
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Sustainable Cities and Society 67 (2021) 102752

Contents lists available at ScienceDirect

Sustainable Cities and Society

journal homepage: www.elsevier.com/locate/scs

# Association of built environment attributes with the spread of COVID-19 at its initial stage in China

Shuangjin Li [a,1], Shuang Ma [b,1], Junyi Zhang [c,*]

[a] Mobilities and Urban Policy Lab, Graduate School for International Development and Cooperation, Hiroshima University, Higashi Hiroshima, 739-8529, Japan
[b] Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, 153-8904, Japan
[c] Prof. Dr. Eng., Mobilities and Urban Policy Lab, Graduate School of Advanced Science and Engineering, Graduate School for International Development and Cooperation, Hiroshima University, Higashi Hiroshima, 739-8529, Japan

ARTICLE INFO

*Keywords:*
COVID-19
Initial stages of pandemic
China
The built environment
Spatial heterogeneity
Mixed GWR

ABSTRACT

Evidence of the association of built environment (BE) attributes with the spread of COVID-19 remains limited. As an additional effort, this study regresses a ratio of accumulative confirmed infection cases at the city level in China on both inter-city and intra-city BE attributes. A mixed geographically weighted regression model was estimated to accommodate both local and global effects of BE attributes. It is found that spatial clusters are mostly related to low infections in 28.63 % of the cities. The density of point of interests around railway stations, travel time by public transport to activity centers, and the number of flights from Hubei Province are associated with the spread. On average, the most influential BE attribute is the number of trains from Hubei Province. Higher infection ratios are associated with higher values of between-ness centrality in 70.98 % of the cities. In 79.22 % of the cities, the percentage of the aging population shows a negative association. A positive association of the population density in built-up areas is found in 68.75 % of county-level cities. It is concluded that the countermeasures in China could have well reflected spatial heterogeneities, and the BE could be further improved to mitigate the impacts of future pandemics.

## 1. Introduction

The spread of the novel coronavirus disease 2019 (COVID-19), as well as other communicable diseases, is due to transmission of viruses from people to people (Ferguson et al., 2005; Cummings et al., 2004; Gog et al., 2014; Poletto et al., 2014; Mousavinia et al., 2019). To capture how viruses spread it is necessary to determine how people communicate face-to-face. Unfortunately, there exists no published data regarding such face-to-face communication. It is also not easy to obtain such data. Mobile phone data can be used, especially considering its wide spatial coverage and massive volume (Wesolowski et al., 2016; Gonzalez et al., 2008). However, face-to-face communication data at the individual level is not available, even though zone-based spatial agglomeration of communication can be measured. Although questionnaire surveys can be implemented to ask people to report their face-to-face communication with others (Baym et al., 2004), it would be too costly to obtain a massive amount of data from a large proportion of the whole population.

It is obvious that the closer/denser social contacts are, the higher the infection risk will be. Facilitating social contacts is one of the various purposes of improving the built environment (BE) (Megaheda & Ghoneim, 2020). Thus, it is more feasible to make use of the BE attributes to examine the spread of COVID-19 than to directly use data of face-to-face communications, which can be collected based on advanced technologies but is only publicly available on a small scale. The BE concept is very broad, which refers to all types of human-made places for living, working, and playing (Basta and Moroni, 2013; Dearry, 2004; Anderson, 2019). The BE dimensions include both urban attributes (e.g., density, street connectivity, land use) and regional structure (distribution of transportation facilities across regions) (Handy et al., 2002). Some BE attributes were revealed to directly contribute to the spread of viruses. At the inter-city level, Ruan et al. (2015) confirmed that the total frequency and speed of travel affect the spread of communicable diseases. Hogbin (1985) examined the role of South Africa's railways in transmitting and preventing infectious diseases. At the urban level Yashima & Sasaki (2014) verified a relationship between the spread of

* Corresponding author.
*E-mail addresses:* d196414@hiroshima-u.ac.jp (S. Li), mashuang@cd.t.u-tokyo.ac.jp (S. Ma), zjy@hiroshima-u.ac.jp (J. Zhang).
[1] Joint first authors: these authors contributed equally to the work.

https://doi.org/10.1016/j.scs.2021.102752
Received 14 July 2020; Received in revised form 19 December 2020; Accepted 24 January 2021
Available online 3 February 2021

communicable diseases and the urban population size, commuting time and population flow by taking Tokyo, Japan as a case study. However, a later review shows there remain significant research gaps, in terms of the BE attributes targeted, analysis approaches, spatial scope and spatial heterogeneity, etc., regarding the association of BE attributes with the spread of the COVID-19. Addressing these gaps is not only important to identify clusters and mechanisms of infections and consequently take immediate proactive measures to effectively control the current pandemic, but also crucial to explore more scientifically sound methodologies for making policies against future pandemics.

In considering the features of the BE in reflecting direct and indirect face-to-face communications, this study attempts to fill the research gaps by focusing on the spread of COVID-19 in its initial stage in China, measured by the ratio of accumulative confirmed infection cases. Studies on China have attracted various international attentions in terms of epidemic analysis (e.g., Ahmed et al., 2020; Han & Yang, 2020; Peng et al., 2020), policy measures (e.g., Aleta et al., 2020; Hasnain et al., 2020), environmental impacts (e.g., Giani et al., 2020), and global development (e.g., Schindler et al., 2020), etc. Such attentions may be because China's responses can be a better reference for other countries to deal with both the current and future pandemics (Altakarli, 2020; Xu et al., 2020). This study presents additional evidence on China's experiences, but from a new perspective of the BE. Such a new perspective is also useful to research and policymaking in other countries, not only because this study shows how the BE attributes can be used to better understand the spread of the virus as a proxy of social contacts, but also because this study further shows how the BE can be improved to prevent future pandemics. This study focuses on the infection at the city level, by targeting the whole of China, and investigates the influences of both the inter- and intra-city BE attributes. The BE attributes vary across locations and, therefore, it is not unrealistic to assume that the associations between the BE attributes and the spread of COVID-19 will also vary across locations. Such spatial heterogeneity can be captured by estimating a geographically weighted regression (GWR) model. As summarized by Zhang, Hayashi, & Frank (2020), various serious impacts of COVID-19 on human society (e.g. economic activities and people's daily lives) have been found. It is, therefore, important to prepare for future pandemics from various aspects. In this regard, the importance of conducting the present research is obvious. A key contribution to practice is that this study presents initial evidence of the impacts of the various BE attributes on the spread of the virus in the context of China by focusing on the spread of COVID-19 in its initial stage. In this regard, this is the

**Table 1**
Major references about the association of the BE attributes with the spread of COVID-19.

| Source | BE Attributes | Analysis method | Scale | Major findings related to the present study |
|---|---|---|---|---|
| Nguyen et al., 2020 | Presence of a crosswalk, non-single family home, single-lane roads, dilapidated building and visible wires | Google Street View (GSV) images and computer vision; Poisson regression models | 164 million images in the USA | Indicators of mixed land use (non-single-family home), walkability (sidewalks) and physical disorder (dilapidated buildings and visible wires) were connected with higher COVID-19 cases. Indicators of lower urban development (single lane roads and green streets) were associated with fewer COVID-19 cases. |
| Hamidi et al., 2020a | Metropolitan population, activity density (population & employment per square mile), ICU beds per 10,000 population, primary care physicians per 10,000 population | Multi-level linear model | 1165 metropolitan counties in the USA | Larger metropolitan areas lead to significantly higher COVID-19 infection rates and higher mortality rates |
| Lee et al., 2020 | Traffic volumes on roads | Single linear regression | 6307 vehicle detection systems (VDS) in South Korea | In Incheon there was a positive, but insignificant, linear relationship between the increasing numbers of newly confirmed cases and increasing traffic. |
| Ghosh et al., 2020 | Travel distance to London, population density | Mixed-effects model | Distance from London to four other cities (Birmingham, Leeds, Manchester and Sheffield) | As the distance from London increases, the number of COVID-19 cases decreases. |
| Mizumoto & Chowell, 2020 | Occupant density on the Diamond Princess cruise ship | Mathematical modeling | 621epidemiological incidence cases | The increased exposure risks associated with high occupant density were demonstrated in the COVID-19 outbreak that occurred on the ship. |
| Emeruwa et al. 2020 | Building-level variables, including the number of residential units per building and mean assessed value (per square foot), and neighborhood-level variables, including population density, household membership (persons per household) and household crowding. | Bivariable logistic Regression model | 71 infected cases in New York | COVID-19 transmission among pregnant women was associated with neighborhood- and building-level markers of large household membership and household crowding. |
| Dai & Zhao, 2020 | Ventilation rate | Wells–Riley equation | Typical scenarios, including offices, classrooms, buses and aircraft cabins. | An infection probability of less than 1% requires a ventilation rate larger than 100–350 $m^3$/h per infector and 1200–4000 $m^3$/h per infector for 0.25 h and 3 h of exposure. |
| Antony, Velray & Fariborz, 2020 | Population density, climate severity, the volume of indoor spaces and air-conditioning usage | Statistical analysis of correlations | Various states in India | Fast drying and size reduction of respiratory droplets makes the virus more active. |
| Auger, Shah & Richardson, 2020 | Schools | Population-based time series analysis | All USA states | School closure was associated with a significant decline in the incidence of COVID-19 and mortality. |
| Brown et al. 2020 | Nursing homes crowding | Population-based retrospective cohort study | 78,000 residents of 618 distinct nursing homes in Ontario, Canada | Crowding in nursing homes was associated with a higher incidence of COVID-19 infection and mortality. |
| Hamidi et al., 2020b | County activity density and metropolitan area population | Structural equation model | 913 metropolitan counties in the USA | Metropolitan population is one of the most significant predictors of infection rates; larger metropolitan areas have higher infection and higher mortality rates. |

first study in literature, and especially, this study examined more intra-/inter-city BE attributes than all of the existing studies. Methodologically, both global and local effects of the BE attributes are jointly captured by estimating a mixed GWR model. Policy measures against pandemics should pay more attention to various locational contexts. Such a model presents a powerful tool for policymaking sensitive to locations.

In the remaining part of this paper a literature review is firstly provided to better position this study in the literature. Secondly, the BE attributes selected for this study are described. Thirdly, the adopted methods of local indicators of spatial association (LISA) and the mixed GWR model are briefly explained. Fourthly, the associations of the BE attributes with the spread of COVID-19 in its initial stage in China are analyzed in detail, with their policy implications also being discussed. Finally, the findings of this study are summarized, together with a discussion about future research issues.

## 2. Literature review

Although there exist many publications about COVID-19 at the time of conducting this study, for example there are more than 6000 and 8000 references on Web of Science and ScienceDirect, respectively, less than 1% are related to the BE or buildings (Pinheiro & Luís, 2020; Raj, Velray, & Haghighat, 2020). The major references are summarized in Table 1 in order to better position the present study. Road attributes, building and housing attributes, population or its density, medical facilities and services, and schools, etc., have been examined. For instance, based on street view images, Nguyen et al. (2020) explored how sidewalks, dilapidated buildings and visible wires were associated with COVID-19 infection cases in USA. Through a multi-level linear model, Hamidi et al. (2020a) examined the impacts of population, activity density (population & employment per square mile), ICU beds, primary care physicians on infection rates and higher mortality rates in USA. By building a structural equation model, Hamidi et al. (2020b) revealed that metropolitan population is one of the most significant predictors of infection rates in USA. Raj et al. (2020) investigated the relationships between virus viability and populating density, climate severity, the volume of indoor spaces and air-conditioning usage in India, through correlation analyses. Traffic volumes on road and travel distance are identified as important factors to explain newly confirmed cases of COVID-19, by Lee et al. (2020) who focused on Incheon, South Korea and estimated a linear regression model, and by Ghosh et al. (2020) who focused on four UK cities and estimated a mixed-effects model, respectively. Through a bivariate logistic regression model, the household crowding, number of residential units per building, and household structure are found to be associated with COVID-19 transmission in New York (Emeruwa et al., 2020). Auger, Shah & Richardson (2020) showed that school closure was associated with a significant decline in the incidence of COVID-19 and mortality in USA through a population-based time series analysis. Furthermore, Brown et al. (2020) found that crowding in nursing homes was associated with a higher incidence of COVID-19 infection and mortality in Ontario, Canada through a population-based retrospective cohort study.

However, only a few BE attributes were targeted in each study and very limited studies have captured the BE from both inter-city and intra-city perspectives. For inter-city BE attributes, the contributions of travel distance are explored (Ghosh et al., 2020), as well airlines or train flows (Lau et al., 2020). However, no study can be found to deal with the number of trains, the number of flights and population flows across cities at the same time, as done in this study. In addition, for intra-city BE attributes, this study further selected the BE attributes of between-ness centrality of major transport nodes to reflect the influence of major transport demand, and point of interests (POIs) around train stations to reflect activity density. Moreover, the USA is a popular study area, but no studies can be found related to China. In particular, all of the existing studies focus on only the global effects of BE attributes based on some regression models, where the effects of BE attributes sensitive to locations have been ignored. Moreover, spatial clusters of the COVID-19 spread have not been investigated from a geographical perspective.

There are some studies (see Table 2) that only mention the BE attributes in association with COVID-19, such as hospitals (Rothan & Byrareddy, 2020; Gan et al., 2020), prisons and churches (Kim, 2020), public transport (European Commission, 2020), and building and indoor occupants (Dietz et al., 2020; Eykelbosh, 2020; Saadat et al., 2020). For instance, CDC, USA (2020) stated that people living and working in shared buildings might have challenges with social distancing measures to prevent the spread of COVID-19. Capolongo et al. (2020) suggested the importance of planning of smart and sustainable mobility networks to prevent the COVID-19. But all these studies do not include any empirical analyses of the relationships between BE attributes and the spread of the virus.

In order to understand the spread of COVID-19, some modeling approaches have been proposed. The network inference model and the susceptible infected recovered (SIR) model categorize the urban population into susceptible population, infected population and recovered population, by including the strength of the connections within and between the sub-groups of the city (Wang et al., 2018; Kuddus et al., 2014; Mccluskey, 2010). The spatially explicit disease transmission model takes account of the population information, such as age, place of residence and individual activities (Germann et al., 2006). At the individual level Yang et al. (2008) developed a space-time activity-based

**Table 2**
The BE attributes targeted in the existing studies of COVID-19.

| Source | BE attributes | Explanation |
|---|---|---|
| Rothan & Byrareddy 2020; Gan et al., 2020 | Hospital facility | Transmissions of COVID-19 are more likely to occur within the hospital BE. |
| Kim, 2020 | Prisons and churches | Accumulating evidence indicates that COVID-19 can spread widely in confined settings such as prisons, and churches. |
| European Commission, 2020 | Public transport | Public transport is also a high-risk environment for the spread of COVID-19, due to the large number of people gathering together in a confined environment. |
| Eykelbosh, 2020 Dietz, et al., 2020 | Inside buildings | Through building operators, all indoor occupants, ventilation and indoor air quality, lighting and the deposition on the surfaces of materials can reduce the spread of COVID-19. |
| Chang, 2020 Megahed & Ghoneim, 2020 | Population density | Close contact among people is very high in urban areas rather than rural areas. |
| Saadat et al., 2020 | Household size | A household with more members will have a higher chance to bring COVID-19 home, because there are more connections among people. |
| CDC, USA, 2020 | Shared facilities | Shared housing includes a broad range of settings with special considerations. People living and working in this type of building might have challenges with social distancing to prevent the spread of COVID-19. |
| Capolongo et al., 2020 | Accessibility | Re-thinking the accessibility to the places of culture and tourism. |
| Capolongo et al., 2020 | Mobility network | Planning of a smart and sustainable mobility network |
| Capolongo et al., 2020 | Semi-private space | Re-thinking building typologies, fostering the presence of semi-private or collective spaces; |
| Budds, 2020 | Social distancing | Social distancing could change the design and planning process, specifically with the increased acceptance of distance learning, online shopping, and the cultural connection of online entertainment. |

model to capture the spatial adjacency and social network relationships of individual activities. Such an activity-based approach by using data collected via an activity diary is powerful to capture activity-travel behavior in various locations and at different timings if a massive amount of data is available; however, such diary surveys are usually costly and time-consuming. Even though a massive amount of data could be obtained, there remains no scientifically sound data about how spatial-temporal activity-travel behavior is associated with different levels of the spread of COVID-19.

### 2.1. Research gaps in existing studies, and features and contributions of this study

The above literature review revealed major research gaps in the existing studies. First, no studies on China can be found associated to the BE attributes. Second, there are few studies examining both the intra- and inter-city BE attributes at the same time. Third, only very limited BE attributes were targeted in each existing study. Fourth, the effects of the BE attributes have been assumed to be spatially homogeneous (i.e. only global effects are examined). Fifth, spatial clusters of the COVID-19 infection have not been well revealed. In order to fill the above research gaps, this study: (1) investigates the spread of COVID-19 in the whole of China during the initial period of the spread between January 20 and February 3, 2020 when Wuhan city was locked down on January 23, 2020; (2) incorporates both global and local effects of the BE attributes by estimating a mixed GWR model, which is well supported by a careful analysis of spatial clusters of the spread; and (3) focuses on both inter-city and intra-city BE attributes, which is more than the existing studies have done.

## 3. Data

The spread of COVID-19 was measured by the cumulative number of confirmed infection cases in each city collected by the China University of Geoscience[2] from various official websites of national, provincial and municipal health organizations. This study targeted Chinese 368 cities which have publicly accessible data about numbers of confirmed infection cases during the period. Among the 368 cities, only 255 cities have valid data of BE attributes. On January 23, 2020 Wuhan city was locked down. Therefore, we use 255 cities as our study area (for Local indicators of spatial association model, 11 cities in Hubei province are excluded). According to World Health Organization (2020), Yu et al. (2020) and Lauer et al. (2020), the incubation period of COVID-19, which is the length of time between exposure to the virus and the symptom onset, is on average 5–6 days, but can be as long as 14 days. We extracted data from January 20 (three days before the lockdown of Wuhan City) to February 3, 2020 and defined this period as the initial stage of COVID-19 in China. We further collected population data from the *China City Statistical Yearbook* and calculated the ratio of the accumulative infection cases in each city, that is the accumulative number of confirmed infection cases divided by the city's population (hereafter "infection ratio" in short).

Regarding the BE attributes, there exist several available nationwide datasets. The first data set relates to the POIs from the Amap electronic navigation map, which has a high resolution and is frequently updated. The POIs data comprises of a total of 24,274,252 POIs in the above 255 cities and 23 categories, including shopping service, catering service and transportation facility, etc. The POIs are used to calculate the density of the POIs around railway stations, and are also used to identify urban railway stations and residential areas, as well as activity centers. The second data set includes national road networks, derived from the Open Street Map (OSM), and lists a total of 1,198,099 roads, including main roads, secondary roads and pathways, in the 255 cities. We processed this second set of data based on Long & Liu's (2017) three-step method[3]. This data set is used to measure the relative positions of railway stations in a city. The third data set relates to trains information and includes a total of 6264 trains and 67,780 stopover records, with the number of trains, names of stopover stations, departure time and arrival time being available. The fourth data set comprises of the number of flights and includes a total of 12,937 flights departing from, or stopping over in, the cities in Hubei province and going to the above 255 cities between December 1, 2019 and January 20, 2020. The fifth relates to the urban built-up areas and population from the *China City Statistical Yearbook*. We calculated the population density of built-up areas. The final data set provides the population flow from Baidu Migration Production based on its location-based services, which is published by Baidu (available at: http://qianxi.baidu.com). For a city we computed the cumulative outflow population percentage from Hubei Province, which is a cumulative value of the daily proportion of the outflow population from Hubei Province to the city, during the period January 1 to January 20, 2020.

## 4. Specification of the BE attributes

The density information at the level of buildings, such as household density (Emeruwa et al. 2020), nursing homes crowding (nursing homes crowding), or ventilation rate in particular architecture type (Dai & Zhao, 2020) is more useful to capture face-to-face contacts than the built environment attributes at city level. However, information of infections within buildings is not available, widely across a whole country. Considering that available information of infection in this study is accessible at city level, we selected as many publicly available city-level BE attributes as possible. In particular, during the initial stage of the pandemic in Wuhan, inter-city passenger transport systems transported many infected persons. Thus, it is indispensable to incorporate the influence of inter-city connections into the pandemic analysis, such as population flows across cities via railways and flights, etc. Within cities, transport systems also transported many infected persons, via urban transport networks and people-to-people contacts are made through a variety of activities at various locations connected by transport systems. Such intra-city connections should also be reflected in our analysis.

### 4.1. Inter-city BE attributes

The COVID-19 outbreak occurred during the busiest travel season in China, namely the famous Spring Festival. Many people used railways to return to their hometowns. Thus, we first selected the number of trains as an inter-city BE attribute. Concretely, the number of trains departing from and stopping over in all of the cities in Hubei Province (hereafter "number of trains" in short). In addition, some scholars pointed out the contributions of the inter-city flights to the spread of COVID-19 in different countries. For instance, Lau et al. (2020) showed that there was an association between international and domestic air traffic and the COVID-19 outbreak, especially the total number of flights from Hubei, China. Murphy et al. (2020) revealed the link between national outbreak of COVID-19 and air travel in Ireland. Therefore, we also selected the total number of flights departing from and stopping over in all of the cities in Hubei Province ("number of flights" for short). Even though the above numbers of trains and flights can indirectly reflect the total

[2] https://github.com/Estelle0217/COVID-19-Epidemic-Dataset (Accessed on October 2, 2020).

[3] (1) Use the "Merge Divided Roads" under ArcGIS Cartography Toolbox to merge the OSM road network, thereby merging multiple lanes into a single lane; (2) Use the "Thin Road Network" under the ArcGIS Cartography Toolbox to simplify the road network and delete thin side roads; (3) Perform a topology check on the road network and deal with topological errors to ensure that there are no broken roads and that there is only one intersection when the roads intersect.

number of passengers transported, including infected persons, and the frequency of contacts between infected and susceptible individuals, these numbers actually represent supply capacities of major transport systems. In other words, actual transport demand is unknown. This motivated us to search for other data sources and eventually, we found available population flow to reflect the inter-city connections in terms of transport demand. According to the above Baidu website, this flow data covers all purposes of inter-city population mobility, captured via mobile phone cellular signaling data (more than 120 billion times per day of accessing location-based services). Jia et al. (2020) argued the importance of tracking aggregate population flows from an epidemiological perspective. This also supports the use of the inter-city population flow in this study.

### 4.2. Intra-city BE attributes

Regarding the intra-city BE attributes, we selected between-ness centrality of key transport nodes (here, railway stations) and transport accessibility (here, measured in terms of travel time by public transport) between residential areas and major activity centers, as the intra-city BE attributes. Railway stations are selected because they are not only key clusters of population flows within cities, but also the main gates to connect with other cities in China. Thus, railway stations play a dual role of defining both intra- and inter-city BE attributes. The travel time by public transport is used to capture exposure of passengers to infection risks inside public transport vehicles. The pandemic has been caused by massive transmission among people. In addition to the aforementioned connection-related BE attributes, the potential of massive transmission can be further measured from two more angles: the first one focuses on various facilities allowing people to make contacts with others, and the second looks at the population density, which reflects the intensity of potential exposure and transmission via social contacts. Concretely, POIs density around railway stations and population density of built-up areas are adopted.

(1) Betweenness centrality of railway stations

Between-ness centrality is one of the important attributes to measure the centrality of transportation networks, which reflects the times of the passed shortest routes between the railway stations in the entire city road networks (Lin et al., 2018). The higher the between-ness value of a railway station is, the more interactions people can generate in the station. It has been verified that this attribute in urban networks is imperative in the dynamics of the spread of infectious diseases via direct person-to-person transmission (Salathe and Jones, 2010), but its influence on COVID-19 has not yet been examined. Thus, quantifying the position of railway stations within a city is important. Here, we measured the between-ness centrality of the road networks surrounding the railway station under study.

We used the Edge Between-ness Centrality in Python-igragh to calculate the between-ness centrality of the road networks of all of the 255 cities across the whole country. We used the average value of the between-ness centrality of the roads around the railway station (within a radius of 1.5 km) as the station's between-ness centrality. Furthermore, we calculated the between-ness centrality of the city using the average value of the between-ness centrality of the railway stations in a city. We detected a total of 1410 railway stations from the POIs data by further matching with the data of the number of trains and stopover stations.

Let a road network be $G = (N, E)$, which has $N$ nodes and $E$ boundaries. Then, the between-ness centrality of the boundary $e$ is defined below.

$$C_e^b = \sum_{\substack{j,k \\ j \neq k}} \frac{|SP_{jk(e)}|}{|SP_{jk}|}$$

Here, $|SP_{jk}|$ represents the number of the shortest paths between a pair of nodes $j, k$ (, $k \in N$), and $|SP_{jk(e)}|$ indicates the number of the shortest paths between $j$ and $k$ that include the boundary $e$ ($e \in E$)

(2) Travel time by public transport between residential areas and major activity centers

We used travel time by public transport from residential areas to the closest activity centers as a proxy of intra-city transport accessibility; that is the connection between the city's main attractions and residential areas. Activity centers are important nodes in urban networks and usually attract a large number of people. Some studies have found that the travel time by public transport from residential areas to activity destinations (e.g., the city center) has a major impact on the spread of epidemics within cities (Yashima & Sasaki, 2014; Balcan & Vespignani, 2011; Balcan & Vespignani, 2012).

There are a few methods to identify the centers based on employment density and local density peaks (Hajrasouliha, & Hamidi, 2017; Park et al. 2020). We identified activity centers based on the POIs density of shopping and catering services in a city, as follows. First, the Create Fishnet tool in the ArcGIS was used to build a grid of 500 by 500 m. Second, the Spatial Connection tool was adopted to calculate the density of the POIs of shopping and catering services in each grid. After the above two steps, the following two methods were further applied to identify the candidates of activity centers separately. The first method was to use Python to identify the grids that exceed a certain density (25 % of the total density of the shopping and catering services in the city). The second was to conduct a Getis-Ord $G_i^*$ analysis of the density of the POIs for shopping and catering services in the city and to select the hotspots grids (Fig. 1). Finally, activity centers were delineated as the intersections of candidate activity centers. Getis-Ord $G_i^*$ is illustrated below, the details of which refer to Getis and Ord (1992).

$$G_i^*(d) = \frac{\sum_{j=1}^{n} w_{ij}(d)x_j - \overline{X}\sum_{j=1}^{n} w_{ij}(d)}{S \times \sqrt{n\sum_{j=1}^{n} w_{i,j}^2 - \frac{\left(\sum_{j=1}^{n} w_{i,j}(d)\right)^2}{n-1}}}, \quad \overline{X} = \frac{\sum_j x_j}{n}, \quad \mathrm{Z}(G_i^*) = \frac{G_i^* - E(G_i^*)}{\sqrt{Var(G_i^*)}}$$

Here, $S$ is the standard deviation of the POIs density (POIs of shopping and catering services) of all grids; $x_j$ refers to the POIs density of the $j$ th gird; $G_i^*$ is a weighted value of $x_j$, $w_{ij}(d)$ is the spatial weight matrix with values for all grids '$j$' within a distance $d$ from target grid $i$ ; $\mathrm{Z}(\mathrm{G}_i^*)$ is a standardized value of $G_i^*$ ; $n$ is the number of grids; and $E(G_i^*)$ and $Var(G_i^*)$ denote the mean and variance of $G_i^*$, respectively. If $\mathrm{Z}(G_i^*)$ is higher than 1.96, the area is called a hot spot, while, if $\mathrm{Z}(G_i^*)$ is lower than -1.96, the area is called a cold spot.

Eventually, we identified a total number of 976 activity centers in 255 cities across the whole of China. After the above step, we calculated the travel time based on Baidu Maps' path-planning interface. The interface is provided by the Baidu Maps' API and is highly accurate and authentic. It precisely takes account of details, such as the walking time and transfer time during the use of public transport. We selected 04:00–5:00 pm on November 27, 2019 (a weekend) to calculate the travel time.

(3) Density of POIs around railway stations

Many concentrated infection cases of COVID-19 were observed owing to the concentration of activities (Dong et al., 2020; Gao et al., 2020), as the concentration of activities at railway stations allows infected persons transported via inter-city transport modes to quickly transfer to urban areas. To represent such a concentration the density of the POIs around railway stations is measured. It has been claimed that COVID-19 emerged from the South China Wet Market in Wuhan, although this has not been firmly confirmed and it is also unclear whether or not the market was the only source of the outbreak. This market is actually less than 1.0 km away from the Hankou Railway station. In this study we firstly established a buffer zone in ArcGIS through the "spatial connection" tool with 2.0 km to calculate the number of POIs, which can represent the distribution of urban facilities

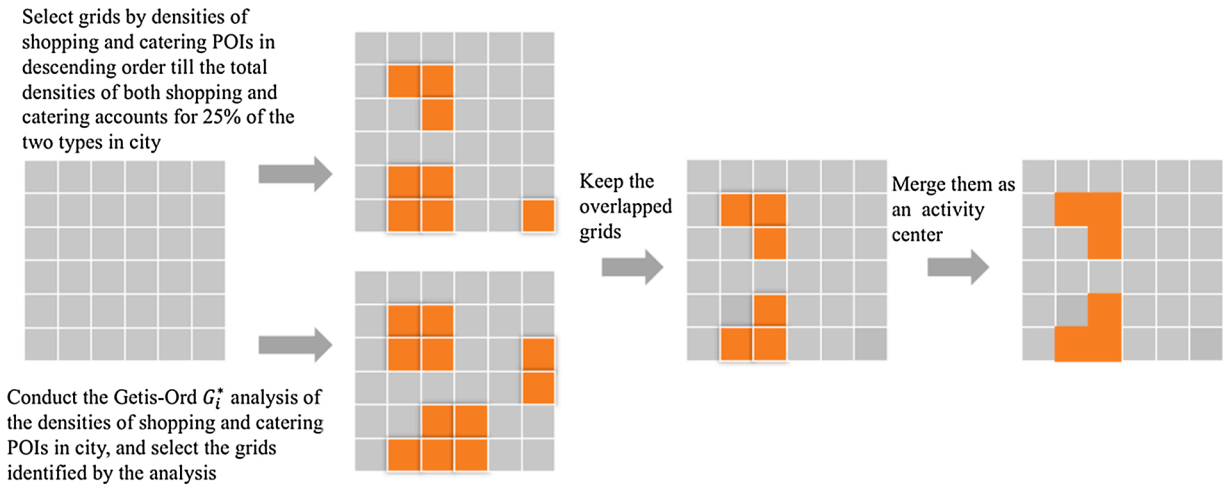



**Fig. 1.** A process of identifying activity centers in city.

around railway stations, and then calculated the POIs density of each station. The concentration of activities around the railway stations in a city is the average value for the POIs density around railway stations.

(4) Population density of built-up areas

Population density of built-up areas reflects possible interaction intensity among people. It is a major attribute affecting the spread and decay patterns of COVID-19 (Rashed et al., 2020). Similarly, Hamidi et al. (2020b) revealed the impact of population in metropolitan counties in USA on COVID-19 infection rates. Considering such evidence, we applied the population density of built-up areas as an additional intra-city BE attribute to examine its influence on COVID-19 spread in initial stages in China.

### 4.3. Control variable

COVID-19 has highlighted the vulnerability of elderly people (Wayne et al., 2020). However, the percentages of the aging population (65+ years old) vary across cities in China. Therefore, we chose the percentage of the aging population as a control variable in the following modeling analyses to capture the associations of BE attributes with the spread of COVID-19 in a proper way.

## 5. Methods

### 5.1. Local indicators of spatial association (LISA)

The LISA approach can detect the spatial autocorrelation. Compared with the conventional Moran's I measuring the global spatial autocorrelation, LISA can help to detect potential clusters of local spatial units, even though there is only a weak global spatial autocorrelation (Chen et al., 2017). LISA has been widely used in various research fields, such as geography (Cao et al., 2019), medical research (Hendricks & Mark-Carew, 2017) and city planning (Talen, & Anselin, 1998). In this study we use LISA to detect the spatial clusters of infection cases. LISA uses the following formulation:

$$I_i = Z_i \sum_{j=1}^{N} w_{ij} Z_j$$

Here, $N$ is the number of spatial units. $Z_i$ is the standardized value of the log-transformed infection ratio at locations $i$: that is $Z_i = \frac{x_i - \bar{x}}{\delta}$, where $\delta$ is the standard deviation of the log-transformed infection cases $x_i$ with its mean being $\bar{x}$. $w_{ij}$ is the spatial weight between the $i$ th and $j$ th locations. The significance of $I_i$ can be assessed using the conditional randomization approach (Anselin, 1995). Here, we set the threshold of significance at the 0.1 level. As stated by Chen, Liu & Li, (2017), if an above-average value is surrounded by above-average neighbors, then the location is categorized as High-High, while, if a below-average value is surrounded by below-average neighbors, then the location is categorized as Low-Low. In contrast, if an above-average value is surrounded by below-average neighbors, then the location is categorized as High-Low, and vice-versa (i.e., Low-High). The High-High and Low-Low types indicate spatial clusters, while the other two types represent spatial outliers.

LISA is more logical to identify spatial clusters and outliers than a direct comparison of infection levels at different locations. As stated in the Tobler's First Law of Geography, a formulation of the concept of spatial autocorrelation proposed by the geographer Waldo Tobler (Tobler, 1970), "everything is related to everything else, but near things are more related than distant things". Use of the four types of cities identified by LISA allows us to reflect the association of the infection rates in surroundings cities of a city under study.

### 5.2. A mixed GWR model

A mixed GWR model jointly estimates the explanatory variables with both local and global effects (e.g., Kang et al., 2010), which can be formulated as follows:

$$y_i = \sum_k \beta_k(u_i, v_i) x_{ik} + \sum_m \gamma_m Z_{im} + \varepsilon_i$$

where $y_i$ is the dependent variable at location $i$; $x_{ik}$ is the $k$ th explanatory variables at location $i$; $\beta_k(u_i, v_i)$ is the coefficient of $x_{ik}$ with a geographical coordinate $(u_i, v_i)$; $Z_{il}$ is the $m$th explanatory variable without local effects, that is its coefficient $\gamma_m$ is invariant across locations (showing global effects); and $\varepsilon_i$ is an error.

The above mixed GWR model turns to the standard GWR model when global effects can be ignored and, when local effects can be ignored, the mixed GWR model turns to the traditional multiple linear regression (MLR) model. The MLR model (e.g., Rath et al., 2020) assumes that the explanatory variables are spatially stationary over the whole study area, which, therefore, only provides global estimates (Yu & Peng, 2019).

In this study the above three regression models were all estimated, with the dependent variables being the infection ratio. The explanatory variables include the following seven BE attributes and a control variable. The first BE attribute is the number of trains, which is used to capture the large-scale population movement from Hubei Province. The second is the between-ness centrality, which is used to reflect the relative positions of the railway stations in a city. The third is the density of

the POIs around the railway stations, which services as a proxy of the intensity of activity participation and face-to-face contact. The fourth is the travel time by public transport (from residences to the nearby activity centers). The fifth is the population flow from Hubei Province. The sixth is the number of flights from Hubei, which also reflects the population movement between cities. The last BE attribute is the population density of built-up areas, which indicates the intensity of the connections between people. Finally, the percentage of the aging population is used as the control variable for better capturing of the associations of the BE attributes with the spread of COVID-19 at its initial stage in China.

## 6. Results

Here, the data from the 255 cities with the available spatial information is used for two types of analysis, based on the LISA and regression models, respectively. LISA is used to identify the spatial clusters of the spread of COVID-19 and the regression models are used to examine the associations of the BE attributes with the spread of the virus. LISA can be used for all of the 255 cities; however, the cities in Hubei province (11 cities in total) are excluded, as these cities had an extremely high number of infection cases, which are obviously spatial clusters. Thus, LISA was conducted with only cities outside Hubei (244 cities in total).

### 6.1. Spatial patterns of the spread of COVID-19

In this study we used the GeoDa software to calculate the local spatial autocorrelations of the spread of COVID-19. As shown in Fig. 2, the LISA approach found spatial clusters of the spread of COVID-19 (measured in terms of the ratio of accumulative infection cases). Among the 244 cities 112 showed positive local autocorrelations, with 39 cities belonging to the High-High type and 73 belonging to the Low-Low type. Only 10 cities showed negative local autocorrelations (5: High-Low; 5: Low-High) indicating spatial outliers. The High-High type (denoted in the red color in Fig. 2) indicates that a city with a high infection ratio is surrounded by cities with high infection ratios. In other words, the High-High type reveals serious infection clusters, which are mainly located at the southeastern region, such as in Shanghai and in the cities at the north and south of Hubei Province. The Low-Low clusters (denoted in the blue color: clusters being less serious) are mainly found at the west of the Heihe-Tengchong line, which is an imaginary line that divides China into two roughly equal parts according to area. The High-Low outliers (denoted in the pink color) tend to be scattered, while the Low-High outliers tend to be gathered, around Hubei Province.

We further compared the BE attributes in the above four types of spatial patterns (Fig. 3). The cities of the High-High type have more trains departing from and stopping over in Hubei Province. This is also true for the Low-High cities. The between-ness centrality has higher values in the Low-High cities. Differently, the travel time by public transport and the POIs around railway stations show smaller variations across the city types. The travel times by public transport in cities with positive local autocorrelations (High-High and Low-Low) are longer than they are in the other two types of spatial patterns, while the High-High and Low-Low cities have more dense POIs around the railway stations. Similar to the number of trains, the population flow is also higher in the High-Low and Low-High types. In contrast, there are fewer flights from Hubei to the cities of these two types of spatial patterns. The population density of built-up areas is higher in the High-Low type.

### 6.2. Analysis of influential attributes based on the mixed GWR model

Table 3 presents the results of the Global Moran's I test. Five variables were identified to be statistically significant at the 0.01 level, indicating that they are spatially clustered and, therefore, the local parameters for them should be estimated. In contrast, three variables were tested to have no spatial cluster and, consequently, global parameters should be introduced to these three variables. With these results we

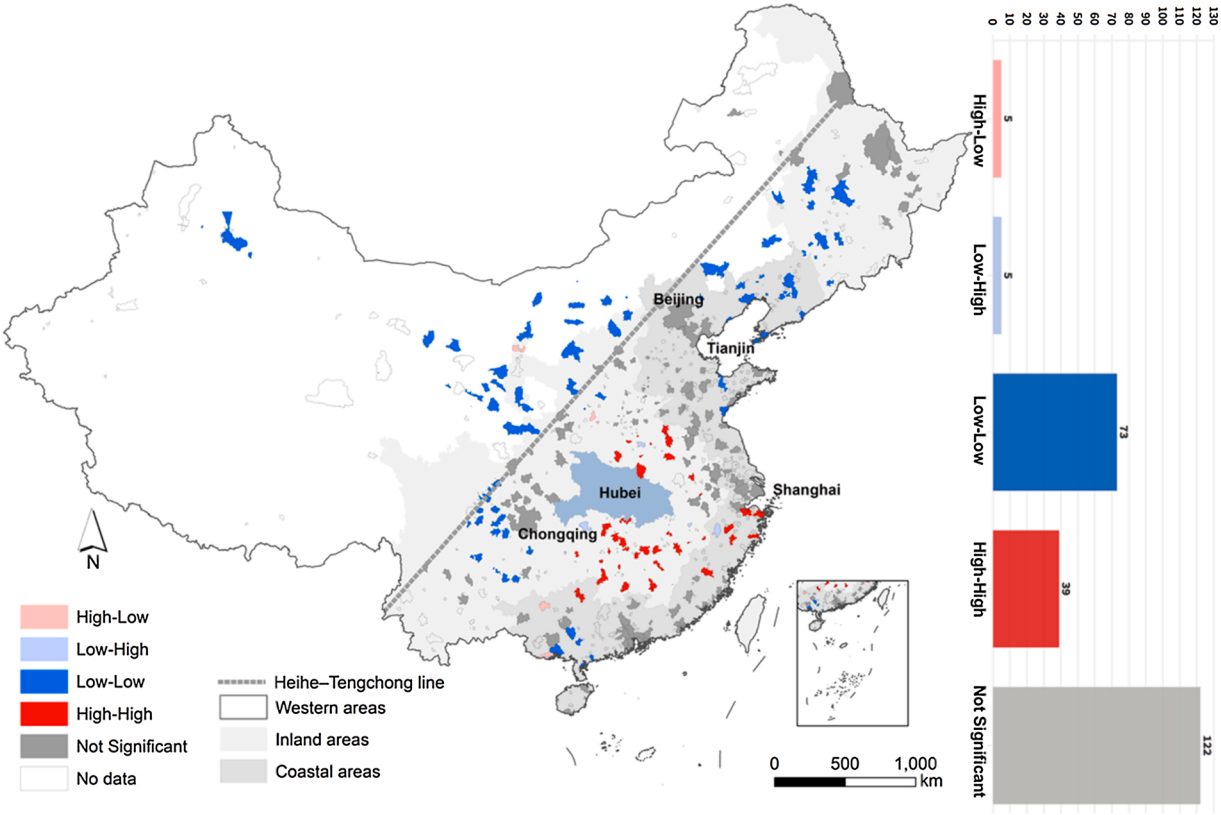



**Fig. 2.** Spatial patterns of the spread of COVID-19 based on LISA.

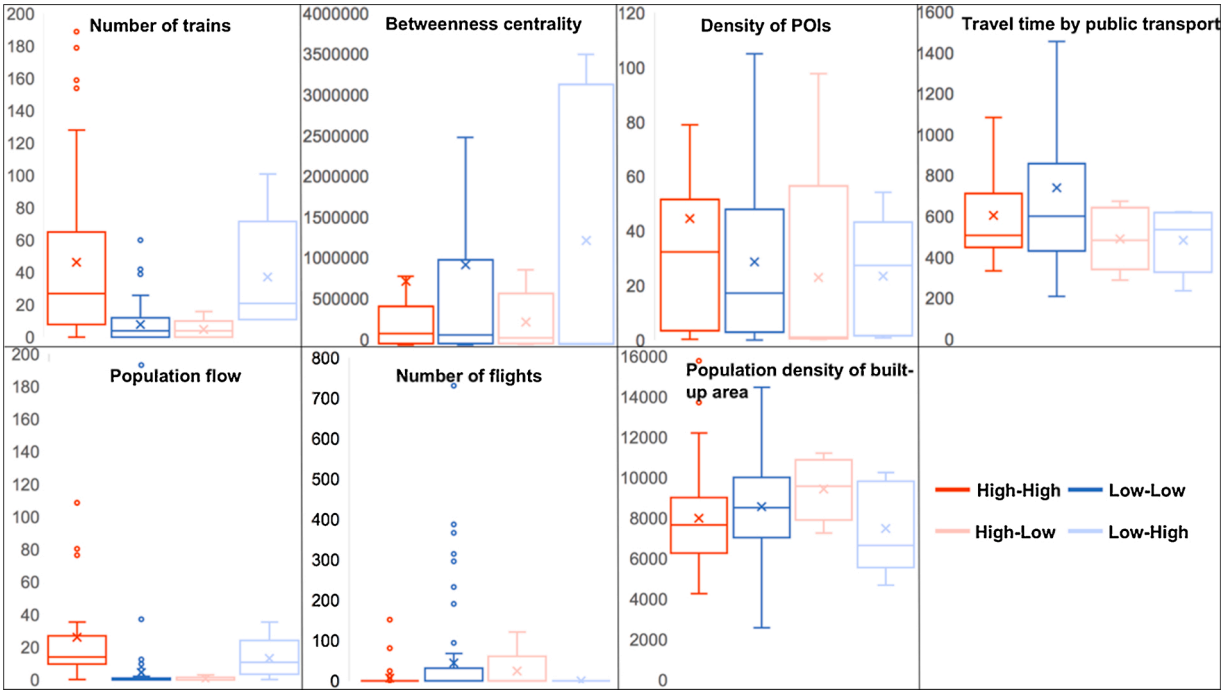



**Fig. 3.** The built environment attributes in four city types.

**Table 3**
Global Moran's I test results.

| Variable | Moran's index | Z-score | Sig. |
|---|---|---|---|
| Infection ratio | 0.596 | 25.990 | Local effects*** |
| Total number of trains | 0.343 | 16.185 | Local effects *** |
| Between-ness centrality | 0.212 | 9.822 | Local effects *** |
| POI density around railway stations | 0.029 | 1.457 | Global effects |
| Travel time by public transport | 0.016 | 0.919 | Global effects |
| Percentage of the aging population (over 65 years of age) | 0.420 | 18.300 | Local effects *** |
| Population flow | 0.191 | 8.982 | Local effects *** |
| Number of flights | 0.009 | 0.567 | Global effects |
| Population density of built-up areas | 0.174 | 7.680 | Local effects *** |

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

decided to use the mixed GWR model to evaluate the local and global effects of the BE attributes.

The standardized estimation results of the three regression models (multiple regression model (MRL), GWR and mixed GWR) are shown in Table 4, where the dependent variable is the infection ratio and the explanatory variables are the number of trains, betweenness centrality, density of the POIs around the railway stations, travel time by public transport, percentage of the aging population, population flow, number of flights, and the population density of built-up areas. The GWR and MLR models are used for comparison, where GWR is a special case of the mixed GWR and MLR is a special case of both GWR and mixed GWR. Here, all the variables are standardized before the estimation and their parameter values are, therefore, directly comparable. The VIF (variance inflation factor) values in Table 4 are all much smaller than 5.0, suggesting that none of the introduced explanatory variables suffers from the collinearity issues. The adjusted $R^2$s in the MLR, GWR and mixed GWR models are 0.282, 0.502 and 0.538, and the AICs in the three models are 792, 738, 710, respectively. Both the adjusted $R^2$ and AIC values suggest that the mixed GWR model outperforms the MLR and GWR models.

For the three variables with global effects the density of the POIs around the railway stations is positively associated with the infection ratio; however, the association is not statistically significant. The number of flights has a negative association with the infection ratio, differing to our expectations, and the association is also not significant. One reason for this negative association of flights may be that the airline coverage from Hubei Province to county-level cities in China is low. Among the 255 cities under study only 68 cities (26.7 %) have available flights from Hubei Province. There are a few prefecture-level cities, such as Xinyang and Nanyang, that are close to Wuhan and have high infection ratios, but have no airlines to Wuhan. The travel time by public transport is negatively related to the infection ratio and is significant at the 0.01 level. This means that the longer the travel time by public transport is from residences to the nearby activity centers, the higher the infection ratio will be.

Table 4 further summarizes the information of the coefficients estimated from the mixed GWR model, including the average, maximum, minimum and median values. By comparing all of the coefficients it was found that, on average, the number of trains shows the highest association (0.461) with the infection ratio, followed by the population flow (0.270) and percentage of the aging population (-0.165). The population flow has the highest positive value of the maximum coefficient, while the between-ness centrality shows the highest negative minimum coefficient. The mixed GWR estimation results indicate that all of the variables with local effects have both negative and positive estimates. Higher infection ratios are associated with more trains (departing from or stopping over in Hubei Province) in 98.04 % of the cities under study and higher between-ness centrality in 70.98 % of the cities. For the population flow and population density of built-up areas, the percentages of the negative and positive estimates are similar. With regard to the percentage of the aging population, there are more negative estimates (in 79.22 % of the cities) than positive parameters. As all of the coefficients with local effects were estimated with regard to each location, their spatial distributions are illustrated below.

**Table 4**
Standardized estimation results of the MLR, GWR and Mixed GWR models.

| Variables | MLR | | | GWR | | | | | Mixed GWR | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Coefficient | t-value | VIF | Coefficient | | | | Sig. at 10% | Coefficient | | | | Sig. at 10% |
| | | | | Average | Min | Median | Max | | Average | Min | Median | Max | |
| Total number of trains | 0.694 | 9.492*** | 1.091 | 0.492 | -0.067 [1.18%] | 0.524 | 1.091 [98.82%] | 193 [75.69%] | 0.461 | -0.094 [1.96%] | 0.433 | 1.180 [98.04%] | 175 [68.63%] |
| Between-ness centrality | 0.072 | 1.019 | 1.029 | -0.012 | -0.951 [25.10%] | 0.046 | 0.666 [74.90%] | 10 [3.92%] | -0.029 | -0.900 [29.02%] | 0.057 | 0.778 [70.98%] | 18 [7.06%] |
| POI density around railway stations | -0.012 | -0.169 | 1.111 | -0.014 | -0.472 [47.45%] | 0.023 | 0.176 [52.55%] | 22 [8.63%] | 0.019 | | | | |
| Travel time by public transport | -0.196 | -2.644*** | 1.120 | -0.239 | -0.739 [95.29%] | -0.15 6 | 0.062 [4.71%] | 73 [28.63%] | -0.136 | | | | *** |
| Percentage of the aging population | -0.133 | -1.790* | 1.122 | -0.147 | -0.593 [76.86%] | -0.12 7 | 0.214 [23.14%] | 81 [31.76%] | -0.165 | -0.677 [79.22%] | -0.143 | 0.279 [20.87%] | 84 [32.94%] |
| Population flow | 0.016 | 0.229 | 1.034 | 0.201 | -0.396 [49.80%] | 0.000 | 2.029 [50.20%] | 117 [45.88%] | 0.270 | -0.542 [45.10%] | 0.056 | 2.493 [54.90%] | 100 [39.22%] |
| Number of flights | -0.068 | -0.911 | 1.122 | -0.081 | -0.575 [70.98%] | -0.07 9 | 0.221 [29.02%] | 8 [3.14%] | -0.020 | | | | |
| Population density of built-up areas | -0.052 | -0.721 | 1.067 | 0.041 | -0.254 [39.22%] | 0.061 | 0.397 [60.78%] | 40 [25.69%] | 0.034 | -0.221 [46.67%] | 0.041 | 0.466 [53.33%] | 26 [10.20%] |
| Constant term | -1.852 | -26.440*** | | -1.671 | -2.558 [100%] | -1.78 8 | -0.470 [0%] | 130 [50.98%] | -1.634 | -2.560 [100%] | -1.798 | -0.156 [0%] | 247 [96.86%] |
| Sample size (cities) | 255 | | | 255 | | | | | 255 | | | | |
| Adjusted $R^2$ | 0.282 | | | 0.502 | | | | | 0.538 | | | | |
| AIC | 792 | | | 738 | | | | | 710 | | | | |
| Kernel function | | | | Gaussian | | | | | Gaussian | | | | |
| Bandwidth | | | | 102 | | | | | 83 | | | | |

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; the values in parentheses of the Min column indicate the shares of negative coefficients and those of the Max column refer to the shares of positive coefficients; MLR: Multiple Linear Regression; GWR: Geographically Weighted Regression; VIF: Variance Inflation Factor.

### *6.3. Spatial distributions of the associations of the BE attributes with the spread of COVID-19*

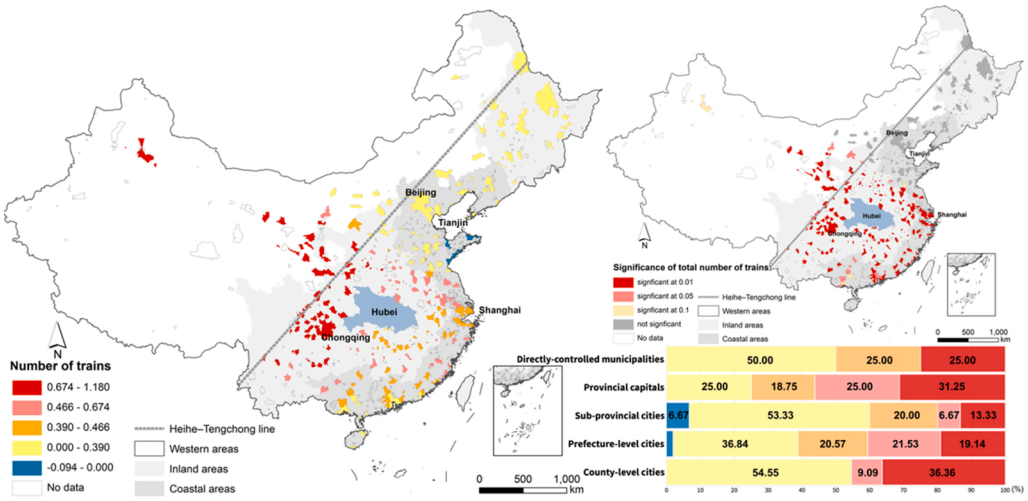
For the number of trains Fig. 4a shows that the coefficients are positive for most of the cities across the country and statistically significant at the 0.01 level, except for the northwestern cities. Higher coefficients appear in the cities surrounding Hubei Province. Negative coefficients are observed in only a handful of sub-provincial cities and prefecture-level cities, but are not statistically significant. There are further differences between the administrative types of cities. The coefficients are higher in county-level cities and 36.36 % of these cities show the highest coefficient range (between 0.674 and 1.180). These results may suggest that, for most cities in China, restricting the intercity connections, via railway transport, contributed to slowing down the spread of COVID-19, especially in the cities surrounding Hubei Province and county-level cities.

As shown in Fig. 4b, positive associations of the between-ness centrality with the spread of COVID-19 are found in 70.98 % of the cities under study, which are widely distributed in the whole of China. The cities with the highest coefficient range (between 0.100 and 0.778) are concentrated in the south and north of Hubei Province, but the coefficients are only statistically significant for most of the cities in the south. The negative impacts of between-ness centrality on the spread of COVID-19 are mainly observed in southwestern China and the Yangtze River delta (around Shanghai). Other coefficients with statistical significance appear along the Heihe-Tengchong line in the middle of China. Positive associations of between-ness centrality are found in half of the directly controlled municipalities and 62.5 % of the sub-provincial cities, but in more than 80 % of the county-level cities. These observations may imply that the control measures implemented at the railway stations with higher between-ness centrality, especially in the southern cities of Hubei Province, significantly contributed to preventing the spread of COVID-19.

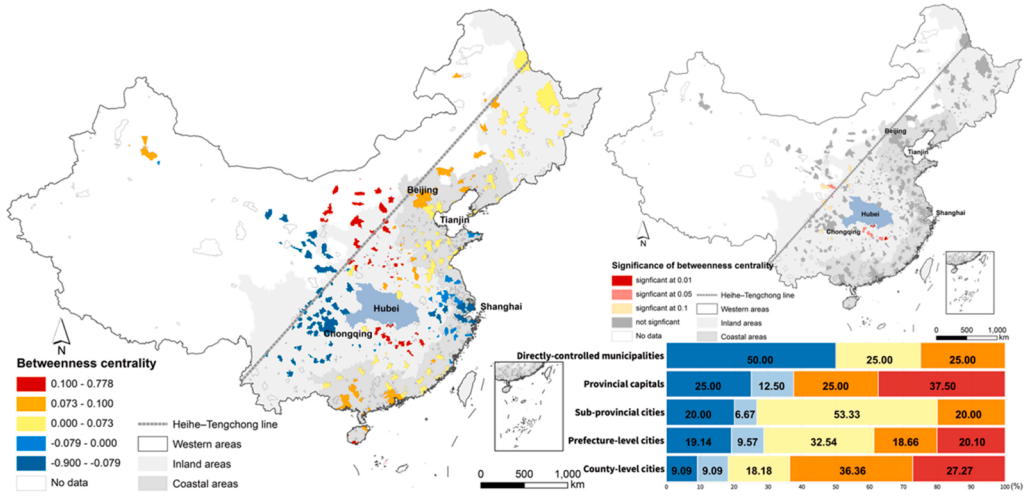
Fig. 4c illustrates the coefficient distributions of the percentage of the aging population. In the cities located at the north of Hubei Province and in southeastern China, the higher the percentage of the aging population is, the more serious the spread of COVID-19 will be. Negative associations are found in 79.22 % of the cities, with the coefficients in southwestern China and the Yangtze River delta (around Shanghai) being statistically significant. The above results suggest that younger people's careless behavior is more likely to spread the virus, while older people are more cautious about the virus. 25 % of the province-capital cities have positive coefficients that are the highest among all of the city levels. This means that, in these province-capital cities, older people may contribute to the spread of COIVD-19.

Fig. 4d shows that both the positive and negative coefficients of the population flow are observed to vary across locations and 39.22 % of them are statistically significant. The highest values of the positive coefficients (between 0.483 and 2.493) are mainly observed in northeast China, most of which are statistically significant. Negative associations appear in the cities surrounding Hubei Province. This may be because the population flow is greater in larger cities, where the medical conditions are better and the local governments are more serious in controlling the spread of COVID-19.

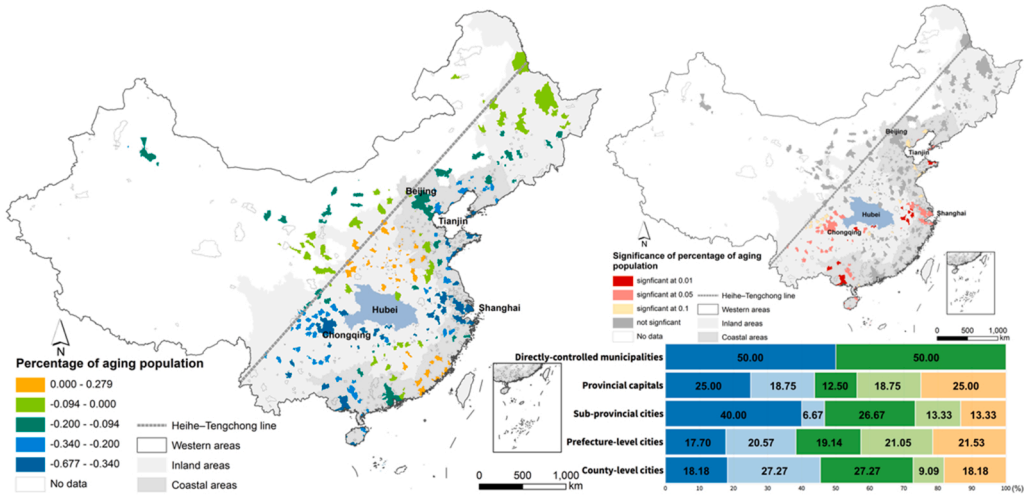
In 53.33 % of the cities under study (Fig. 4e) the population density of the built-up areas is estimated to be positively related to the infection ratio, while 46.67 % of the cities show an opposite relationship. Positive coefficients are mainly concentrated in the west, east and south of Hubei Province. Differences in the relationships are remarkable between inland and coastal areas. The coefficients are statistically significant for a few cities located in the south of Hubei Province. Focusing on the administrative type of cities, positive relationships are obvious in provincial capitals. Thus, for the cities at the south of Hubei Province reducing the population densities of the built-up areas could be effective to mitigate the impacts of future pandemics in China in their initial stage, especially in provincial capitals.




a. Parameters of number of trains



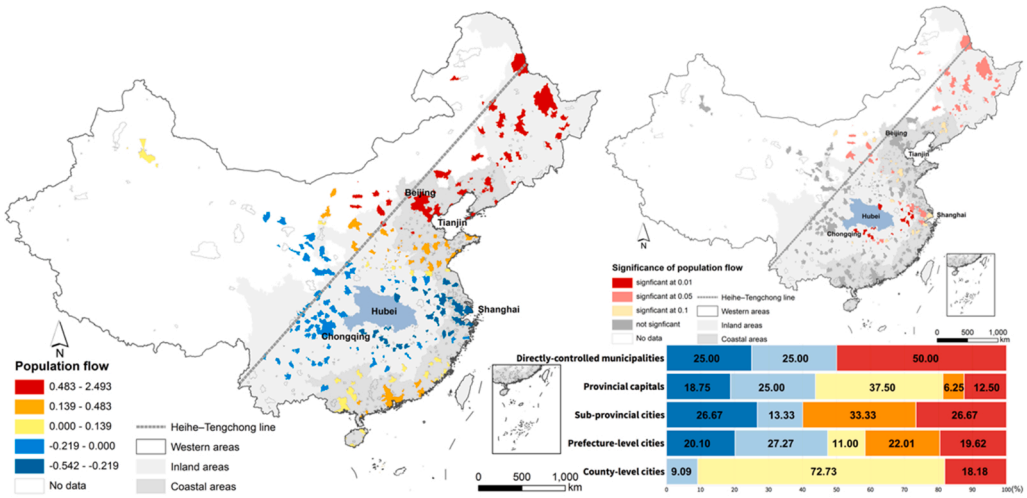

b. Parameters of betweenness centrality



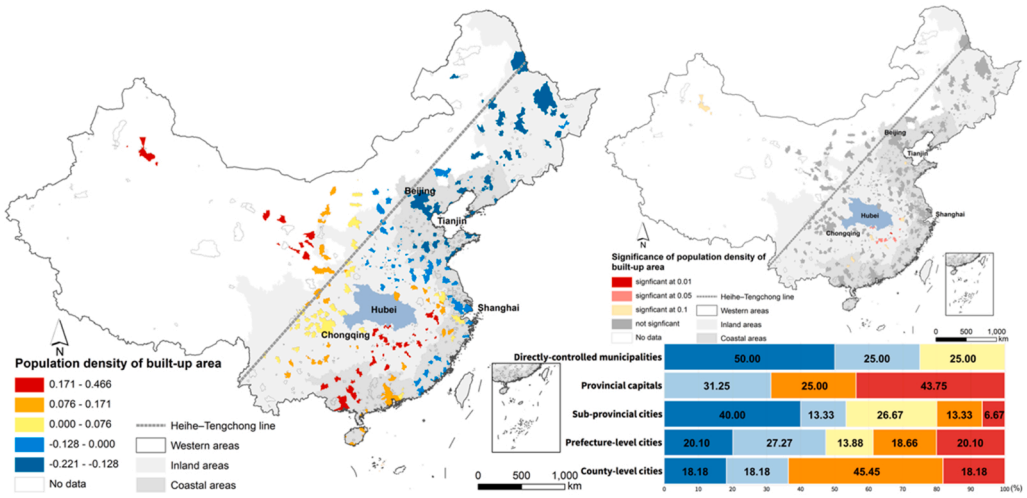

c. Parameters of percentage of aging population

**Fig. 4.** Spatial distributions of impacts of the built environment attributes on the spread of COVID-19..
a. Parameters of number of trains.
b. Parameters of betweenness centrality.
c. Parameters of percentage of aging population.
d. Parameters of population flow.
e. Parameters of population density of built-up area.

**Fig. 4.** (*continued*).




d. Parameters of population flow




e. Parameters of population density of built-up area

## 7. Discussion and conclusion

### 7.1. Discussion

Five BE attributes were identified to have local effects by the Global Moran's I test. On average the number of trains is mostly associated with the spread of COVID-19, followed by the population flow and percentage of the aging population. The coefficients for the number of trains and between-ness centrality are positive in most of the cities. In contrast, the coefficients for the percentage of the aging population are negative in most of the cities. For the population flow and population density of built-up areas the percentages of the negative and positive estimates are similar. On the whole, all of the five BE attributes have mixed effects on the spread of COVID-19, as captured by the location-specific coefficients. It is further observed that there are many spatial clusters of the spread of COVID-19 and these clusters vary with all of the seven BE attributes. For the global-effect coefficients the associations are positive for the density of the POIs around railway stations and negative for the number of flights and travel time by public transport; however, the association is significant for only the travel time by public transport.

The above findings have important policy implications. First, restricting the inter-city connections via railways is likely to prevent a further spread of COVID-19 in most cities of all levels. For sub-provincial cities and country-level cities control of the population flow is probably effective to suppress the spread of COVID-19. Lowering the between-ness centrality could slow down the spread of future pandemics in more than 70.00 % of prefectural-level cities, while reducing the population densities of built-up areas could be effective to prevent the spread of future pandemics in about 70.00 % of the province capitals. Effective countermeasures seem not to be the same across locations, showing geographical differences, as suggested by the coefficients of the BE attributes with local effects for directly controlled municipalities, provincial capitals and sub-provincial cities (most developed cities in China) (Supplementary Table 1). For instance, in Shanghai, only the coefficient of the number of trains is positive, implying that restricting the inter-city connections could be sufficient to control the spread of COVID-19. Reducing the travel time by public transport from residences to nearby activity centers could be significantly effective to control the spread of COVID-19.

### 7.2. Conclusion

This study empirically examined how the BE attributes were associated with the spread of COVID-19 in China in its initial stage across the whole country. We conducted a mixed GWR-based analysis by focusing on both the inter- and intra-city BE attributes, through which spatial

heterogeneities across locations are reflected. To the authors' best knowledge this study has presented one of the initial investigations about the above association in the field of urban and regional research linked with a public health pandemic.

The contributions of this study can be summarized as follows. First, we conducted the first study in literature on the whole of China by associating the BE attributes with the spread of COIVD-19. Second, we examined more intra-/inter-city BE attributes than all of the existing studies by focusing on the spread of COVID-19 in its initial stage. Third, we conducted a joint analysis of both the global and local effects of the BE attributes by estimating a mixed GWR model.

The introduced seven BE attributes indirectly reflect the intensity of activity participation and face-to-face contact. Although the GWR analysis does not capture true causalities between the BE attributes and the spread of COVID-19, the findings of this study have important practical implications. Infection clusters occurred not only at activity centers, but also in less dense areas. At the initial stage of the pandemic in China cutting the inter-city connections may have largely contributed to mitigating the spread of COVID-19, especially in the cities surrounding Hubei Province. The associations of the seven BE attributes differ remarkably in different geographical locations. This suggests that countermeasures against COVID-19 in China could have been implemented by considering such spatial heterogeneities, which may have the same effects as countrywide uniform measures. Countrywide uniform measures are not unique in China, as they can also be observed in other countries, such as India, Japan and some European countries. Implementing uniform measures across a whole country is easier; however, the social and economic costs are enormous. Our findings suggest that geographical differences should be better addressed in policy measures against COVID-19.

The current pandemic has raised important questions for national, regional and urban planning, such as how to make our nations, regions, cities and communities resilient to future public health pandemics, as argued by Zhang et al. (2021). The findings of this study have long-term urban planning and policymaking implications, for example the BE could be improved to mitigate the impacts of future public health pandemics, such as planning a more flexible transport network and city schedules, and rethinking about the location of landmarks and urban activity centers, as well as integrating the telecommuting work plans.

Having summarized the findings, we have to admit there are limitations of this study. First, due to the focus being on the initial stage of the spread of COIVD-19, we did not investigate shelter-in-place variables (such as the lockdown time in different cities). Wuhan was locked down first, followed by other Chinese cities, and the whole country was locked down before January 30, 2020. Our target period (January 20 – February 3, 2020) can reflect the spread of COVID-19 before the lockdown of the whole of China. Thus, excluding the shelter-in-place variables in this study has its own rationality. Nevertheless, in considering the various different situations of the spread of COVID-19 in other countries, the shelter-in-place variables should be paid sufficient attention in research of COVID-19. Second, there are some unintuitive associations between the BE attributes and the ratios of accumulative infection cases. To overcome this shortcoming analysis approaches incorporating the causalities between the dependent and independent variables should be developed. Third, more spatially fine data may be needed to measure the influence of the BE attributes. Fourth, as well as the BE attributes, other socio-demographic characteristics of households and healthcare facilities should be considered in future studies. Finally, it is important to make international comparisons to derive more scientifically sound evidence and, consequently, support more effective policy recommendations, as the virus cannot distinguish between people and between countries.

## 8. Role of funding source

This study was funded by two funding bodies, as shown below, but nobody in the funding bodies was involved in this study.

## Declaration of Competing Interest

No potential conflicts of interest were disclosed.

## Acknowledgement

This research was financially supported by Japan Science and Technology Agency (JST) (JST RISTEX Grant Number JPMJRX20J6 and JST J-RAPID Grant Number JPMJJR2006) and Japan Society for the Promotion of Science (Grants-in-Aid for Scientific Research (B), No. 18KT0007).

## Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.scs.2021.102752.

## References

Ahmed, M. Z., Ahmed, O., Zhou, A., Sang, H., Liu, S., & Almad, A. (2020). Epidemic of COVID-19 in China and associated psychological problems. *Asian Journal of Psychiatry, 51*, Article 102092.

Aleta, A., Hu, Q., Ye, J., Ji, P., & Moreno, Y. (2020). A data-driven assessment of early travel restrictions related to the spreading of the novel COVID-19 within mainland China. *Chaos, Solitons, and Fractals, 139*, Article 110068.

Altakarli, N. S. (2020). China's response to the COVID-19 outbreak: A model for epidemic preparedness and management. *Dubai Medical Journal, 3*, 44–49.

Anderson, L. (2019). *Planning the built environment.* The Routledge Press.

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis, 27*, 93–115.

Auger, K. A., Shah, S. S., Richardson, T., Hartley, D., Hall, M., Warniment, A., Timmons, K., Bosse, D., Ferris, S. A., Brady, P. W., Schondelmeyer, A. C., & Thomson, J. E. (2020). Association between statewide school closure and COVID-19 incidence and mortality in the US. *JAMA, 324*(9), 859–870.

Balcan, D., & Vespignani, A. (2011). Phase transitions in contagion processes mediated by recurrent mobility patterns. *Nature Physics, 7*, 581–586.

Balcan, D., & Vespignani, A. (2012). Invasion threshold in structured populations with recurrent mobility patterns. *Journal of Theoretical Biology, 293*, 87–100.

Basta, C., & Moroni, S. (2013). *Ethics, design and planning of the built environment.* The Springer Press.

Baym, N. K., Zhang, Y. B., & Lin, M. C. (2004). Social interactions across media: Interpersonal communication on the Internet, telephone and face-to-face. *New Media & Society, 6*(3), 299–318.

Brown, K. A., Jones, A., Daneman, N., Chan, A. K., Schwartz, K. L., Garber, G. E., … Stall, N. M. (2020). *Association between nursing home crowding and COVID-19 infection and mortality in Ontario, Canada.*. Preprint at medRxiv. https://doi.org/10.1101/2020.06.23.20137729.

Budds, D. (2020). *Design in the age of pandemics*. available at: https://www.curbed.com/2020/3/17/21178962/design-pandemics-coronavirus-quarantine (Accessed 27 March 2020.

Cao, X., Liu, Y., Li, T., & Liao, W. (2019). Analysis of spatial pattern evolution and influencing factors of regional land use efficiency in china based on ESDA-GWR. *Scientific Reports, 9*(1), 520.

Capolongo, S., Rebecchi, A., Buffoli, M., Letizia, A., & Carlo, S. (2020). COVID-19 and cities: From urban health strategies to the pandemic challenge. A decalogue of public health opportunities. *Acta Biomedica, 91*(2), 13–22. https://doi.org/10.23750/abm.v91i2.9515

CUSA CDC. (2020). *COVID-19 guidance for shared or congregate housing*. Centers for Disease Control and Prevention. available at : https://www.cdc.gov/coronavirus/2019-ncov/community/shared-congregate-house/guidance-shared-congregate-housing.html (Accessed 1 April 2020).

Chang, V. (2020). *The post-pandemic style*. available at: https://slate.com/business/2020/04/coronavirus-architecture-1918-flu-cholera-modernism.html (Accessed 28 April 2020).

Chen, Y., Liu, X., & Li, X. (2017). Analyzing parcel-level relationships between urban land expansion and activity changes by integrating Landsat and nighttime light data. *Remote Sensing, 9*(2), 164.

Cummings, D. A., Irizarry, R. A., Huang, N. E., Endy, T. P., Nisalak, A., Ungchusak, K., & Burke, D. S. (2004). Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand. *Nature, 427*, 344–347.

Dai, H., & Zhao, B. (2020). Association of the infection probability of COVID-19 with ventilation rates in confined spaces. *Building Simulation*. https://doi.org/10.1007/s12273-020-0703-5

Dearry, A. (2004). Impacts of our built environment on public health: Editorial. *Environmental Health Perspectives, 112*(11), 601.

Dietz, L., Horve, P., Coil, D., Fretz, M., Eisen, J., & Van Den Wymelenberg, K. (2020). 2019 novel coronavirus (COVID-19) pandemic: Built environment considerations to

reduce transmission. *MSystems, 5*(2). https://doi.org/10.1128/mSystems.00245-20. e00245-20.
Dong, L., Hu, S., & Gao, J. (2020). Discovering drugs to treat coronavirus disease 2019 (COVID-19). *Drug Discoveries & Therapeutics, 14*(1), 58–60.
Emeruwa, U. N., Ona, S., Shaman, J. L., Turitz, A., Wright, J. D., Gyamfi-Bannerman, C., & Melamed, A. (2020). Associations between built environment, neighborhood socioeconomic status, and SARS-CoV-2 infection among pregnant women in New York City. *JAMA, 324*(4), 390–392.
European Commission. (2020). *COVID-19: Guidelines on the progressive restoration of transport services and connectivity*. Brussels, Belgium: European Commission.
Eykelbosh, A. (2020). *Physical barriers for COVID-19 infection prevention and control in commercial settings [blog]*. Vancouver, BC: National Collaborating Center for Environmental Health. available at: https://ncceh.ca/content/blog/physical-barriers-covid-19-infection-prevention-and- control-commercial-settings. (Accessed 13 May 2020).
Ferguson, N. M., Cummings, D. A., Cauchemez, S., Fraser, C., Riley, S., Meeyai, A. A., Iamsirithaworn, S., & Burke, D. S. (2005). Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature, 437*, 209–214.
Gan, N., Thomas, N., & Culver, D. (2020). *CNN. Over 1,700 frontline medics infected with coronavirus in China, presenting new crisis for the government*. available at: https://edition.cnn.com/2020/02/13/asia/coronavirus-health-care-workers-infected-intl-hnk/index.html. (Accessed 24 February 2020).
Gao, J., Tian, Z., & Yang, X. (2020). Breakthrough: Chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies. *Bioscience Trends, 14*(1), 72–73.
Germann, T. C., Kadau, K., Longini, I. M., & Macken, C. (2006). Mitigation strategies for pandemic influenza in the United States. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 5935–5940.
Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis, 24*(3), 189–206.
Ghosh, A., Nundy, S., Ghosh, S., & Mallick, T. K. (2020). Study of COVID-19 pandemic in London (UK) from urban context. *Cities, 106*, Article 102928.
Giani, P., Castruccio, S., Anav, A., Howard, D., Hu, W., & Crippa, P. (2020). Short-term and long-term health impacts of air pollution reductions from COVID-19 lockdowns in China and Europe: A modelling study. *The Lancet Planetary Health, 4*(10). e474-e482.
Gog, J. R., Ballesteros, S., Viboud, C., Simonsen, L., Bjornstad, O. N., Shaman, J., Chao, D. L., Khan, F., & Grenfell, B. T. (2014). Spatial transmission of 2009 pandemic influenza in the US. *PLoS Computational Biology, 10*, Article e1003635.
Gonzalez, M. C., Hidalgo, C. A., & Barabasi, A. L. (2008). Understanding individual human mobility patterns. *Nature, 453*, 779–782.
Hajrasouliha, A. H., & Hamidi, S. (2017). The typology of the American metropolis: monocentricity, polycentricity, or generalized dispersion? *Urban Geography, 38*(3), 420–444.
Hamidi, S., Ewing, R., & Sabouri, S. (2020a). Longitudinal analyses of the relationship between development density and the COVID-19 morbidity and mortality rates: Early evidence from 1,165 metropolitan counties in the U.S. *Health & Place, 64*, Article 102378.
Hamidi, S., Sabouri, S., & Ewing, R. (2020b). Does density aggravate the COVID-19 pandemic? Early findings and lessons for planners. *Journal of the American Planning Association, 86*, 1–15.
Han, Y., & Yang, H. (2020). The transmission and diagnosis of 2019 novel coronavirus infection disease (COVID-19): A Chinese perspective. *Journal of Medical Virology, 92*, 639–644.
Handy, S. L., Boarnet, M. G., Ewing, R., & Killingsworth, R. E. (2002). How the built environment affects physical activity: Views from urban planning. *American Journal of Preventive Medicine, 23*(2S), 64–73.
Hasnain, M., Pasha, M. F., & Ghani, I. (2020). Combined measures to control the COVID-19 pandemic in Wuhan, Hubei, China: A narrative review. *Journal of Biosafety and Biosecurity, 2*(2), 51–57.
Hendricks, B., & Mark-Carew, M. (2017). Using exploratory data analysis to identify and predict patterns of human Lyme disease case clustering within a multistate region, 2010-2014. *Spatial and Spatio-temporal Epidemiology, 20*, 35–43.
Hogbin, V. (1985). Railways, disease and health in South Africa. *Social Science & Medicine, 9*(20), 933–938.
Jia, J. S., Lu, X., Yuan, Y., Xu, G., Jia, J., & Christakis, N. A. (2020). Population flow drives spatio-temporal distribution of COVID-19 in China. *Nature, 582*, 389–394.
Kang, D. W., Kim, M., Cho, D., & Lee, S. (2010). The effects of urban development pressure on agricultural land price: Application of a mixed GWR model. *Journal of Rural Development, 33*(4), 63–83.
Kim, N. (2020). 'It's ruining everyone': Eerie quiet reigns in coronavirus-hit south Korean city. *The guardian*. available at: https://www.theguardian.com/world/2020/feb/23/its-ruining-everyone-coronavirus-hit-south-korean-city-daegu. (Accessed 24 February 2020).
Kuddus, M. A., Rahman, A., Talukder, M. R., & Hoque, A. (2014). A modified SIR model to study on physical behavior among smallpox infective population in Bangladesh. *American Journal of Mathematics and Statistics, 4*(5), 231–239.
Lau, H., Khosrawipour, V., Kocbach, P., Mikolajczyk, A., Ichii, H., Zacharski, M., Bania, J., & Khosrawipour, T. (2020). The association between international and domestic air traffic and the coronavirus (COVID-19) outbreak. *Journal of Microbiology Immunology and Infection, 53*(3), 467–472.
Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R., … Lessler, J. (2020). The incubation period of Coronavirus Disease 2019 (COVID-19) from publicly reported confirmed cases: Estimation and application. *Annals of Internal Medicine, 172*, 577–582.
Lee, H., Park, S. J., Lee, G. R., Kim, J. E., Lee, J. H., Jung, Y., & Nam, E. W. (2020). The relationship between trends in COVID-19 prevalence and traffic levels in South Korea. *International Journal of Infectious Diseases, 96*, 399–407.
Lin, G., Chen, X., & Liang, Y. (2018). The location of retail stores and street centrality in Guangzhou, China. *Applied Geography, 100*, 12–20.
Long, Y., & Liu, L. (2017). How green are the streets? An analysis for central areas of Chinese cities using Tencent Street View. *PloS One, 12*(2), 1–18.
Mccluskey, C. C. (2010). Complete global stability for an SIR epidemic model with delay - distributed or discrete. *Nonlinear Analysis: Real World Application, 11*(1), 55–59.
Megaheda, N. A., & Ghoneimb, E. M. (2020). Antivirus-built environment: Lessons learned from COVID-19 pandemic. *Sustainable Cities and Society, 61*, Article 102350.
Mizumoto, K., & Chowell, G. (2020). Transmission potential of the novel coronavirus (COVID-19) onboard the diamond Princess Cruises Ship. *Infectious Disease Modelling, 5*, 264–270.
Mousavinia, S., Pourdeihimi, S., & Madani, R. (2019). Housing layout, perceived density and social interactions in gated communities: Mediational role of territoriality. *Sustainable Cities and Society, 51*, Article 101699.
Murphy, N., Boland, M., Bambury, N., Fitzgerald, M., Comerford, L., Dever, N., … O'Connor, L. (2020). A large national outbreak of COVID-19 linked to air travel, Ireland, summer 2020. *Euro Surveillance : Bulletin Europeen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin, 25*(42), Article 2001624.
Nguyen, Q. C., Huang, Y., Kumar, A., Duan, H., Keralis, J. M., Dwivedi, P., … Tasdizen, T. (2020). Using 164 million Google Street View images to derive built environment predictors of COVID-19 cases. *International Journal of Environmental Research and Public Health, 17*, 6359.
Park, K., Ewing, R., Sabouri, S., Choi, D., Hamidi, S., & Tian, G. (2020). Guidelines for a polycentric region to reduce vehicle use and increase walking and transit use. *Journal of the American Planning Association, 86*(2), 236–249.
Peng, L., Yang, W., Zhang, D., Zhuge, C., & Hong, L. (2020). Epidemic analysis of COVID-19 in China by dynamical modeling. *Preprint at medRxiv*. https://doi.org/10.1101/2020.02.16.20023465
Pinheiro, M. D., & Luís, N. C. (2020). COVID-19 could leverage a sustainable built environment. *Sustainability, 12*(14), 5863. https://doi.org/10.3390/su12145863
Poletto, C., Gomes, M. F., Piontti, A., Rossi, L., Bioglio, L., Chao ref_ellipsis, D. L., Jr., & Vespignani, A. (2014). Assessing the impact of travel restrictions on international spread of the 2014 West African Ebola epidemic. *Eurosurveillance, 19*(42), 20936.
Raj, V. A. A., Velraj, R., & Haghighat, F. (2020). The contribution of dry indoor built environment on the spread of Coronavirus: Data from various Indian states. *Sustainable Cities and Society, 62*, Article 102371. https://doi.org/10.1016/j.scs.2020.102371
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.
Rath, S., Tripathy, A., & Tripathy, A. R. (2020). Prediction of new active cases of coronavirus disease (COVID-19) pandemic using multiple linear regression model. *Diabetes and Metabolic Syndrome Clinical Research and Reviews, 14*(5), 1467–1474.
Rothan, H. A., & Byrareddy, S. N. (2020). The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *Journal of Autoimmunity, 109*, Article 102433.
Ruan, Z. Y., Wang, C. Q., Hui, M. P., & Liu, Z. H. (2015). Integrated travel network model for studying epidemics: Interplay between journeys and epidemic. *Scientific Reports, 5*, 11401.
Saadat, S., Rawtani, D., & Hussain, C. (2020). Environmental perspective of COVID-19. *The Science of the Total Environment, 728*, Article 138870.
Salathe, M., & Jones, J. H. (2010). Dynamics and control of diseases in networks with community structure. *PLoS Computational Biology, 6*(4), Article e1000736.
Schindler, S., Jepson, N., & Cui, W. (2020). Covid-19, China and the future of global development. *Research in Globalization, 2*, Article 100020.
Talen, E., & Anselin, L. (1998). Assessing spatial equity: An evaluation of measures of accessibility to public playgrounds. *Environment and Planning A: Economy and Space., 30*(4), 595–613.
Tobler, W. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography, 46*, 234–240.
Wang, J. Y., Wang, X. J., & Wu, J. J. (2018). Inferring metapopulation propagation network for intra-city epidemic control and prevention. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 830–838.
Wesolowski, A., Buckee, C. O., Engo-Monsen, K., & Metcaalf, C. J. E. (2016). Connecting mobility to infectious diseases: The promise and limits of mobile phone data. *The Journal of Infectious Diseases, 214*(4), 414–420.
World Health Organization. (2020). *Transmission of SARS-CoV-2: Implications for infection prevention precautions*. available at: https://www.who.int/news-room/commentaries/detail/transmission-of-sars-cov-2-implications-for-infection-prevention-precautions. (Accessed 9 July 2020).
Xu, W., Wu, J., & Cao, L. (2020). COVID-19 pandemic in China: Context, experience and lessons. *Health Policy and Technology, 9*(4), 639–648.
Yang, Y., Atkinson, P., & Ettema, D. (2008). Individual space-time activity-based modelling of infectious disease transmission within a city. *Journal of the Royal Society, Interface, 5*(24), 759–772.
Yashima, K., & Sasaki, A. (2014). Epidemic process over the commute network in a metropolitan area. *PloS One, 9*(6). No. e98518.
Yu, H., & Peng, Z. R. (2019). Exploring the spatial variation of ridesourcing demand and its relationship to built environment and socioeconomic factors with the

geographically weighted Poisson regression. *Journal of Transport Geography, 75*, 147–163.

Yu, P., Zhu, J., Zhang, Z., & Han, Y. (2020). A familial cluster of infection associated with the 2019 Novel Coronavirus indicating possible person-to-person transmission during the incubation period. *The Journal of Infectious Diseases, 221*(11), 1757–1761.

Zhang, J., Hayashi, Y., & Frank, L. D. (2021). COVID -19 and transport: Findings from a world-wide expert survey. *Transport Policy*. https://doi.org/10.1016/j.tranpol.2021.01.011. Available online 27 January 2021.