Uncertainties in the Assessment of COVID-19 Risk: A Study of People's Exposure to High-Risk Environments Using Individual-Level Activity Data

Jianwei Huang & Mei-Po Kwan

Uncertainties in the Assessment of COVID-19 Risk: A Study of People's Exposure to High-Risk Environments Using Individual-Level Activity Data

Jianwei Huang[*] and Mei-Po Kwan[*,†]

[*]*Institute of Space and Earth Information Science, The Chinese University of Hong Kong, China*
[†]*Department of Geography and Resource Management, The Chinese University of Hong Kong, China*

Based on different conceptualizations and measures of individual-level environmental exposure, this study examines how the uncertain geographic context problem (UGCoP) and the neighborhood effect averaging problem (NEAP) might affect the assessment of COVID-19 risk. Using the COVID-19 data on an open-access government Web site and the individual-level activity data of sixty confirmed COVID-19 cases (infected persons) in Hong Kong, we first represent COVID-19 risk environments using case-based and venues-based high-risk locations. The COVID-19 risk of each of the sixty selected cases is then evaluated by three approaches based on their exposures to the case-based or venues-based risk environments: the mobility-based approach, the residence-based approach, and the activity space–based approach. The results indicate that the UGCoP and the NEAP exist in the assessment of COVID-19 risk, which has significant implications: Ecological COVID-19 studies need to address the uncertainties due to the UGCoP and the NEAP by considering people's daily mobility. Otherwise, ignoring peoples' daily mobility and its interactions with complex and dynamic COVID-19 risk environments could lead to misleading results and misinform government nonpharmaceutical intervention measures. *Key Words: COVID-19, environmental exposure, neighborhood effect averaging problem, individual-level health risk analysis, uncertain geographic context problem.*

The COVID-19 pandemic has been the most serious threat to global public health in more than 100 years, with more than 126.3 million confirmed cases and 2.7 million confirmed deaths worldwide by early May 2021 (World Health Organization 2021). The pandemic has led scholars to investigate the environment's role in the incidence and mortality of the disease (e.g., Hamidi, Sabouri, and Ewing 2020; Huang et al. 2020; Kan et al. 2021; Kwok et al. 2021). Numerous studies have linked the incidence or mortality of COVID-19 to specific physical and social environmental contexts using public COVID-19 and environmental data that are spatially aggregated to fixed administrative units (e.g., counties or census tracts; Coccia 2020; Das et al. 2020; Desjardins, Hohl, and Delmelle 2020; Kodera, Rashed, and Hirata 2020; Tian et al. 2020). Using aggregated space–time data, however, is susceptible to the modifiable spatiotemporal unit problem (Cheng and Adepeju 2014; Kwan 2018a), which means that research findings from ecological studies might be different due to the use of areal units of different spatial scales (e.g., counties or census tracts) and different temporal scales (e.g., weeks or months). For instance, Huang et al. (2020) and Amram et al. (2020) concluded that urban density is an important factor affecting the spatial incidence rate of COVID-19 in Hong Kong and Washington State (United States), and Hamidi, Sabouri, and Ewing (2020) found that metropolitan size plays a more important role than urban density in the COVID-19 transmission in the United States. The differences in these results might be due to the use of areal units of different spatial scales: The former two studies used ZIP code (Washington State) and Tertiary Planning Unit (Hong Kong) level data, whereas Hamidi, Sabouri, and Ewing (2020) used county-level data.

To address the modifiable spatiotemporal unit problem in ecological COVID-19 studies, Helbich, Browning, and Kwan (2021) recently urged researchers to apply retrospective case–control and conduct high-resolution, individual-centered studies to investigate the environmental determinants of disease transmission based on registered residential address data and location tracking technologies. As the authors argued, one of the critical tasks for these studies is to accurately assess how and to what

extent infected people were exposed to the coronavirus and various environmental factors or contexts in their recent daily life. Assessment of environmental exposure in most studies on health–environment relationships to date, however, tends to use conventional delineations of people's residential neighborhoods (Kwan 2013). These studies tend to link people's health to their residential contexts, assuming that the residential neighborhood is the most relevant area affecting people's health (Kwan 2018a). For example, studies have found that air pollution (Fattorini and Regoli 2020; Wu et al. 2020; Yu et al. 2021), sociodemographics (Y. Xiong et al. 2020), neighborhood disadvantage (Durfey et al. 2019; Alkhamis et al. 2020), temperature (Azuma et al. 2020), and the food environment (Holsten 2009) are associated with different health outcomes (e.g., spatial incidence and mortality of COVID-19, obesity, and chronic diseases) using data based on people's residential neighborhoods.

Although previous studies provide a useful foundation for understanding the possible causal pathways through which environmental context might affect health, their conclusions could be misleading due to methodological issues. Because most people travel to areas outside of their residential neighborhoods in their daily lives for various activities, they are exposed to different neighborhood contexts (Kwan 2012). Hence, ignoring people's daily mobility in environmental exposure assessment could lead to unreliable results (Park and Kwan 2017; Kim and Kwan 2019). Among these methodological issues is the *uncertain geographic context problem* (UGCoP), which stresses that inferences about the effects of environmental exposure on people's health outcomes might be different due to how contextual areas are geographically delineated (Kwan 2012; Kwan et al. 2019).

To address the UGCoP, recent studies have investigated the effects of environmental exposure (e.g., air pollution, green space, and the food environment) on health outcomes based on fine-grained mobility and environmental data (Zhang et al. 2018; Zhao et al. 2018; Kou, Kwan, and Chai 2020; J. Ma et al. 2020). A particular phenomenon has been observed in these studies: People's mobility-based environmental exposures tend to converge to the average exposure value in the study area. This phenomenon has been articulated as the *neighborhood effect averaging problem* (NEAP), which suggests that

residence-based exposure assessments might underestimate or overestimate people's environmental exposure if it overlooks human daily mobility (Kwan 2018b; Kim and Kwan 2021a, 2021b, 2021c). Specifically, people who have higher residence-based exposures might have lower mobility-based exposures and vice versa, and the distribution of mobility-based exposures would deviate less from the mean than that of residence-based exposures. The NEAP has been regarded as "an elusive confounder of the neighborhood effect" (Kwan 2018b, 1). Further, the NEAP not only suggests that residence-based exposures might be different from mobility-based exposures but also identifies neighborhood effect averaging as a specific source of the UGCoP that leads to such differences (Kim and Kwan 2021b). Recent studies have identified the impacts of the NEAP on people's exposure to air pollution, traffic congestion, and different ethnic groups (Kim and Kwan 2019, 2021a, 2021b; X. Ma et al. 2020; Tan, Kwan, and Chen 2020). Very few studies to date, however, have examined how the UGCoP and the NEAP might influence the results of ecological COVID-19 studies when human mobility is not considered.

It is well known that the spread of infectious disease is influenced by human mobility and people's interactions in space and time (Viboud et al. 2006; Stoddard et al. 2009; Bian et al. 2012; Wesolowski et al. 2012; Lai et al. 2019; Li et al. 2019). As they move around to undertake their daily activities, susceptible individuals could be infected via direct contact with infected individuals. Overlapped activity spaces of individuals can thus contribute to the spread of infectious disease. For instance, infected individuals might leave the virus on certain surfaces (e.g., door handles, elevator buttons, and tableware) in the venues they visited or in which they stayed (e.g., restaurants or hotels). Contacting or touching these surfaces greatly increases the infection risk of susceptible individuals. Further, an infectious disease can spread through people's movements via the public transportation system and social interactions. In the case of COVID-19, transmission occurs mainly through people's face-to-face interactions via respiratory droplets produced by an infected person's mouth or nose when they cough, sneeze, speak, sing, or breathe heavily. Hence, people could catch COVID-19 if they have been in close contact with an infected person in specific locations (e.g.,

restaurants, bars, fitness facilities, and workplaces; Chang et al. 2021; Harapan et al. 2020). Thus, in this study, we conceptualize COVID-19 risk as the likelihood of a person contracting COVID-19 through contacts with infected individuals based on the notion of environmental exposure. In this framework, an individual with a higher COVID-19 risk means that the individual has a higher level of exposure to infected people or venues visited by infected people and thus will have a higher likelihood of contracting COVID-19. The individual is considered more vulnerable in the pandemic, and areas with higher concentrations of infected people or venues visited by infected people are high-risk locations.

Further, as recent studies have found, people's COVID-19 risk has a positive association with the duration and frequency of their exposure to high-risk locations. For instance, Baker, Peckham, and Seixas (2020) and St-Denis (2020) assessed the COVID-19 risk of workers with different occupations and socioeconomic characteristics based on how frequently these workers are exposed to infections in various types of workplaces in the United States and Canada. These studies found that workers with high-risk occupations (e.g., health care–related and low-income occupations) are more frequently exposed to COVID-19 risk in their workplaces and have significantly higher average risks of exposure than other occupational groups. Similarly, people with certain high-risk behaviors or activity patterns (e.g., frequently visiting high-risk venues like bars or fitness facilities) also tend to have a higher risk of exposure than other groups. Thus, venues recently visited by confirmed cases (i.e., where transmission of COVID-19 might occur) are venues with a high risk of COVID-19 transmission for a certain period, and people's exposure to these venues in their daily lives can be used to evaluate their COVID-19 risk.

Previous studies have used individual-based and spatially explicit epidemiological frameworks to evaluate individual risk to infectious diseases. For instance, Bian (2004) and Bian et al. (2012) used an individual-based simulation approach that incorporates four important considerations: individuals are different, individuals interact with each other locally, individuals are mobile, and the environment for individuals is heterogeneous. Wesolowski et al. (2012) used spatially explicit mobile phone data to identify the movement patterns and travel networks of infected persons who drive the spread of malaria in Kenya. These studies focused mainly on examining how human mobility affects the spread of infectious diseases. Since the onset of the COVID-19 pandemic, many studies have used mobile phone data to examine how nonpharmaceutical control measures (e.g., travel restrictions and stay-at-home orders) affect human mobility and the effectiveness of these measures in mitigating the spread of COVID-19 (e.g., Badr et al. 2020; Gao et al. 2020; Kraemer et al. 2020; Pullano et al. 2020; C. Xiong et al. 2020; Chang et al. 2021; Kim and Kwan 2021c; Lee, Qian, and Schwanen 2021; Willberg et al. 2021). In general, these studies observed that nonpharmaceutical control measures tend to reduce people's mobility and are effective in mitigating the spread of COVID-19. There are considerable disparities, however, between different social groups (with respect to income, race, ethnicity, and class) in their mobility change and ability to limit exposure to infected persons and high-risk places. For instance, Chang et al. (2021) observed higher infection rates among disadvantaged racial and socioeconomic groups because members of these groups cannot considerably reduce their mobility and the places and venues they visit are more crowded and riskier. Kim and Kwan (2021c) found that, perhaps due to quarantine fatigue, restricting people's mobility to control the pandemic is effective only for a short period, and low-income people keep traveling during the COVID-19 pandemic because they are mostly essential workers who are required to be physically present at their workplaces.

Although these studies represent significant advances in the study of individual risk of contracting infectious diseases, including COVID-19, they did not examine how different methods for assessing individual exposure risk might affect the results due to the UGCoP and the NEAP (i.e., contextual uncertainties), which were discovered in recent years. Specifically, they did not compare risk levels based on residence-based assessments with those obtained using mobility-based assessments. Thus, the goal of this article is to provide a methodological investigation of the UGCoP and the NEAP when assessing individual COVID-19 exposure risk. Further, although the UGCoP and NEAP are now established concepts, the limited number of studies on the NEAP to date have examined it only in individual exposure to air pollution, traffic congestion,

and ethnic groups. As a result, we currently have very limited knowledge about whether other kinds of exposures would face the NEAP and what social groups are most affected in those situations (Kim and Kwan 2021b). Thus, more empirical evidence is needed to expand our knowledge on how the NEAP might affect the experiences of different social groups for other kinds of environmental exposures and what mechanisms underly those exposures. The study thus seeks to increase our empirical knowledge about the NEAP (in addition to the UGCoP).

This article thus seeks to bridge this significant knowledge gap by examining how contextual uncertainties due to the UGCoP and the NEAP might affect the assessment of COVID-19 risk using three approaches: (1) a mobility-based approach, (2) a residence-based approach, and (3) an activity space–based approach. Specifically, we use the individual-level activity data of sixty infected persons in Hong Kong and different measures based on the three approaches to assess people's COVID-19 risk, which in turn is based on people's exposure to two types of high-risk locations: case-based high-risk locations (CHLs) and venues-based high-risk locations (VHLs). We compare the risk levels obtained with different combinations of the COVID-19 risk measures and high-risk locations. The results indicate that the UGCoP and the NEAP exist in the assessment of COVID-19 risk, which has significant implications for the ecological study of COVID-19 and other infectious diseases. Finally, we use regression models to evaluate the associations between the CHL-based and VHL-based COVID-19 risk obtained with different risk measures based on the three approaches. The results reveal the highest association between CHL-based and VHL-based COVID-19 risk when using the mobility-based approach, which suggests that using a mobility-based approach to assess people's COVID-19 risk can generate more reliable results when compared to other methods. The results imply that ecological studies of COVID-19 and other infectious diseases need to address the uncertainties due to the UGCoP and the NEAP by considering peoples' daily mobility. The study advances our understanding of how the UGCoP and the NEAP might affect the assessment of individual-based exposure risk to COVID-19 and other infectious diseases based on an environmental exposure framework.

Data and Method

This study seeks to examine the uncertainties in the assessment of COVID-19 risk. We first compile the individual-level activity data and construct the space–time trajectories of sixty selected infected persons. Second, the COVID-19 risk environments are represented based on CHLs and VHLs using kernel density surfaces, space–time cubes, and space–time kernel density estimation (STKDE). Third, six widely used activity space methods are applied to delineate the exposure spaces of the sixty infected persons. Then, three approaches are used to assess their COVID-19 risk. Finally, we examine the effects of the UGCoP and the NEAP on the assessments of COVID-19 risk. These methods and analytical tasks are described in detail in this section, and the results obtained are reported later.

Data Collection and Preprocessing

The study area for this research is Hong Kong, where the government conducts careful contact tracing of all confirmed COVID-19 cases and provides much of this information via an open-access Web site (this information is readily available to other researchers for replicating the study). Such detailed information allows us to examine COVID-19 risk at a high spatial and temporal resolution (e.g., at the building and venue level) that few other study areas can offer. The individual-level activity data of the infected persons used in this study cover the period from 27 January to 14 April 2020 (i.e., the first and second waves of the COVID-19 pandemic in Hong Kong). They came from two public sources: the Hong Kong Department of Health and online news articles. Table 1 shows the data items in the data set. These items include the case number, activity location (i.e., geographic coordinates), type of buildings visited, visit date, activity type, the start time of the activity, the end time of the activity, and the duration of the activity. Note that the publicly available data set from the government does not contain any information that reveals the identity of the infected persons. To protect personal privacy, the data were deidentified by the government by replacing the names of the infected persons with a case number (e.g., 1, 2, 3 …). Although the names of the residential buildings in which the infected persons live are disclosed in the data, these building

Table 1. Examples of an infected person's activities

| Case no.	Visited date	Start time	End time	Building type	Duration	Activity	Report date	x	y
1	13/3/2020	9:00	18:00	Office	540 minutes	Work	20/3/2020	828,665.46	838,395.21
1	13/3/2020	19:00	20:30	Commercial	90 minutes	Dinner	20/3/2020	815,836.27	834,010.28
1	13/3/2020	20:50	21:30	Bar	40 minutess	Entertainment	20/3/2020	815,913.55	833,856.76

names do not allow the reidentification of the infected persons in any straightforward manner because all of the residential buildings involved are multistory or high-rise structures with many residents. Despite the deidentified nature of the data, we have adhered to human subjects protection guidelines when conducting this research and preparing the article. Also, the institutional review board of the authors' university does not require ethics review or approval for the project because the data set is publicly available. The steps used to construct the data set are described as follows.

We first downloaded the activity data of the infected persons reported from 27 January to 14 April 2020 from the Hong Kong government's open-data Web site, which provides the data freely (see https://data.gov.hk). The data include 419 local cases and 591 imported cases. The imported cases are excluded from this study because these cases were infected before they entered Hong Kong and we cannot assess the many factors in the source countries or cities that had influenced the COVID-19 risk relevant to them. The data of the local cases contain the following information: the number of confirmed cases, some demographic data for each case (e.g., age and gender), the name of the buildings visited (without detailed activity information) by the cases during the incubation period (i.e., their visits to these venues in the past fourteen days before they were confirmed to be infected), and the name of the building in which they live. Note that the data only contain the names of the buildings visited without geographic coordinates. Hence, the geographic coordinates of each building visited by the confirmed cases are obtained by using the Google Place application programming interface (API). These buildings visited by infected persons (i.e., all local cases) were used to represent one of the high-risk environments in the study: CHLs (i.e., locations with high COVID-19 risk due to visits by people who were later confirmed to be infected).

Second, because local online media frequently reported the details of new COVID-19 cases in Hong Kong, news articles from these sources were used in the study to enrich the data obtained from the government's open-data Web site. These sources include *South China Morning Post* (see www.scmp.com), Mingpao (see news.mingpao.com), Hong Kong 01 (see www.hk01.com), and Radio Television Hong Kong (see news.rthk.hk). Using and cross-checking the information in these articles, we derived the activity locations, activity types (e.g., shopping, dining, or working), activity durations, and other attributes of the confirmed cases. For example, the data from the Hong Kong government's open-data Web site indicate which buildings were visited by Case 1, and the news articles further disclose when and what kind of activities the case had undertaken in the building. The duration each confirmed case stayed at specific locations can thus be roughly estimated based on the types of activities undertaken. Note that although there is some uncertainty in these activity duration estimates, we can still use these data for a methodological investigation because the study compares the individual-based risk levels obtained by different exposure measures based on the same COVID-19 risk environment. The differences in the results among these measures are unlikely to be significantly affected because such differences are influenced more by how different measures represent or capture the risk environment than by the estimation errors in the activity durations, given the same risk environment. The same can also be said concerning how the results might be affected by how busy different high-risk venues are. Although proxy measures (e.g., the square footage of high-risk venues) can be used to estimate how busy they are, the differences in the results among the exposure measures are unlikely to be significantly affected because such differences are influenced more by how different measures represent or capture the risk environment than by how busy the venues are, given the same risk environment.

Using these two public data sources, we construct an activity data set of sixty infected persons (14.3 percent of the total local cases in Hong Kong during
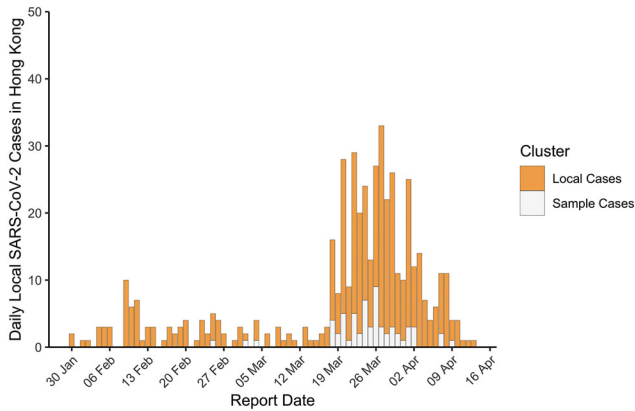
Figure 1. The temporal distribution of the COVID-19 cases in Hong Kong from 27 January to 14 April 2020.

the study period), with an average age of thirty-eight years (with a range of nineteen to sixty-nine years), including twenty-four women and thirty-six men. The data set includes all reported superspreaders (i.e., infected persons suspected to have infected many others through their activities at different venues), who caused 80 percent of all local cases in Hong Kong during the study period (Adam et al. 2020). Note that only the confirmed infected persons were admitted to hospitals for isolation and treatment in Hong Kong during the study period; people who had been in close contact with infected persons could still conduct their daily activities in various locations before they were confirmed as new cases. Further, the sixty selected cases are representative of the local cases during the study period in terms of their spatiotemporal distributions (Figure 1 and Figure 2A). Thus, we use the activity data of the sixty selected infected persons before they were confirmed as new cases for methodological investigation (i.e., uncertainties in the assessment of COVID-19 risk) in the study. Specifically, exposures of the sixty selected persons to the CHLs are used to assess people's exposures to COVID-19 risk.

In addition to the case-based high-risk environment, we use the data of all bars and restaurants in Hong Kong to represent another high-risk environment for assessing people's exposures to COVID-19 risk: the VHLs (i.e., locations with high COVID-19 risk due to the risky social gatherings and interactions that often occur there). These data were collected from the OpenRice Web site (see https://www.openrice.com/) and are used to represent the spatial distribution of the high-risk venues of social gathering (Figure 2B). OpenRice is the most

popular open-dining application in Hong Kong, providing the public with comprehensive dining information (e.g., location of restaurants or bars, opening hours, etc.). A Python program was developed and implemented to crawl the OpenRice data. There are 26,703 records in total (i.e., 885 bars and 25,818 restaurants) after data clearing and preprocessing (e.g., removal of takeout bakeries and closed stores). Each record includes the store name, location (i.e., latitude and longitude), type (e.g., tea restaurant, BBQ restaurant, or bar), and opening hours. We chose bars and restaurants to represent the VHLs because they were the venues with the highest COVID-19 risk during the study period (i.e., 27 January to 14 April) in Hong Kong according to the government's daily press briefings.
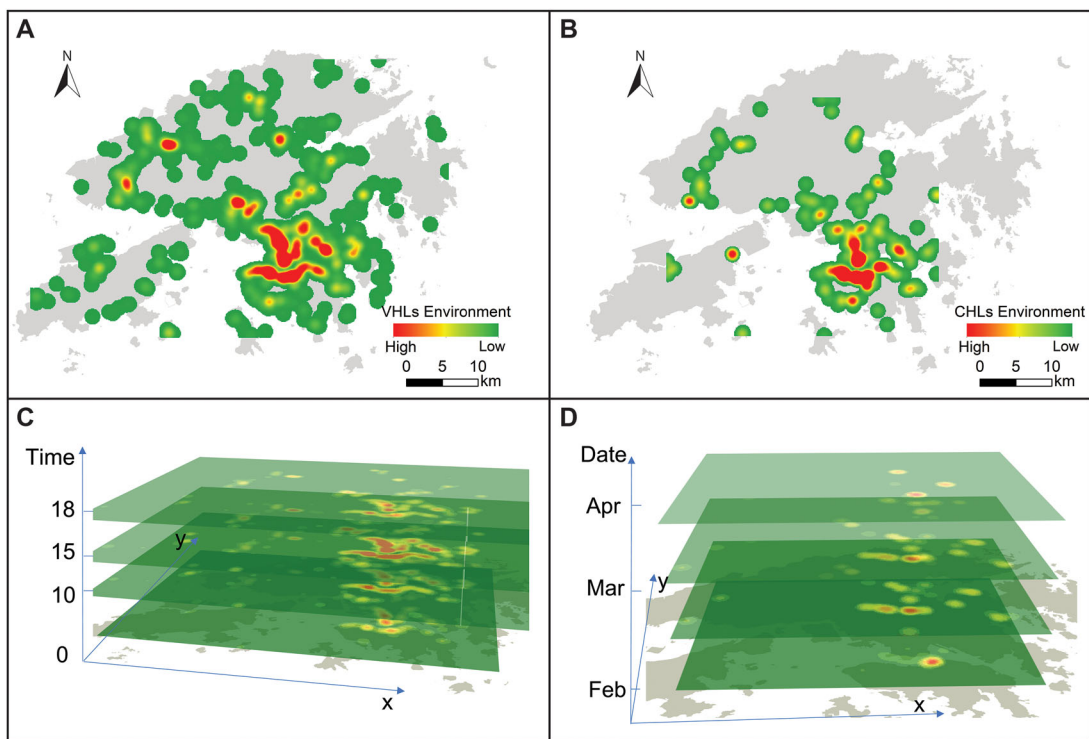
Case-Based and Venues-Based High-Risk Locations

As previously described, we use two types of high-risk locations to assess people's exposure to COVID-19 risk in this study. We call the 1,075 activity locations visited by all local cases (including the sixty selected cases) before they were confirmed as new cases the CHLs. These high-risk locations include various types of establishments, such as hair salons, schools, grocery stores, religious establishments, bars, and restaurants. On the other hand, the 26,703 bars and restaurants are called VHLs. These locations are venue based because the high COVID-19 risk at these locations is largely the result of their characteristic as popular venues of social gatherings. Note that 637 of the CHLs have geographic coordinates similar to those of some of the bars and restaurants among the VHLs, because they might be located in the same buildings (e.g., a building might have a fitness room visited by people who were later confirmed to be infected on the seventh floor and have bars and restaurants on the third floor). Further, some of the VHLs had been visited by people who were later confirmed to be infected and thus are also CHLs (i.e., 217 locations are both CHLs and VHLs). This suggests that assessments of people's exposure to COVID-19 risk based on the CHLs or VHLs should have a strong correlation.

We represent the space–time distributions of the CHLs and VHLs using three methods: the KDE method, the space–time cube method, and the space–time kernel density estimation (STKDE) method. First, the spatial distributions of the CHLs
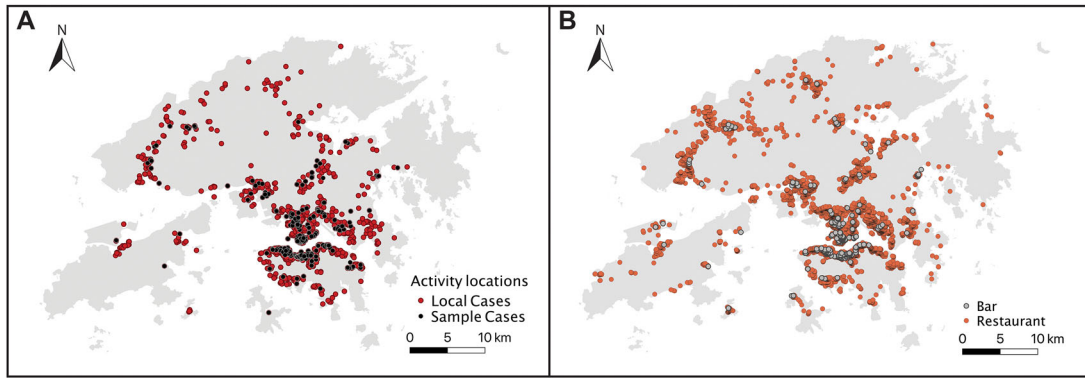
Figure 2. The spatial distribution of the selected confirmed cases and high-risk venues: (A) The spatial distribution of the activity locations of the sixty selected confirmed cases. (B) The spatial distribution of bars and restaurants in Hong Kong.

Figure 3. Methods for representing the VHL and CHL environments: (A) The VHL environment; (B) the CHL environment; (C) the VHL cube; (D) the CHL cube. VHL = venues-based high-risk location; CHL = case-based high-risk location.

or VHLs are represented by density surfaces derived using KDE, which generates a density surface from the locations of a set of points using a kernel function and a predetermined search radius (i.e., bandwidth). Kernel density surfaces take into account the decreasing effect of a location on nearby locations as distance increases from that location (i.e., distance decay) and have been widely used in environmental health studies (Thornton, Pearce, and Kavanagh 2011; Shi et al. 2019). In the context of this study, locations farther from a high-risk location (CHL or

VHL) are less influenced by that high-risk location than nearer locations are. Hence, KDE is applied to represent the spatial distribution of the CHLs and VHLs (Figures 3A, 3B). In addition, most people in Hong Kong are willing to walk for a distance of 3 km or less during normal times, but mobility in the city has decreased more than 50 percent since the end of January 2020 (Hung, Manandhar, and Ranasinghege 2010; Google LLC 2020). We thus used 1 km as the search radius and 100 m × 100 m as the spatial resolution in the KDE.

Second, the space–time cube method is applied to represent the spatiotemporal distribution of the CHLs and VHLs. Note that the risk-related characteristics of the CHLs and VHLs change across space and time. For instance, bars and restaurants (the VHLs) might operate on specific schedules and offer their services only during certain hours. Meanwhile, the coronavirus can survive on different environmental surfaces outside of its host organisms with the effect of time decay, which means that the concentration of coronavirus outside of its host organisms will decrease over time (van Doremalen et al. 2020). Thus, the space–time cube, which is constituted by a collection of 3D voxels in a regular grid in 3D space, can be used to capture the complex dynamics of the CHLs and VHLs (J. Wang and Kwan 2018). The value of each voxel represents the density of the CHLs or VHLs at a specific geographic location (x- and y-coordinates) and at a specific time (z-coordinate). To create the VHL space–time cube, the KDE method is first applied based on the number of the VHLs that operate during each of the twenty-four time slots of a day. The twenty-four raster VHL layers are then voxelized with each layer representing one hour and mapped to the z-axis. Hence, a 3D space–time cube is constructed to represent the VHL environment in Hong Kong during the study period (Figure 3C). Note that the government decided to close all bars from 2 April to control the spread of COVID-19. Thus, the VHL space–time cube was constructed to represent its environment before 2 April.

Third, the STKDE method is used to generate the CHL space–time cube (Figure 3D). As mentioned previously, the concentration of coronaviruses on different environmental surfaces decreases over time. Thus, the CHL space–time cube should consider the decay effects of distance and time. The STKDE method developed by Brunsdon et al. (2007) multiplies a bivariate kernel placed over the x–y (spatial) domain by a univariate kernel along the temporal dimension z to estimate the density of an event while taking into account the decay effects of distance and time. It is formulated as shown here:

$$\hat{f}(x, y, t) = \frac{1}{n h_s^2 h_t} \sum_{i=1}^{n} I(d_i < h_s, t_i < h_t)$$
$$K_s\left(\frac{x - x_i}{h_s}, \frac{y - y_i}{h_s}\right) K_t\left(\frac{t - t_i}{h_t}\right), \quad (1)$$

where $\hat{f}(x, y, t)$ is the estimated density at location (x, y, t), n is the number of points, d_i and d_t are the spatial and temporal distance between any point i and the location (x, y, t), and h_s and h_t are the spatial and temporal search radius (i.e., bandwidth). In this study, the kernel functions K_s and K_t are defined using the Epanecknikov kernel (Epanechnikov 1969), which is a common type of kernel function used in ArcGIS (Nakaya and Yano 2010):

$$K_s(u, v) = \begin{cases} \frac{2}{\pi}(1 - (u^2 + v^2)) & (u^2 + v^2) < 1, \\ 0 & otherwise \end{cases} \quad (2)$$

$$K_t(w) = \begin{cases} \frac{3}{4}(1 - w^2) & w^2 < 1, \\ 0 & otherwise. \end{cases} \quad (3)$$

In addition, the indicator function $I(d_i < h_s, t_i < h_t)$ takes on a value of one when d_i and d_t are smaller than the spatial and temporal search radius (i.e., h_s and h_t); otherwise, its value is zero. Similar to the KDE method mentioned earlier, the spatial search radius is 1 km and the spatial resolution is 100 m × 100 m. The temporal resolution is one hour and the temporal search radius is three days (i.e., 72 hours), because the coronavirus can survive on some surfaces (e.g., plastic and stainless steel) up to three days (van Doremalen et al. 2020). Note that although previous studies have provided useful information on STKDE calibration (e.g., Hohl et al. 2016), these studies largely used STKDE to detect significant space–time clusters of infectious diseases. Their primary purpose is different from ours: We used STKDE to represent the COVID-19 risk environment by considering the space–time decay effect. Thus, the parameters we selected (i.e., search radius = 1 km and search duration = 72 hours) is determined by the specific context (i.e., people in Hong Kong are willing to walk for a distance of 3 km but mobility in the city has decreased more than 50 percent in the study period, and the novel coronavirus can survive on some surfaces up to three days).

Constructing the Exposure Space of the Infected Persons Based on Their Activity Data

To assess the COVID-19 risk and exposures of the sixty infected persons to the CHLs and VHLs during the study period, we construct their exposure space based on their activity data. First, the
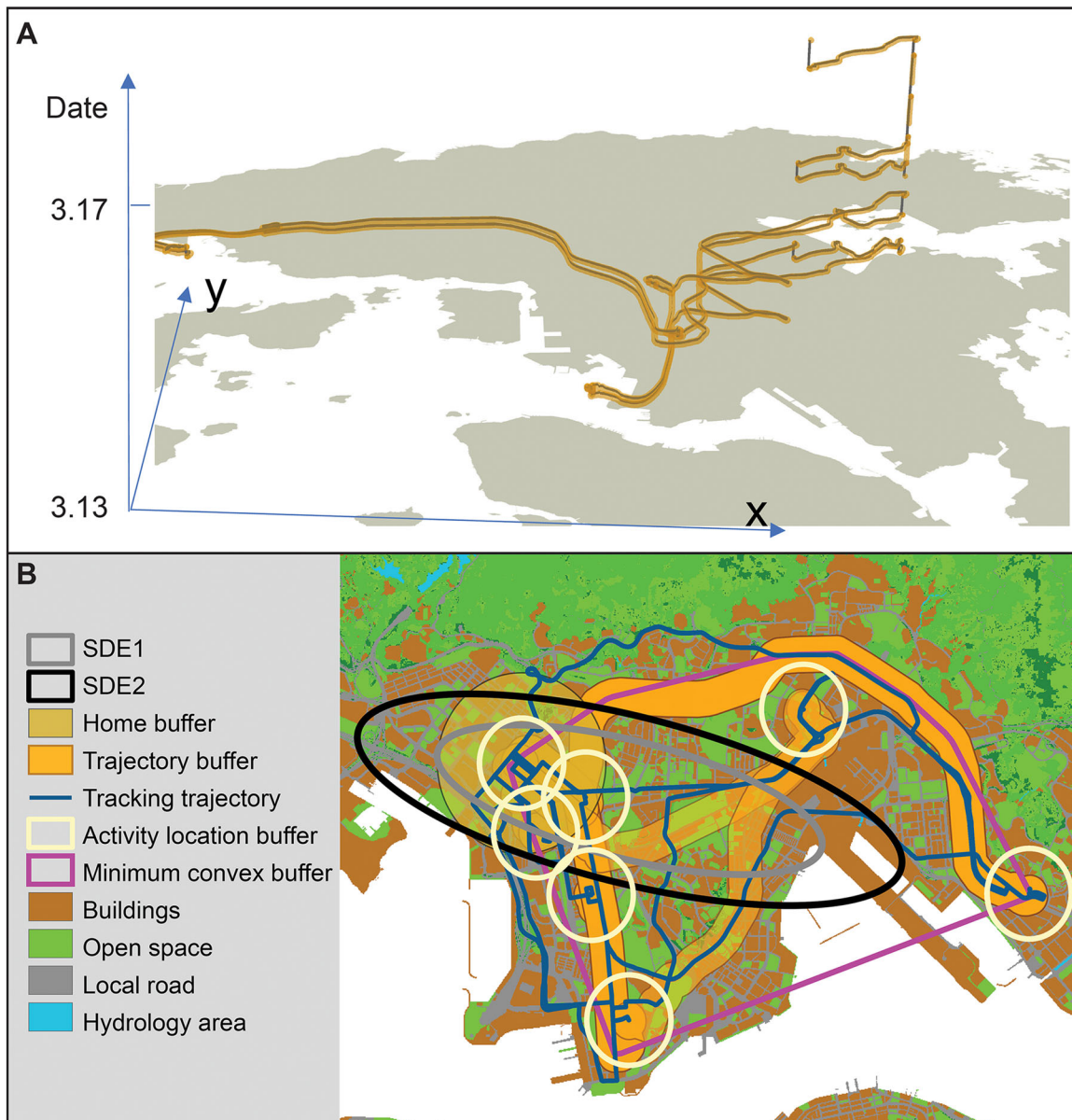
Figure 4. The exposure space of an infected person constructed by different methods: (A) The space–time trajectories and (B) the six activity space methods used in the study.

space–time trajectory between each pair of subsequent activity locations of an infected person is constructed using the Google Maps Directions API (Figure 4A). Note that because the activity data did not include actual Global Positioning System (GPS) trajectories of each infected person, the exact routes traveled are unknown. Thus, we use the Google Maps Directions API to estimate the shortest time routes because its algorithm reflects on-ramps, one-way streets, traffic conditions, the speed limits of different road segments, and so on (e.g., Kim and Kwan 2019). Further, note that we do not have data

on how risky or busy different public transport modes (i.e., bus, the Mass Transit Railway, or ferry) are during the pandemic and thus assume that the levels of COVID-19 risk of these travel modes are the same. The government data of the COVID-19 cases in Hong Kong do not indicate that using public transport (or undertaking activities in crowded venues like shopping) is a high-risk activity.

Then, six widely used activity space methods (Crawford et al. 2014; J. Wang, Kwan, and Chai 2018; Kwan et al. 2019) are implemented to delineate the exposure spaces of the sixty infected

Table 2. The three approaches of COVID-19 risk measurements and the respective measures in the study

COVID-19 risk environment	Mobility-based approach	Residence-based approach	Activity space–based approach
Case-based high-risk locations	ME^C	RE^C	AS^C (i.e., $2DTBs^C$, MCH^C, HB^C, $SDE1^C$, $SDE2^C$, $ALBs^C$)
Venues-based high-risk locations	ME^V	RE^V	AS^V (i.e., $2DTBs^V$, MCH^V, HB^V, $SDE1^V$, $SDE2^V$, $ALBs^V$)

persons based on their activity locations and the constructed space–time trajectories (Figure 4B). These methods are 2D trajectory buffers (2DTBs), standard deviation ellipses with one or two standard deviation(s) (SDE1 and SDE2), minimum convex polygons (MCPs), activity-location buffers (ALBs), and home buffers (HBs). The 2DTBs are created by a 200-m buffer along the infected persons' trajectories. The SDEs represent the spatial distribution and directional trends of an infected person's activity locations. The MCP is the smallest convex polygon that contains all the activity locations of an infected person. The ALB is a 500-m area around each of the activity locations of an infected person. The HB is a 1-km buffer area around an infected person's home location. Note that these different delineations of individual activity spaces have their respective strengths and weaknesses (e.g., the MCP tends to include a considerable area that is actually not in a person's activity space). They are used in this study largely for comparative purposes to examine which can best capture a person's COVID-19 risk.

Measuring People's COVID-19 Risk Exposures

Using these delineated exposure spaces of the sixty infected persons, we assess their COVID-19 risk based on their exposures to the CHLs or VHLs using three approaches (see Table 2): the mobility-based approach, the residence-based approach, and the activity space–based approach. Thus, the COVID-19 risk assessed by each of these three approaches has two versions: One is based on the CHLs and the other is based on the VHLs. Note that the exposure measures used in the study are based on the notion that an individual's COVID-19 risk is heavily influenced by the duration and frequency of his or her exposure to high-risk locations (Baker, Peckham, and Seixas 2020; St-Denis 2020). Equation 4 measures ME_i^k, which is person i's mobility-based COVID-19 risk due to exposure to the CHLs (i.e., ME_i^C) or VHLs (i.e., ME_i^V):

$$ME_i^k = \sum_{t=T_s i}^{T_e i} S^k(x_t, y_t, t), \qquad (4)$$

where $T_s i$ and $T_e i$ denote the start and end time of the activities of individual i. S^k is the risk density S^C or S^V obtained from the CHL and VHL cubes, and (x_t, y_t, t) is person i's location at time t. Equation 5 measures RE_i^k, which is person i's residence-based risk due to exposure to the CHLs (i.e., RE_i^C) or VHLs (i.e., RE_i^V):

$$RE_i^k = \sum_{t=T_s i}^{T_e i} S^k(x_r, y_r, t), \qquad (5)$$

where (x_r, y_r) denotes person i's residential location reported in the collected activity data. Equation 6 measures ASE_i^k, which is the person's activity space–based COVID-19 risk evaluated by the person's exposure to the CHLs (i.e., AS_i^C) or VHLs (i.e., AS_i^V) in his or her activity space:

$$AS_i^k = \sum C^k, \qquad (6)$$

where C^k is the sum of the values derived from the KDEs of the CHL (C^C) or VHL (C^V) environment within the activity space of individual i. In this study, all six activity space methods described earlier in the article are implemented to estimate AS_i^k. Thus, the AS_i^k includes $2DTBs_i^k$, $SDE1_i^k$, $SDE2_i^k$, $ALBs_i^k$, MCH_i^k, and HB_i^k. Again, note that the COVID-19 risk assessed by each of these activity space–based methods has two versions: One is based on the CHLs and the other is based on the VHLs. Further, we use the CHL-based measures to further identify the existence of the UGCoP and the NEAP in the assessments of COVID-19 risk, while using the associations between CHL-based and VHL-based COVID-19 risk measures to examine the effects of the UGCoP and the NEAP on such assessments.

Analytical Approach

After assessing people's COVID-19 risk using these three approaches, we further examine the effects of the UGCoP and the NEAP on such assessments. First, we compare the CHL-based measures (i.e., ME^c, RE^c, and AS^c) to evaluate whether the UGCoP exists. Note that the ME^c and RE^c measures are based on high-risk environments different from that of the AS^c. The ME^c and RE^c measures are based on the CHL space–time cube, whereas the AS^c measure is based on the KDEs of the CHLs. To compare them for examining the UGCoP, we convert each of these three assessed COVID-19 risk levels to standardized scores (i.e., dividing each risk level by the standard deviation of all risk levels obtained with the same measure).

Further, the NEAP is examined by comparing the ME^C and RE^C measures based on three different methods (Kim and Kwan 2021b): (1) using descriptive statistics to examine whether there is a decrease in the standard deviation of the ME^C measure when compared to the RE^C; (2) comparing the probability distribution functions (PDFs) of the ME^C and RE^C to assess whether there is a tendency for the ME^C to converge to the average exposure value; and (3) using scatterplots to explore the relationship between the RE^C (i.e., x-axis values) and how much the RE^C are higher or lower than the ME^C (i.e., using RE^C-ME^C as y-axis values). The scatterplots are used to analyze whether persons with high RE^C tend to experience lower ME^C and whether individuals with low RE^C tend to experience higher ME^C. As mentioned earlier, people's COVID-19 risk assessed based on the CHLs and VHLs should be highly correlated. Hence, we use multivariate linear regression models to examine the associations between the CHL-based and VHL-based COVID-19 risk measures.

Results

Variations in COVID-19 Risk Levels Obtained with Different Methods

In this subsection, we use the CHL-based measures to identify the existence of the UGCoP in the assessment of COVID-19 risk. Figure 5 presents the standardized risk measures for each of the sixty infected persons. In Figure 5, the horizontal axis indicates these persons, whereas the vertical axis shows their COVID-19 risk levels obtained with different methods. Figure 5 shows that different methods give considerably different COVID-19 risks for the same individual.

Bivariate correlation analysis is used to investigate the relationships among all COVID-19 risks assessed based on the CHLs. Table 3 presents the results: More than half of the pairs do not have significant correlations, including the pairs of ME^C–$2DTBs^C$, ME^C–$SDE1^C$, ME^C–$SDE2^C$, ME^C– $ALBs^C$, ME^C–MCH^C, ME^C–HB^C, RE^C– $2DTBs^C$, RE^C–$SDE1^C$, RE^C–$SDE2^C$, RE^C– $ALBs^C$, RE^C–MCH^C, RE^C–HB^C, MCH^C–HB^C, HB^C– $SDE1^C$, and HB^C– $SDE2^C$. Although the other pairs show significant correlations, most of them have correlation coefficients smaller than 0.6, which indicate moderate to low associations. Specifically, only the
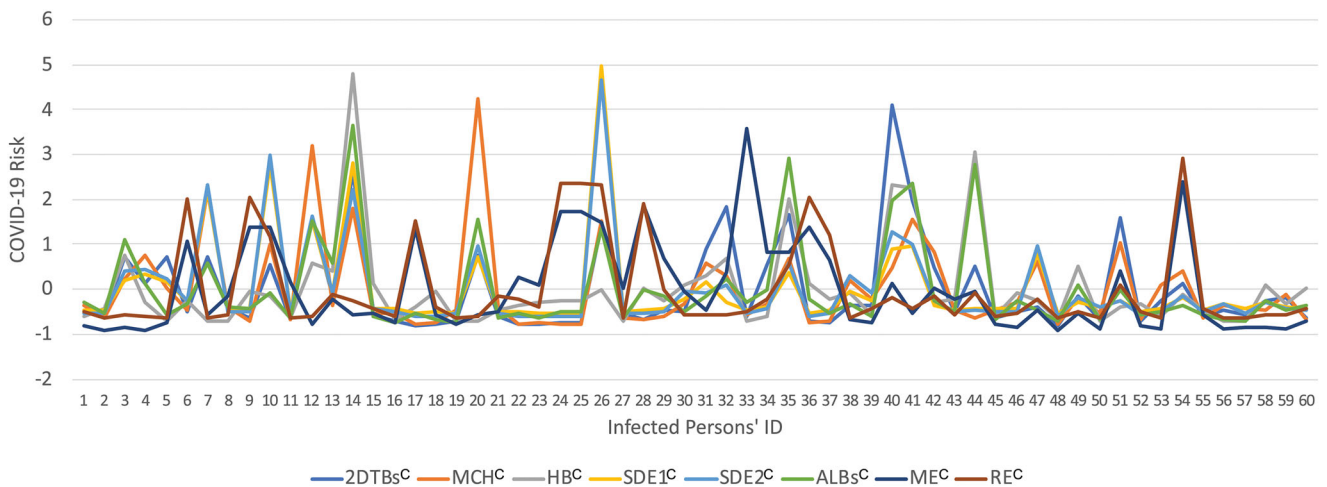

Figure 5. Comparison of the COVID-19 risk obtained with different methods for each infected person.

Table 3. The results of the bivariate Pearson correlation analysis between each pair of the assessed COVID-19 risk measures

	$2DTBs^C$	MCH^C	HB^C	$SDE1^C$	$SDE2^C$	AB^C	ME^C	RE^C
$2DTBs^C$	—	0.57*	0.44*	0.53*	0.53*	0.64*	0.15	0.11
MCH^C	—	—	0.27	0.62*	0.62*	0.59*	0.13	0.10
HB^C	—	—	—	0.36	0.35	0.67*	0.11	0.12
$SDE1^C$	—	—	—	—	0.94*	0.52*	0.07	0.16
$SDE2^C$	—	—	—	—	—	0.53*	0.11	0.12
AB^C	—	—	—	—	—	—	0.13	0.16
ME^C	—	—	—	—	—	—	—	0.90*
RE^C	—	—	—	—	—	—	—	—

Note: *Significant at the 5 percent level.

Table 4. Descriptive statistics of ME^C and RE^C

		M	SD
ME^C	Mobility-based approach	3,633.15	3,723.94
RE^C	Residence-based approach	3,065.90	4,192.10
	Difference	2.46*a	

Note: [a]Paired sample *t* test.
*Significant at the 5 percent level.

pairs of MCH^C–$SDE1^C$, MCH^C–$SDE2^C$, $SDE1^C$–$SDE2^C$, AB^C–$2DTBs^C$ AB^C–HB^C, and ME^C–RE^C show strong associations (i.e., correlation coefficients larger than 0.6). It is reasonable that these pairs have strong associations because most of the sixty infected persons had relatively circumscribed activity spaces (e.g., they spent most of their time at home) before they were confirmed as new COVID-19 cases under the government's nonpharmaceutical interventions (e.g., work-from-home order). Note that the sixty infected persons were free to conduct their daily activities before they were confirmed as new cases, even if they might have been in close contact with infected persons. The results show that the COVID-19 risk measures obtained with different approaches are different, indicating the existence of the UGCoP (i.e., different delineations of individual exposure space yield different COVID-19 risk levels).

Comparisons of Mobility-Based and Residence-Based COVID-19 Risk

In this subsection, we examine whether the assessment of COVID-19 risk is affected by the NEAP by comparing mobility-based and residence-based COVID-19 risk due to exposure to the CHLs (i.e., ME^C and RE^C). Table 4 presents the descriptive statistics of the ME^C and RE^C measures of the infected persons. As shown in Table 4, the mean value of ME^C is higher than that of RE^C, which indicates that people's COVID-19 risk can be underestimated if their daily mobility is ignored. The pairwise differences between ME^C and RE^C are significant. Meanwhile, the standard deviation of ME^C is smaller than that of RE^C. Further, Figure 6A indicates that the PDF of ME^C is less deviated than that of RE^C, indicating that people's COVID-19 risk tends to converge toward the average risk level when their daily mobility is taken into account. Moreover, Figure 6B shows that there is a positive linear relationship between RE^C (i.e., x-axis values) and values obtained by subtracting the RE^C from the ME^C (i.e., y-axis values). The results provide strong evidence for the presence of the NEAP in the assessment of COVID-19 risk.

In addition, the results suggest that the NEAP might not affect the entire group of infected persons in a similar way. Specifically, note that individuals whose RE^C levels are very high or very low and who move around considerably in their daily life are exposed to risk environments that are considerably different from those in their residential locations (marked as A and B in Figure 6B). This means that people who live in residential locations with many high-risk venues might work or conduct their daily activities at low-risk locations outside of their residential locations but could catch COVID-19 and transmit it to others in these low-risk locations (i.e., A in Figure 6B). On the other hand, people who live in residential locations with few high-risk venues might work or conduct their daily activities at high-risk locations in areas outside of their residential locations and get infected in these high-risk areas (i.e., B in Figure 6B). The results further imply that ignoring
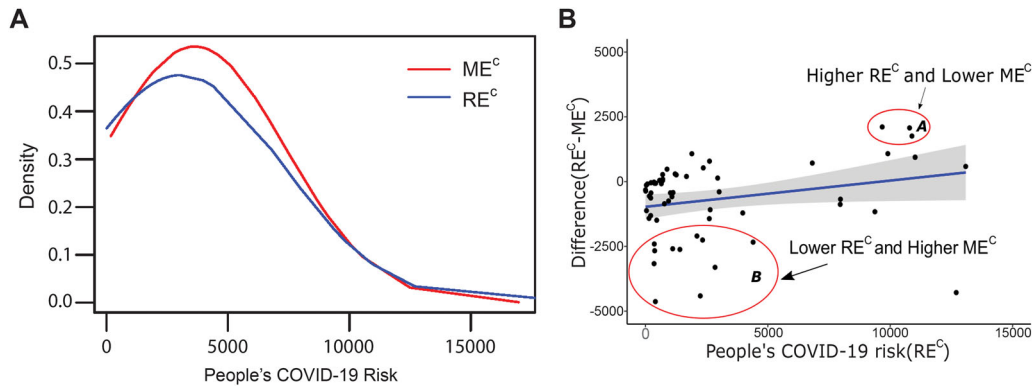
Figure 6. Methods for identifying the neighborhood effect averaging problem in the assessment of COVID-19 risk: (A) Probability density functions of RE^C (blue line) and ME^C (red line). (B) A scatterplot with the x-axis values as the RE^C and the y-axis values obtained by subtracting the ME^C from the RE^C.

people's daily mobility might lead to misleading results in the assessment of COVID-19 risk.

Further, we conduct 3D geovisualizations of the space–time trajectories of two infected persons who have low RE^C and high ME^C or high RE^C and low ME^C using ArcGIS Pro (Figure 7). Figure 7A shows an individual who lives in a residential location with a low density of high-risk venues (i.e., a suburban area on Lantau Island) but she or he has to go to work and conduct other daily activities in nonresidential locations with a high density of high-risk venues (i.e., low RE^C and high ME^C). On the other hand, Figure 7B shows an opposite scenario where an individual who lives in a residential location with a high density of high-risk venues works and conducts other daily activities in nonresidential locations with a low density of high-risk venues (i.e., high RE^C and low ME^C). These two cases corroborate and help illustrate our findings concerning the effects of the UGCoP and the NEAP in the assessment of COVID-19 risk. The results of the preceding analysis thus clearly indicate that the UGCoP and the NEAP exist in ecological COVID-19 studies.

Association between CHL-Based and VHL-Based Measures

In this subsection, we examine the associations between CHL-based and VHL-based COVID-19 risk using eight multivariate linear regression models. In these models, the dependent variable is the CHL-based COVID-19 risk measures, and the independent variable is the VHL-based COVID-19 risk measures (i.e., evaluating how venues-based risk influences case-based risk). The models are controlled for the age (a continuous variable) and

activity diversity (the number of different types of activities conducted) of the sixty infected persons. A total of eight models are estimated to examine the associations between CHL-based and VHL-based COVID-19 risk levels obtained with the mobility-based approach, the residence-based approach, and six activity space methods. The performance of these models is compared using the Akaike's information criterion (AIC), the adjusted R^2, and the corresponding p value. Furthermore, the models are compared to see whether a significant association exists between a CHL-based measure and the corresponding VHL-based measure. A higher significant association indicates a more reliable COVID-19 risk assessment.

Table 5 shows the results of the regression models. Table 5 shows that all models are statistically significant with a p value <0.001. Among the models, the most robust one is the mobility-based model, which has the lowest AIC (81.94) and p value <0.001. The model explains 79 percent of the variance in people's CHL-based COVID-19 risk. Meanwhile, the least robust model is the one based on the ALBs (AIC = 156.85), which only explains 28 percent of the variance.

The associations between the CHL-based and VHL-based COVID-19 risk obtained by different exposure-space methods are shown in Table 6. In all models, VHL-based risk has a significant positive association with CHL-based risk, which means that people who conduct their daily activities in locations with a higher density of VHLs also have a higher CHL-based risk. The ME model, in particular, yields the highest association between CHL-based and VHL-based COVID-19 risk. Recall that a higher association indicates a more reliable COVID-19 risk
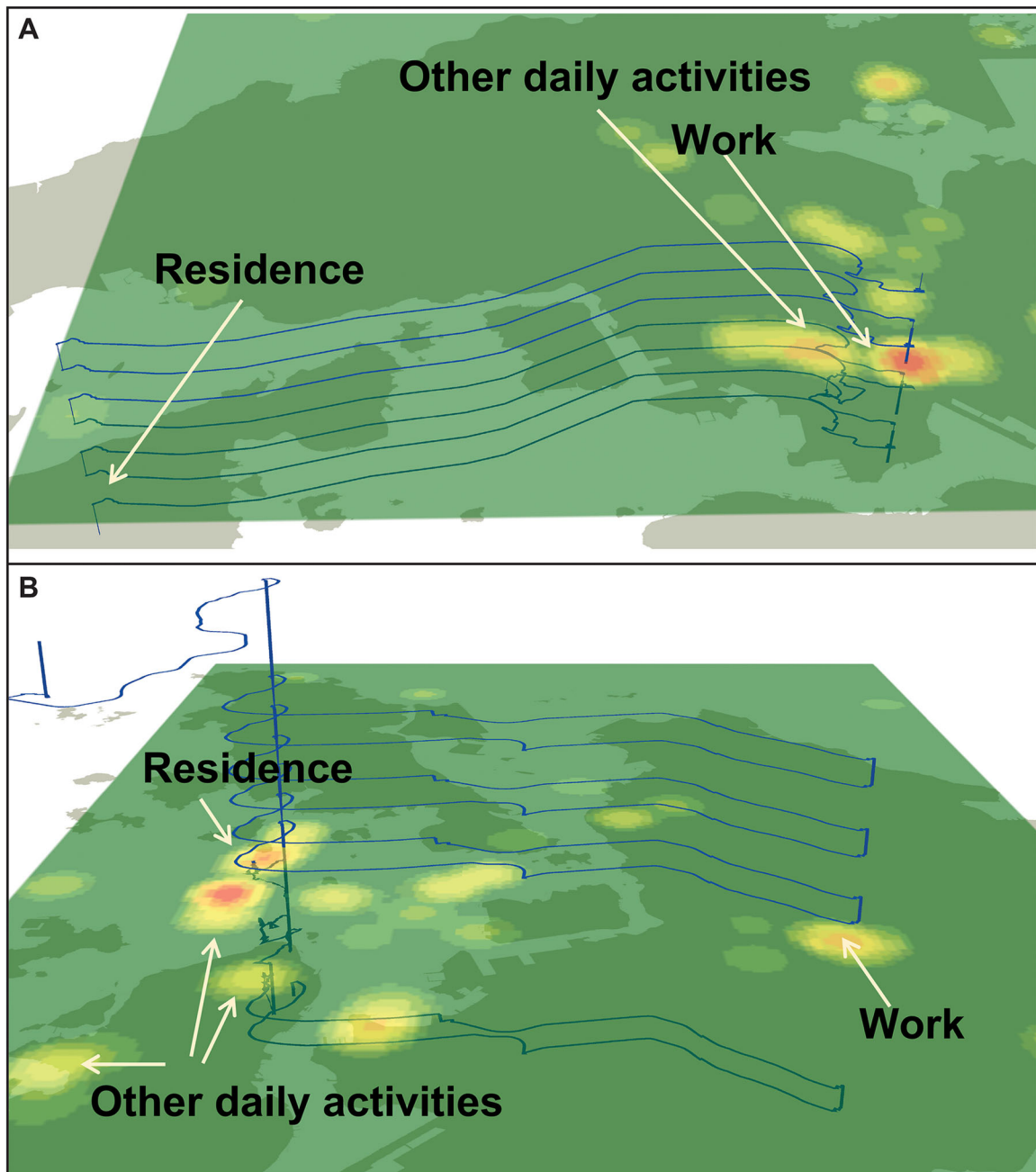
Figure 7. The space–time trajectories of the two different individuals exposed to different high-risk environments: (A) Low RE^C and high ME^C and (B) high RE^C and low ME^C.

assessment. Thus, the results suggest that using the mobility-based approach to assess people's COVID-19 risk can generate more reliable results when compared to other methods. Meanwhile, there is a significant positive association between people's activity diversity and their COVID-19 risk assessed based on the CHLs using the 2DTBs, MCH, and ALBs models. This suggests that people who conduct more diverse activities in their daily lives tend to

have a higher COVID-19 risk, probably due to their higher probability of exposure to high-risk locations. Finally, only in the 2DTBs model does the age of people have a positive association with COVID-19 risk.

These results indicate that using the mobility-based approach to assess people's COVID-19 risk generates more reliable results when compared to other methods. Thus, the inconsistent findings in

previous studies might be partly due to ignoring people's daily mobility and its interactions with the complex and dynamic COVID-19 risk environment. Further, the results imply that taking into account people's daily mobility when measuring COVID-19 risk can help address the NEAP and mitigate the UGCoP.

Discussion and Conclusion

This study used the activity data of sixty infected persons to examine how the assessment of COVID-19 risk might be affected by the UGCoP and the NEAP. Three approaches (i.e., the mobility-based approach, the residence-based approach, and the activity space–based approach) were used to assess

Table 5. The results of the regression models for examining the associations between case-based high-risk locations–based measures and venues-based high-risk locations–based measures

Model	AIC	Adjusted R^2	p Value
2DTBs	150.11	0.35	0.000***
MCH	135.37	0.49	0.000***
HB	152.66	0.34	0.000***
SDE1	134.30	0.50	0.000***
SDE2	129.86	0.54	0.000***
ALBs	156.85	0.28	0.000***
ME	81.94	0.79	0.000***
RE	88.45	0.72	0.000***

Note: AIC = Akaike's information criterion.
*** Significant at the 0.1 percent level, or *** p-value < 0.001.

people's exposure based on CHLs and VHLs. By comparing the CHL-based COVID-19 risk obtained with different approaches, the results indicate that the UGCoP exists in the assessment of COVID-19 risk. Further, by comparing the mobility-based and residence-based COVID-19 risk, the NEAP was also observed in the assessment of COVID-19 risk. Finally, the highest association between the CHL-based and VHL-based COVID-19 risk levels was obtained by the mobility-based approach, which suggests that using the mobility-based approach to assess people's COVID-19 risk generates more reliable results when compared to those obtained by other methods. The study thus contributes to the ecological study of infectious diseases by showing how the UGCoP and the NEAP might affect the assessment of the risk of COVID-19 and other infectious diseases.

The results of the study have important implications for COVID-19 studies and control measures. First, ecological COVID-19 studies need to address the uncertainties due to the UGCoP and the NEAP by considering people's mobility in their daily lives because a small number of infected persons can lead to a large wave of COVID-19 transmission (e.g., 80 percent of the local cases in Hong Kong can be traced back to 19 percent of all infected persons who are the superspreaders; Adam et al. 2020). Further, ignoring people's daily mobility and its interactions with the complex and dynamic COVID-19 risk environment can underestimate COVID-19 risk, which might lead to ineffective control

Table 6. The association between case-based high-risk locations–based and VHL-based COVID-19 risk analyzed by the regression models for different methods of representing exposure space

Independent variables	Model	Coefficient	Standard	Model	Coefficient	Standard
Age	2DTBs	0.03*	0.11	MCH	0.01	0.09
Activity diversity		0.20*	0.11		0.31**	0.10
VHL-based risk		0.53***	0.14		0.53***	0.10
Age	HB	0.05	0.11	SDE1	−0.01	0.09
Activity diversity		0.14	0.11		0.03	0.10
VHL-based risk		0.48***	0.12		0.72***	0.10
Age	SDE2	0.01	0.09	ALBs	−0.03	0.11
Activity diversity		0.03	0.09		0.20*	0.11
VHL-based risk		0.74***	0.09		0.49***	0.11
Age	ME	0.18	0.11	RE	0.18	0.12
Activity diversity		−0.18	0.11		0.01	0.12
VHL-based risk		0.82***	0.12		0.78***	0.12

Note: VHL = venues-based high-risk location.
*Significant at the 10 percent level.
**Significant at the 1 percent level.
***Significant at the 0.1 percent level.

measures. Policymakers should thus be aware of the effects of the NEAP on the assessment of COVID-19 risk when formulating measures to control the transmission of COVID-19. Because there is some evidence that people with similar mobility patterns tend to have similar levels of exposure to environmental risks (e.g., S. Ma et al. 2021), future studies on individual-level risk to infectious diseases could use this association to tailor specific control measures to different social groups with distinctive mobility characteristics (e.g., office workers).

Second, examining people's COVID-19 exposure risk based on the NEAP revealed important mechanisms that greatly enhance our understanding of the possibilities for different groups of people to mitigate their risk through mobility. Specifically, the NEAP enabled us to identify specific social groups that do not have an option to decide which trips to make or to forego: (1) people with low income who live in low-risk neighborhoods and have to work in locations with a high density of high-risk venues and (2) people with low income who live in high-risk neighborhoods but cannot reduce their COVID-19 exposure risk (e.g., moving out of high-risk neighborhoods). On the other hand, individuals who have the option of deciding which trips to make or to forego can limit their exposure to high-risk areas (e.g., working from home or conducting daily activities in low-risk neighborhoods). Although previous studies have observed these phenomena in China (Kraemer et al. 2020), Hong Kong (Huang et al. 2020), the United States (Chang et al. 2021; Kim and Kwan 2021c), the United Kingdom (Lee, Qian, and Schwanen 2021), France (Pullano et al. 2020), and Chile (Gozzi et al. 2021), examining them via the NEAP enables us to not only identify the specific ways in which different social groups are affected because of their different abilities for mobility change but also to highlight the advantaged or disadvantaged situations of different social groups.

Third, the mobility- and activity-based perspective of the study enables us to more readily identify vulnerable social groups and high-risk areas and behaviors. This knowledge in turn can help health authorities to develop more targeted place-based and area-specific intervention measures for effectively controlling the spread of COVID-19 and future pandemics (e.g., health authorities can advise people to avoid visiting high-risk areas or undertaking high-risk activities and implement area-specific lockdowns and mass

testing). For instance, Jordan is a densely packed neighborhood in Hong Kong and is known for the city's highest concentrations of "coffin homes" and "cages" (i.e., apartments that are parceled out into two or more smaller subunits; V. Wang and May 2021). The neighborhood witnessed 160 COVID-19 infected persons in the first two weeks of January 2021 out of about 1,100 cases citywide in the same period. The Hong Kong government responded by locking down 10,000 residents in Jordan for two days to conduct mass testing to control the spread of COVID-19 (V. Wang and May 2021). In this light, policymakers can invest more resources to improve the capacity for contact tracing, testing, and vaccination in neighborhoods where people might have a low ability to reduce their COVID-19 exposure risk.

In addition to the empirical contribution and policy implications, the study aligns with a developing line of research in geography and public health that shifts the emphasis from residence-based assessment of environmental exposures to mobility-based dynamic assessment (which often involves the use of real-time tracking and mobile sensing). This is a highly promising, nascent area of research that has tremendous potential to enhance our understanding of a wide range of environmental exposures and their health impacts. The study contributes to this emerging area of research.

Although this study is significant because it is one of the first studies that examine the UGCoP and NEAP on the assessment of COVID-19 risk, it has several limitations that should be addressed in future research. First, due to data limitations, the activity data of the infected persons are for an early stage of the pandemic, which might not be able to capture the detailed spatiotemporal variations of the pandemic under different nonpharmaceutical interventions (Kan et al. 2021). For instance, the Hong Kong government mandated the closing of certain social gathering venues (e.g., bars) at the beginning of the third wave (i.e., from July to September 2020) due to the lessons learned from the first and second waves (i.e., from 29 January to 14 April 2020). It did not stop the third wave of the COVID-19 outbreak in other social gathering venues, however (e.g., nursing homes; Law, Leung, and Xu 2020). Thus, future studies should further examine how the NEAP affects the assessment of COVID-19 risk under different nonpharmaceutical interventions. Although certain characteristics of the COVID-19 pandemic in Hong Kong

might have changed after the second wave as a result of different nonpharmaceutical interventions, weather, and other factors that had changed over time, these differences are unlikely to significantly affect the conclusion of the study: The UGCoP and the NEAP exist in the assessment of COVID-19 exposure risk (because the differences between residence-based assessments and mobility-based or activity space–based assessments will likely remain, although the levels of exposure risk might be different).

Second, although this study used the Google Place API to construct the space–time trajectories of the infected persons, the constructed trajectories might not be the actual ones (i.e., different travel routes could be used). The results from earlier studies on the food environment and air pollution exposure, however, have identified the UGCoP and the NEAP based on real-time GPS tracking data (J. Wang and Kwan 2018; J. Ma et al. 2020). The results of these studies are consistent with those in our study. Thus, the methods used in this study based on public data could still be used to further explore the UGCoP and the NEAP in ecological COVID-19 studies and research on other infectious diseases.

Finally, it should be noted that the NEAP might lead to different conclusions concerning the experiences of socially vulnerable groups in the pandemic. For instance, improving mobility might mitigate the effects of the NEAP and thus might reduce the air pollution exposure for low-income people who live in a residential location with a high level of air pollution (X. Ma et al. 2020). People who live in a suburban low-risk location with high mobility, however, might experience significantly higher COVID-19 risk if they work in high-risk locations, as we demonstrated earlier. Hence, it is important to conduct careful analysis to understand the effects of the NEAP on people's exposure to COVID-19 risk and other environmental factors (e.g., air pollution).

Acknowledgment

Funding

ORCID

Jianwei Huang ⓘ http://orcid.org/0000-0002-2230-3446
Mei-Po Kwan ⓘ http://orcid.org/0000-0001-8602-9258

References

Adam, D., P. Wu, J. Wong, E. Lau, T. Tsang, S. Cauchemez, G. Leung, and B. Cowling. 2020. Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. *Nature Medicine* 26 (11):1714–19. doi: 10.1038/s41591-020-1092-0.

Alkhamis, M. A., S. Al Youha, M. M. Khajah, N. B. Haider, S. Alhardan, A. Nabeel, S. Al Mazeedi, and S. K. Al-Sabah. 2020. Spatiotemporal dynamics of the COVID-19 pandemic in the State of Kuwait. *International Journal of Infectious Diseases: Official Publication of the International Society for Infectious Diseases* 98:153–60. doi: 10.1016/j.ijid.2020.06.078.

Amram, O., S. Amiri, R. B. Lutz, B. Rajan, and P. Monsivais. 2020. Development of a vulnerability index for diagnosis with the novel coronavirus, COVID-19, in Washington State, USA. *Health & Place* 64:102377. doi: 10.1016/j.healthplace.2020.102377.

Azuma, K., N. Kagi, H. Kim, and M. Hayashi. 2020. Impact of climate and ambient air pollution on the epidemic growth during COVID-19 outbreak in Japan. *Environmental Research* 190:110042. doi: 10.1016/j.envres.2020.110042.

Badr, H. S., H. Du, M. Marshall, E. Dong, M. M. Squire, and L. M. Gardner. 2020. Association between mobility patterns and COVID-19 transmission in the USA: A mathematical modelling study. *The Lancet Infectious Diseases* 20 (11):1247–54. doi: 10.1016/S1473-3099(20)30553-3.

Baker, M. G., T. K. Peckham, and N. S. Seixas. 2020. Estimating the burden of United States workers exposed to infection or disease: A key factor in containing risk of COVID-19 infection. *PLoS ONE* 15 (4):e0232452. doi: 10.1371/journal.pone.0232452.

Bian, L. 2004. A conceptual framework for an individual-based spatially explicit epidemiological model.

Environment and Planning B: Planning and Design 31 (3):381–95. doi: 10.1068/b2833.

Bian, L., Y. Huang, L. Mao, E. Lim, G. Lee, Y. Yang, M. Cohen, and D. Wilson. 2012. Modeling individual vulnerability to communicable diseases: A framework and design. *Annals of the Association of American Geographers* 102 (5):1016–25. doi: 10.1080/00045608.2012.674844.

Brunsdon, C., J. Corcoran, and G. Higgs. 2007. Visualising space and time in crime patterns: A comparison of methods. *Computers, Environment and Urban Systems* 31 (1):52–75. doi: 10.1016/j.compenvurbsys.2005.07.009.

Chang, S., E. Pierson, P. W. Koh, J. Gerardin, B. Redbird, D. Grusky, and J. Leskovec. 2021. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* 589 (7840):82–87. doi: 10.1038/s41586-020-2923-3.

Cheng, T., and M. Adepeju. 2014. Modifiable temporal unit problem (MTUP) and its effect on space-time cluster detection. *PLoS ONE* 9 (6):e100465. doi: 10.1371/journal.pone.0100465.

Coccia, M. 2020. Factors determining the diffusion of COVID-19 and suggested strategy to prevent future accelerated viral infectivity similar to COVID. *Science of the Total Environment* 729:138474. doi: 10.1016/j.scitotenv.2020.138474.

Crawford, T. W., S. B. J. Pitts, J. T. McGuirt, T. C. Keyserling, and A. S. Ammerman. 2014. Conceptualizing and comparing neighborhood and activity space measures for food environment research. *Health & Place* 30:215–25. doi: 10.1016/j.healthplace.2014.09.007.

Das, A., S. Ghosh, K. Das, T. Basu, M. Das, and I. Dutta. 2020. Modelling the effect of area-deprivation on COVID-19 incidences: A study of Chennai Megacity, India. *Public Health* 185:266–69. doi: 10.1016/j.puhe.2020.06.011.

Desjardins, M. R., A. Hohl, and E. M. Delmelle. 2020. Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: Detecting and evaluating emerging clusters. *Applied Geography* 118:102202. doi: 10.1016/j.apgeog.2020.102202.

Durfey, S. N., A. J. Kind, W. R. Buckingham, E. H. DuGoff, and A. N. Trivedi. 2019. Neighborhood disadvantage and chronic disease management. *Health Services Research* 54:206–16. doi: 10.1111/1475-6773.13092.

Epanechnikov, V. A. 1969. Non-parametric estimation of a multivariate probability density. *Theory of Probability & Its Applications* 14 (1):153–58. doi: 10.1137/1114019.

Fattorini, D., and F. Regoli. 2020. Role of the chronic air pollution levels in the COVID-19 outbreak risk in Italy. *Environmental Pollution* 264:114732. doi:10.1016/j.envpol.2020.114732Get.

Gao, S., J. Rao, Y. Kang, Y. Liang, J. Kruse, D. Dopfer, A. K. Sethi, J. F. M. Reyes, B. S. Yandell, and J. A. Patz. 2020. Association of mobile phone location data indications of travel and stay-at-home mandates with COVID-19 infection rates in the U.S. *JAMA Network Open* 3 (9):e2020485. doi: 10.1001/jamanetworkopen.2020.20485.

Google LLC. 2020. Google COVID-19 community mobility reports. Accessed January 7, 2020. https://www.google.com/covid19/mobility/.

Gozzi, N., M. Tizzoni, M. Chinazzi, L. Ferres, A. Vespignani, and N. Perra. 2021. Estimating the effect of social inequalities on the mitigation of COVID-19 across communities in Santiago de Chile. *Nature Communications* 12 (1):2429. doi: 10.1038/s41467-021-22601-6.

Hamidi, S., S. Sabouri, and R. Ewing. 2020. Does density aggravate the COVID-19 pandemic? Early findings and lessons for planners. *Journal of the American Planning Association* 86 (4):495–509. doi: 10.1080/01944363.2020.1777891.

Harapan, H., N. Itoh, A. Yufika, W. Winardi, S. Keam, H. Te, D. Megawati, Z. Hayati, A. L. Wagner, and M. Mudatsir. 2020. Coronavirus disease 2019 (COVID-19): A literature review. *Journal of Infection and Public Health* 13 (5):667–73. doi: 10.1016/j.jiph.2020.03.019.

Helbich, M., M. H. M. Browning, and M.-P. Kwan. 2021. Time to address the spatiotemporal uncertainties in COVID-19 research: Concerns and challenges. *The Science of the Total Environment* 764:142866. doi: 10.1016/j.scitotenv.2020.142866.

Hohl, A., E. Delmelle, W. Tang, and I. Casas. 2016. Accelerating the discovery of space–time patterns of infectious diseases using parallel computing. *Spatial and Spatiotemporal Epidemiology* 19:10–20. doi: 10.1016/j.sste.2016.05.002.

Holsten, J. E. 2009. Obesity and the community food environment: A systematic review. *Public Health Nutrition* 12 (3):397–405. doi: 10.1017/S1368980008002267.

Huang, J., M.-P. Kwan, Z. Kan, M. S. Wong, C. Y. T. Kwok, and X. Yu. 2020. Investigating the relationship between the built environment and relative risk of COVID-19 in Hong Kong. *ISPRS International Journal of Geo-Information* 9 (11):624. doi: 10.3390/ijgi9110624.

Hung, W. T., A. Manandhar, and S. A. Ranasinghege. 2010. A walkability survey in Hong Kong. In *Proceedings of TRANSED 2010: 12th International Conference on Mobility and Transport for Elderly and Disabled Persons*, June 1–4, 2010, Hong Kong, China. https://trid.trb.org/view/1126939.

Kan, Z., M.-P. Kwan, M. S. Wong, J. Huang, and D. Liu. 2021. Identifying the space-time patterns of COVID-19 risk and their associations with different built environment features in Hong Kong. *The Science of the Total Environment* 772 (10):145379. doi: 10.1016/j.scitotenv.2021.145379.

Kim, J., and M.-P. Kwan. 2019. Beyond commuting: Ignoring individuals' activity-travel patterns may lead to inaccurate assessments of their exposure to traffic congestion. *International Journal of Environmental Research and Public Health* 16 (1):89. doi: 10.3390/ijerph16010089.

Kim, J., and M.-P. Kwan. 2021a. Assessment of sociodemographic disparities in environmental exposure might be erroneous due to neighborhood effect averaging: Implications for environmental inequality research. *Environmental Research* 195:110519. doi: 10.1016/j.envres.2020.110519.

Kim, J., and M.-P. Kwan. 2021b. How neighborhood effect averaging might affect assessment of individual exposures to air pollution: A study of ozone exposures

in Los Angeles. *Annals of the American Association of Geographers* 111 (1):121–40. doi: 10.1080/24694452.2020.1756208.

Kim, J., and M.-P. Kwan. 2021c. The impact of the COVID-19 pandemic on people's mobility: A longitudinal study of the U.S. from March to September of 2020. *Journal of Transport Geography* 93:103039. doi: 10.1016/j.jtrangeo.2021.103039.

Kodera, S., E. A. Rashed, and A. Hirata. 2020. Correlation between COVID-19 morbidity and mortality rates in Japan and local population density, temperature, and absolute humidity. *International Journal of Environmental Research and Public Health* 17 (15):5477. doi: 10.3390/ijerph17155477.

Kou, L., M.-P. Kwan, and Y. Chai. 2020. The effects of activity-related contexts on individual sound exposures: A time–geographic approach to soundscape studies. *Environment and Planning B: Urban Analytics and City Science*. Advance online publication. doi: 10.1177/2399808320965243.

Kraemer, M. U. G., C.-H. Yang, B. Gutierrez, C.-H. Wu, B. Klein, D. M. Pigott, L. Du Plessis, N. R. Faria, R. Li, W. P. Hanage, et al. 2020. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 368 (6490):493–97. doi: 10.1126/science.abb4218.

Kwan, M.-P. 2012. The uncertain geographic context problem. *Annals of the Association of American Geographers* 102 (5):958–68. doi: 10.1080/00045608.2012.687349.

Kwan, M.-P. 2013. Beyond space (as we knew it): Toward temporally integrated geographies of segregation, health, and accessibility: Space–time integration in geography and GIScience. *Annals of the Association of American Geographers* 103 (5):1078–86. doi: 10.1080/00045608.2013.792177.

Kwan, M.-P. 2018a. The limits of the neighborhood effect: Contextual uncertainties in geographic, environmental health, and social science research. *Annals of the American Association of Geographers* 108 (6):1482–90. doi: 10.1080/24694452.2018.1453777.

Kwan, M.-P. 2018b. The neighborhood effect averaging problem (NEAP): An elusive confounder of the neighborhood effect. *International Journal of Environmental Research and Public Health* 15 (9):1841. doi: 10.3390/ijerph15091841.

Kwan, M.-P., J. Wang, M. Tyburski, D. H. Epstein, W. J. Kowalczyk, and K. L. Preston. 2019. Uncertainties in the geographic context of health behaviors: A study of substance users' exposure to psychosocial stress using GPS data. *International Journal of Geographical Information Science* 33 (6):1176–95. doi: 10.1080/13658816.2018.1503276.

Kwok, C. Y. T., M. S. Wong, K. L. Chan, M.-P. Kwan, J. E. Nichol, C. H. Liu, J. Y. H. Wong, A. K. C. Wai, L. W. C. Chan, Y. Xu, et al. 2021. Spatial analysis of the impact of urban geometry and socio-demographic characteristics on COVID-19: A study in Hong Kong. *Science of the Total Environment* 764:144455. doi: 10.1016/j.scitotenv.2020.144455.

Lai, S., A. Farnham, N. W. Ruktanonchai, and A. J. Tatem. 2019. Measuring mobility, disease connectivity and individual risk: A review of using mobile phone data and mHealth for travel medicine. *Journal of Travel Medicine* 26 (3):1–9. doi: 10.1093/jtm/taz019.

Law, S., A. W. Leung, and C. Xu. 2020. "Third wave" of COVID-19 pandemic in Hong Kong. *Bangladesh Journal of Infectious Diseases* 7:S61–S62. doi: 10.3329/bjid.v7i00.50165.

Lee, W. D., M. Qian, and T. Schwanen. 2021. The association between socioeconomic status and mobility reductions in the early stage of England's COVID-19 epidemic. *Health & Place* 69:102563. doi: 10.1016/j.healthplace.2021.102563.

Li, M., X. Shi, X. Li, W. Ma, J. He, and T. Liu. 2019. Epidemic forest: A spatiotemporal model for communicable diseases. *Annals of the American Association of Geographers* 109 (3):812–36. doi: 10.1080/24694452.2018.1511413.

Ma, J., Y. Tao, M.-P. Kwan, and Y. Chai. 2020. Assessing mobility-based real-time air pollution exposure in space and time using smart sensors and GPS trajectories in Beijing. *Annals of the American Association of Geographers* 110 (2):434–48. doi: 10.1080/24694452.2019.1653752.

Ma, S., L. Yang, M.-P. Kwan, Z. Zuo, H. Qian, and M. Li. 2021. Do individuals' activity structures influence their $PM_{2.5}$ exposure levels? Evidence from human trajectory data in Wuhan city. *International Journal of Environmental Research and Public Health* 18 (9):4583. doi: 10.3390/ijerph18094583.

Ma, X., X. Li, M.-P. Kwan, and Y. Chai. 2020. Who could not avoid exposure to high levels of residence-based pollution by daily mobility? Evidence of air pollution exposure from the perspective of the neighborhood effect averaging problem (NEAP). *International Journal of Environmental Research and Public Health* 17 (4):1223. doi: 10.3390/ijerph17041223.

Nakaya, T., and K. Yano. 2010. Visualising crime clusters in a space-time cube: An exploratory data-analysis approach using space-time kernel density estimation and scan statistics. *Transactions in GIS* 14 (3):223–39. doi: 10.1111/j.1467-9671.2010.01194.x.

Park, Y. M., and M.-P. Kwan. 2017. Individual exposure estimates may be erroneous when spatiotemporal variability of air pollution and human mobility are ignored. *Health Place* 43:85–94. doi: 10.1016/j.healthplace.2016.10.002.

Pullano, G., E. Valdano, N. Scarpa, S. Rubrichi, and V. Colizza. 2020. Evaluating the effect of demographic factors, socioeconomic factors, and risk aversion on mobility during the COVID-19 epidemic in France under lockdown: A population-based study. *The Lancet Digital Health* 2 (12):e638–e649. doi: 10.1016/S2589-7500(20)30243-0.

Shi, X., M. Li, O. Hunter, B. Guetti, A. Andrew, E. Stommel, W. Bradley, and M. Karagas. 2019. Estimation of environmental exposure: Interpolation, kernel density Estimation, or snapshotting. *Annals of GIS* 25 (1):1–8. doi: 10.1080/19475683.2018.1555188.

St-Denis. X. 2020. Sociodemographic determinants of occupational risks of exposure to COVID-19 in Canada. *Canadian Review of Sociology/Revue Canadienne de Sociologie* 57 (3):399–452. doi: 10.1111/cars.12288.

Stoddard, S. T., A. C. Morrison, G. M. Vazquez-Prokopec, V. P. Soldan, T. J. Kochel, U. Kitron, J. P. Elder, and T. W. Scott. 2009. The role of human movement in the transmission of vector-borne pathogens. *PLoS Neglected Tropical Diseases* 3 (7):e481. doi: 10.1371/journal.pntd.0000481.

Tan, Y., M.-P. Kwan, and Z. Chen. 2020. Examining ethnic exposure through the perspective of the neighborhood effect averaging problem: A case study of Xining, China. *International Journal of Environmental Research and Public Health* 17 (8):2872. doi: 10.3390/ijerph17082872.

Thornton, L. E., J. R. Pearce, and A. M. Kavanagh. 2011. Using geographic information systems (GIS) to assess the role of the built environment in influencing obesity: A glossary. *International Journal of Behavioral Nutrition and Physical Activity* 8 (1):71. doi: 10.1186/1479-5868-8-71.

Tian, H., Y. Liu, Y. Li, C.-H. Wu, B. Chen, M. U. G. Kraemer, B. Li, J. Cai, B. Xu, Q. Yang, et al. 2020. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* 368 (6491):638–42. doi: 10.1126/science.abb6105.

van Doremalen, N., T. Bushmaker, D. H. Morris, M. G. Holbrook, A. Gamble, B. N. Williamson, A. Tamin, J. L. Harcourt, N. J. Thornburg, S. I. Gerber, et al. 2020. Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *New England Journal of Medicine* 382 (16):1564–67. doi: 10.1056/NEJMc2004973.

Viboud, C., O. N. Bjørnstad, D. L. Smith, L. Simonsen, M. A. Miller, and B. T. Grenfell. 2006. Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science* 312 (5772):447–51. doi: 10.1126/science.1125237.

Wang, J., and M.-P. Kwan. 2018. An analytical framework for integrating the spatiotemporal dynamics of environmental context and individual mobility in exposure assessment: A study on the relationship between food environment exposures and body weight. *International Journal of Environmental Research and Public Health* 15 (9):2022. doi: 10.3390/ijerph15092022.

Wang, J., M.-P. Kwan, and Y. Chai. 2018. An innovative context-based crystal-growth activity space method for environmental exposure assessment: A study using GIS and GPS trajectory data collected in Chicago. *International Journal of Environmental Research and Public Health* 15 (4):703. doi: 10.3390/ijerph15040703.

Wang, V., and T. May. 2021. In "coffin homes" and "cages," Hong Kong lockdown exposes inequality. *New York Times*, January 26. Accessed May 10, 2021. https://www.nytimes.com/2021/01/26/world/asia/hong-kong-coronavirus-lockdown-inequality.html?_ga=2.185757673.1763379367.1617417488-133072336.1607935695.

Wesolowski, A., N. Eagle, A. J. Tatem, D. L. Smith, A. M. Noor, R. W. Snow, and C. O. Buckee. 2012. Quantifying the impact of human mobility on malaria. *Science* 338 (6104):267–70. doi: 10.1126/science.1223467.

Willberg, E., O. Järv, T. Väisänen, and T. Toivonen. 2021. Escaping from cities during the COVID-19 crisis: Using mobile phone data to trace mobility in Finland. *ISPRS International Journal of Geo-Information* 10 (2):103. doi: 10.3390/ijgi10020103.

World Health Organization. 2021. Coronavirus disease (COVID-19): Weekly epidemiological, update—30 March 2021. https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19—31-march-2021.

Wu, X., R. C. Nethery, M. B. Sabath, D. Braun, and F. Dominici. 2020. Air pollution and COVID-19 mortality in the United States: Strengths and limitations of an ecological regression analysis. *Science Advances* 6 (45):eabd4049. doi: 10.1126/sciadv.abd4049.

Xiong, C., S. Hu, M. Yang, W. Luo, and L. Zhang. 2020. Mobile device data reveal the dynamics in a positive relationship between human mobility and COVID-19 infections. *Proceedings of the National Academy of Sciences of the United States of America* 117 (44):27087–89. doi: 10.1073/pnas.2010836117.

Xiong, Y., Y. Wang, F. Chen, and M. Zhu. 2020. Spatial statistics and influencing factors of the COVID-19 epidemic at both prefecture and county levels in Hubei Province, China. *International Journal of Environmental Research and Public Health* 17 (11):3903. doi: 10.3390/ijerph17113903.

Yu, X., M. S. Wong, M.-P. Kwan, J. E. Nichol, R. Zhu, J. Heo, P. W. Chan, D. C. W. Chin, C. Y. T. Kwok, and Z. Kan. 2021. COVID-19 infection and mortality: Association with PM2.5 concentration and population size—An exploratory study. *ISPRS International Journal of Geo-Information* 10 (3):123. doi: 10.3390/ijgi10030123.

Zhang, L., S. Zhou, M.-P. Kwan, F. Chen, and R. Lin. 2018. Impacts of individual daily greenspace exposure on health based on individual activity space and structural equation modeling. *International Journal of Environmental Research and Public Health* 15 (10):2323. doi: 10.3390/ijerph15102323.

Zhao, P., M.-P. Kwan, and S. Zhou. 2018. The uncertain geographic context problem in the analysis of the relationships between obesity and the built environment in Guangzhou. *International Journal of Environmental Research and Public Health* 15 (2):308. doi: 10.3390/ijerph15020308.

JIANWEI HUANG is a PhD Student in the Institute of Space and Earth Information Science at the Chinese University of Hong Kong, Shatin, Hong Kong, China. E-mail: Jianwei.Huang@link.cuhk.edu.hk. His research interests include human mobility, environmental health, and GIScience.

MEI-PO KWAN is Choh-Ming Li Professor of Geography and Resource Management and Director of the Institute of Space and Earth Information Science at the Chinese University of Hong Kong, Shatin, Hong Kong, China. E-mail: mpk654@gmail.com. Her research interests include environmental health, human mobility, sustainable cities, transport and health issues in cities, and GIScience.