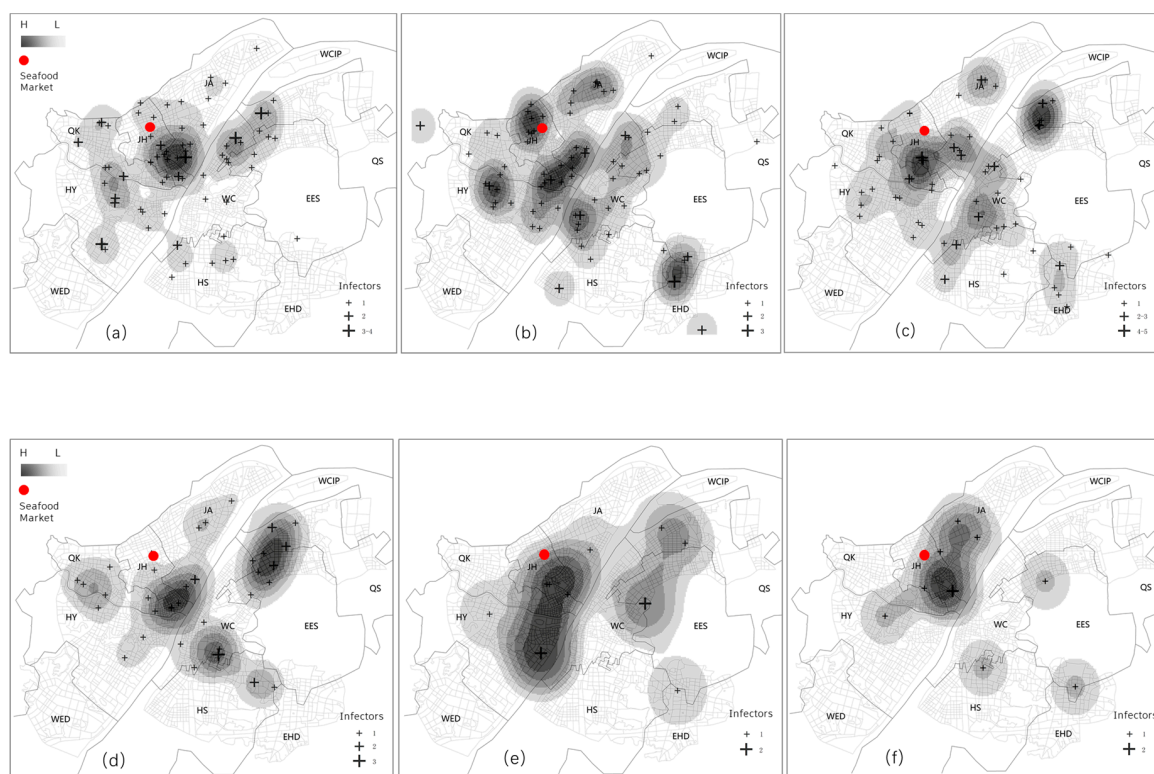
Figure 14. Spatial distribution of help seekers from January 29th to February 10th, 2020: (**a**) from January 29th to 30th, 2020; (**b**) from January 31st to February 1st, 2020; (**c**) from February 2nd to 3rd, 2020; (**d**) from February 4th to 5th, 2020; (**e**) from February 6th to 7th, 2020; and (**f**) from February 8th to 10th, 2020.

5. Discussion

The spatial distribution of COVID-19 cases extracted by Weibo data is highly correlated with the density of population, especially of the elderly population. As the first region of the outbreak in China, Wuhan entered the community transmission stage earlier. Therefore, it can be inferred that after the epidemic development for a period of time, the regeneration index R0 of each district tended to be consistent, leading to a strong correlation between the case and population. The closer relationship between the cases and the elderly population confirms that the elderly population is the high-risk group, consistent with the current medical observations [29,31]. Consequently, the study proposes that the elderly, as the susceptible population with high incidence and high risk of the COVID-19, should be the key target for active response in the epidemic prevention.

Compared with the traditional cases data, spatiotemporal data such as social media data, is more available, more time-sensitive. Differing from the previous study in the data sources [32], the study proposes social media data and mobile phone data to explore spatial features of novel coronavirus transmission. Weibo data is of great significance to identify the spatial distribution of infectors and mobile phone provides fine-scale population with age tags.

The results indicate that:

(1)　When the capita medical resources were extremely scarce, the incidence rates in urban areas tended to be the same after entering the community transmission period. The spatial distribution of help seekers was related to the regional population density. Since the elderly are more likely to convert severe groups, the space distribution of help seekers had a higher correlation with the density of the elderly population.

(2)　The new coronavirus epidemic showed the obvious spatiotemporal characteristic of scattered infection, community spread, and full-scale outbreak in the early stage, which was specifically manifested in the process of mobile diffusion centered on the early cases found in Jiang'an, Jianghan, Qiaokou, Wuchang, Hongshan, and Hanyang districts before January 23rd, the interior spread of each community that forming the polycentric structure after January 23rd, and the explosion process in which the density core area further spread later.

Compared with the research on the country level l [3–5], the spatiotemporal characteristics of the epidemic in emerging city can help us understand the detail inter-regional interactions in the urban space, and help other cities that are likely to have outbreaks or have had outbreaks to better understand the mechanism of disease transmission and the relationship between urban governance and public health epidemic prevention, so as to take appropriate protective measures in each stage, respectively.

During the epidemic period, countries have different principles for the use of privacy data in the general policy of epidemic prevention and control. When the detailed disease data that involves privacy issues is difficult to obtain, the spontaneous data provided by the media can be used as an effective means to provide a certain degree of reference information for the public and government departments. Meanwhile, due to the inaccessibility of data, we used the call detail record data of March 2017 in Wuhan city. The inconsistency of data on the time profile may lead to a certain difference in results, while we consider that such a time period will not change greatly in terms of macroscopic characteristics.

The release time of Weibo help information was mainly from February 3rd to February 12th. With the gradual completion of medical resources, help information tended to be no update. It can be seen from the spatial correlation analysis that the help seeking data in the period had good spatial coverage, consequently, the study believes that it can reflect the early three stages of disease transmission in Wuhan urban space from December 20th, 2019 to February 10th, 2020 to a certain extent. It should also be noted that, due to the small sample size of Weibo cases and the onset time were mostly after 2020, in spite of the good sample coverage, there was still missing information. The significance of our research is to make such an attempt on limited data, to carry out a retrospective analysis of the development of the epidemic situation, and to analyze the possible laws of the spread of disease in space and time, with a certain degree of verification. Therefore, if the detailed spatial distribution information of early patients can be obtained, a complete retrospective deduction of the spatial transmission path of the entire early epidemic can be obtained in a similar way.

Furthermore, the study can be expanded in the following aspects: (1) the further exploration of the influencing factors of epidemic to better understand the transmission mechanism and (2) combined with more detailed flow data to build agent model for further simulation and analysis.

6. Conclusions

In the context of the outbreak of the COVID-19 in Wuhan, the spatiotemporal mapping is of vital importance to understand the spatial features of the new coronavirus transmission mechanism. This study contributes to propose the combination of Weibo help seeking data and mobile phone data to achieve the quantitative study of the new coronavirus epidemic. Help seeking mapping can reflect the actual spatial distribution of the infectors who could not get timely treatment in the epidemic city under the condition of the short run of medical resources in Wuhan, and help us to identify the obvious transmission characteristics of epidemic development, which is specifically manifested in the earlier

scattered infection, community spread, and full-scale outbreak process on the whole. Simultaneously, the population derived from mobile phone data enables to find the high-similarity of distribution patterns between Weibo COVID-19 cases and the elderly population, which is verified by the result of the OLS model. Consequently, it can be used as the evidence for the elderly population groups being the susceptible population of the new coronavirus pneumonia. Moreover, the study proposes that elderly population should be the key target for active response in the epidemic prevention while corresponding measures should be deployed.

In general, the study based on the usage of Weibo data, mobile phone data, and other spatial big data resources can clearly identify the susceptible people and disease-prone spaces, and explore the process of the spatiotemporal dynamic spread of the new coronavirus, which is helpful to provide decision-making basis for disease prevention and control to a certain extent.

Author Contributions: Lingbo Liu and Ru Wang conceived and designed the experiments; Zhenghong Peng and Hao Wu acquired and analyzed the data; Hao Wu and Ru Wang contributed analysis tools; Lingbo Liu and Ru Wang wrote the paper. All authors have read and agreed to the published version of the manuscript.

Funding: The study was funded by National Natural Science Foundation of China (No. 51978535); Humanities and Social Science Project of the Ministry of Education (No. 19YJCZH187); Wuhan University Experiment Technology Project Funding.

Acknowledgments: The authors acknowledge the contribution of all the anonymous reviewers that improved the quality of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wu, F.; Zhao, S.; Yu, B.; Chen, Y.M.; Wang, W.; Song, Z.G.; Hu, Y.; Tao, Z.W.; Tian, J.H.; Pei, Y.Y.; et al. A new coronavirus associated with human respiratory disease in China. *Nature* **2020**, *579*, 265–269. [CrossRef] [PubMed]

2. Tian, H.; Liu, Y.; Li, Y.; Wu, C.-H.; Chen, B.; Kraemer, M.U.; Li, B.; Cai, J.; Xu, B.; Yang, Q. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* **2020**, *368*, 638–642. [CrossRef] [PubMed]

3. Gao, S.; Rao, J.; Kang, Y.; Liang, Y.; Kruse, J. Mapping county-level mobility pattern changes in the United States in response to COVID-19. *Sigspatial Spec.* **2020**, *12*, 16–26. [CrossRef]

4. Huang, R.; Liu, M.; Ding, Y. Spatial-temporal distribution of COVID-19 in China and its prediction: A data-driven modeling analysis. *J. Infect. Dev. Ctries.* **2020**, *14*, 246–253. [CrossRef] [PubMed]

5. Kang, D.; Choi, H.; Kim, J.-H.; Choi, J. Spatial epidemic dynamics of the COVID-19 outbreak in China. *Int. J. Infect. Dis.* **2020**, *94*, 96–102. [CrossRef]

6. Keil, R.; Ali, H. Governing the sick city: Urban governance in the age of emerging infectious disease. *Antipode* **2007**, *39*, 846–873. [CrossRef]

7. Neiderud, C.-J. How urbanization affects the epidemiology of emerging infectious diseases. *Infect. Ecol. Epidemiol.* **2015**, *5*, 27060. [CrossRef]

8. Jaglin, S. Rethinking urban heterogeneity. In *The Routledge Handbook on Cities of the Global South*; Routledge: Abingdon, UK, 2014; pp. 434–446.

9. Breiman, R.F.; Cosmas, L.; Njenga, M.K.; Williamson, J.; Mott, J.A.; Katz, M.A.; Erdman, D.D.; Schneider, E.; Oberste, M.S.; Neatherlin, J.C. Severe acute respiratory infection in children in a densely populated urban slum in Kenya, 2007–2011. *BMC Infect. Dis.* **2015**, *15*, 95. [CrossRef]

10. Huang, C.; Xu, X.; Cai, Y.; Ge, Q.; Zeng, G.; Li, X.; Zhang, W.; Ji, C.; Yang, L. Mining the Characteristics of COVID-19 Patients in China: Analysis of Social Media Posts. *J. Med. Internet Res.* **2020**, *22*, e19087. [CrossRef]

11. Penrose, K.; de Castro, M.C.; Werema, J.; Ryan, E.T. Informal urban settlements and cholera risk in Dar es Salaam, Tanzania. *PLoS Negl. Trop. Dis.* **2010**, *4*. [CrossRef]

12. Li, Q.; Guan, X.; Wu, P.; Wang, X.; Zhou, L.; Tong, Y.; Ren, R.; Leung, K.S.M.; Lau, E.H.Y.; Wong, J.Y.; et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N. Engl. J. Med.* **2020**, *382*, 1199–1207. [CrossRef] [PubMed]

13. Perlman, S. Another decade, another coronavirus. *Mass. Med. Soc.* **2020**. [CrossRef] [PubMed]

14. Bansal, S.; Chowell, G.; Simonsen, L.; Vespignani, A.; Viboud, C. Big data for infectious disease surveillance and modeling. *J. Infect. Dis.* **2016**, *214*, S375–S379. [CrossRef] [PubMed]

15. Lee, E.C.; Asher, J.M.; Goldlust, S.; Kraemer, J.D.; Lawson, A.B.; Bansal, S. Mind the scales: Harnessing spatial big data for infectious disease surveillance and inference. *J. Infect. Dis.* **2016**, *214*, S409–S413. [CrossRef] [PubMed]

16. Li, L.; Zhang, Q.; Wang, X.; Zhang, J.; Wang, T.; Gao, T.-L.; Duan, W.; Tsoi, K.K.-f.; Wang, F.-Y. Characterizing the Propagation of Situational Information in Social Media During COVID-19 Epidemic: A Case Study on Weibo. *IEEE Trans. Comput. Soc. Syst.* **2020**, *7*, 556–562. [CrossRef]

17. Zhao, Y.; Xu, H. Chinese public attention to COVID-19 epidemic: Based on social media. *medRxiv* **2020**. [CrossRef]

18. Li, J.; Xu, Q.; Cuomo, R.; Purushothaman, V.; Mackey, T. Data Mining and Content Analysis of the Chinese Social Media Platform Weibo During the Early COVID-19 Outbreak: Retrospective Observational Infoveillance Study. *Jmir Public Health Surveill* **2020**, *6*, e18700. [CrossRef]

19. Prabhakar Kaila, D.; Prasad, D.A. Informational flow on Twitter–Corona virus outbreak–topic modelling approach. *Int. J. Adv. Res. Eng. Technol.* **2020**, *11*, 128–134.

20. Jahanbin, K.; Rahmanian, V. Using Twitter and web news mining to predict COVID-19 outbreak. *Asian Pac. J. Trop. Med.* **2020**, *13*. [CrossRef]

21. Schild, L.; Ling, C.; Blackburn, J.; Stringhini, G.; Zhang, Y.; Zannettou, S. "Go eat a bat, Chang!": An Early Look on the Emergence of Sinophobic Behavior on Web Communities in the Face of COVID-19. *arXiv* **2020**, arXiv:2004.04046.

22. Li, S.; Wang, Y.; Xue, J.; Zhao, N.; Zhu, T. The Impact of COVID-19 Epidemic Declaration on Psychological Consequences: A Study on Active Weibo Users. *Int. J. Env. Res. Public Health* **2020**, *17*, 2032. [CrossRef] [PubMed]

23. Carroll, L.N.; Au, A.P.; Detwiler, L.T.; Fu, T.-c.; Painter, I.S.; Abernethy, N.F. Visualization and analytics tools for infectious disease epidemiology: A systematic review. *J. Biomed. Inform.* **2014**, *51*, 287–298. [CrossRef] [PubMed]

24. Cromley, E.K.; McLafferty, S.L. *GIS and Public Health*; Guilford Press: New York, NY, USA, 2011.

25. Weibo Reports Robust Q2 User Growth. Available online: http://www.xinhuanet.com/english/2019-08/20/c_138323288.htm (accessed on 1 June 2020).

26. Weibo Novel Coronavirus Pneumonia Help Seeking Channel. Available online: https://s.weibo.com/weibo?q=%23%E7%BA%BF%E4%B8%8A%E8%82%BA%E7%82%8E%E6%82%A3%E8%80%85%E6%B1%82%E5%8A%A9%E4%B8%93%E5%8C%BA%23 (accessed on 13 February 2020).

27. Spencer, J.; Angeles, G. Kernel density estimation as a technique for assessing availability of health services in Nicaragua. *Health Serv. Outcomes Res. Methodol.* **2007**, *7*, 145–157. [CrossRef]

28. Gerber, M.S. Predicting crime using Twitter and kernel density estimation. *Decis. Support Syst.* **2014**, *61*, 115–125. [CrossRef]

29. Response, E.W.G.f.N.E. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China. *Chin. J. Epidemiol.* **2020**, *41*, 145–151.

30. Miller, H.J.; Shaw, S.-L. *Geographic Information Systems for Transportation: Principles and Applications*; Oxford University Press: London, UK, 2001.

31. Huang, C.; Wang, Y.; Li, X.; Ren, L.; Zhao, J.; Hu, Y.; Zhang, L.; Fan, G.; Xu, J.; Gu, X. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **2020**, *395*, 497–506. [CrossRef]

32. Wesolowski, A.; Buckee, C.O.; Engø-Monsen, K.; Metcalf, C.J.E. Connecting mobility to infectious diseases: The promise and limits of mobile phone data. *J. Infect. Dis.* **2016**, *214*, S414–S420. [CrossRef]