

Exploring year-to-year spatiotemporal changes in cycling patterns for bike-sharing system in the pre-, during and post-pandemic periods

Xiaoying Shi ^{a,b}, Junjie Zhao ^a, Jiaming He ^a, Haitao Xu ^{a,b,*}

^a School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou, China

^b Key Laboratory of Discrete Industrial Internet of Things of Zhejiang Province, Hangzhou, China

ARTICLE INFO

Keywords:

Bike-sharing system
COVID-19
Spatiotemporal clustering
Network analysis
Visualization

ABSTRACT

The bike-sharing system plays an important role during the COVID-19 pandemic due to its benefit of reducing the risk of infection. Previous studies have employed bike trip datasets to analyze the impact of COVID-19 on spatiotemporal cycling mobility, which was limited to a short pandemic period. An understanding of the long-term impact of the pandemic on human travel patterns is urgently needed. This paper presents a framework for investigating the long-term spatiotemporal change in cycling patterns in the pre-, during and post-pandemic periods. A spatiotemporal clustering algorithm is combined with spatiotemporal visualization, network analysis, and important location identification to explore latent mobility patterns. The strength of the proposed framework is that it can model a unified representation of cycling networks across different years, and provide insightful spatiotemporal patterns that allow easy interpretation. The experiments are conducted on five years of data collected from the ‘Citibike’ system in New York City. The results demonstrate the efficiency of the proposed approach in exploring the evolutionary cycling patterns of how people respond to the COVID-19 pandemic and are beneficial for various stakeholders.

1. Introduction

The COVID-19 pandemic swept through the world in early 2020, which had significant impacts on human mobility and public transportation (Gkiotsalitis & Cats, 2021; THu et al., 2021). Due to the fear of infection risk, the urban transport system is confronted with unprecedented scenarios. Bike-sharing (Fishman, 2016; DeMaio, 2009) is a healthy and environment-friendly travel mode, to solve the “last-mile problem”. Compared to other travel modes, such as bus or subway, bike-sharing has the advantage of preserving physical distances to reduce infection risks (Jobe & Griffin, 2021; Kim & Cho, 2022; Teixeira & Lopes, 2020). The bike-sharing system provides a resilient transportation option for citizens (Hu et al., 2021; Wang & Noland, 2021), and is an ideal choice to meet people’s travel needs during the pandemic.

To learn about the impact of the pandemic on bike-sharing systems, questionnaire surveys were conducted to identify the motivations for using shared bikes, the influencing factors, and changes in user habits during the pandemic (Nikiforidis et al., 2020; Teixeira et al., 2021; Bergantino et al., 2021). However, the survey samples are limited. The analysis results refer to the selected respondents and are influenced by

people’s subjective opinions on the questions.

Thanks to the bicycle trip data collected by bike-sharing systems, it is possible to study the cycling behavior of massive users. Researchers have used the datasets collected from different cities, to predict daily bike rentals (Kim, 2021; Mehdizadeh Dastjerdi & Morency, 2022), and to analyze the impact of COVID-19 on usage patterns (Kubáľák et al., 2021; Y Chen et al., 2022; Padmanabhan et al., 2021), influencing factors (Bustamante et al., 2022; Yang et al., 2022; Sung, 2023; Chibwe et al., 2021), environmental benefits (Shang et al., 2021), and spatiotemporal mobility (Hu et al., 2021; Pase et al., 2020; Xin et al., 2022; Chai et al., 2021; Li et al., 2021; Bi et al., 2022). Despite the growing research interest in the spatiotemporal mobility affected by the pandemic, previous studies have analyzed only a short period of time. We lack a complete understanding of the long-term influence of COVID-19 on spatiotemporal mobility for large-scale bike-sharing systems. In this study, we address the following three research questions: 1) What is the difference in spatiotemporal mobility in the pre-, during and post-pandemic periods? 2) How do the properties of bicycle networks evolve in the pre-, during and post-pandemic periods? 3) Where are the hotspots for cycling in the pre-, during and post-pandemic periods?

The basis for answering the above questions is to get a unified

* Corresponding author at: School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou, China.

E-mail address: xuhaitao@hdu.edu.cn (H. Xu).

representation of bicycle networks for different time periods. For large-scale bike sharing systems, the station-based networks are difficult to visualize and compare. Stations aggregated by community detection methods would generate a small number of spatial clusters and cannot support a detailed analysis between regions (Hu et al., 2021; Pase et al., 2020; Xin et al., 2022). In addition, the division of time periods is also critical for analyzing the spatiotemporal patterns. Previous studies divided dates into different categories based on workday-weekend segmentation (J Corcoran et al., 2014; Wu et al., 2018), human experience (Hu et al., 2021; Li et al., 2021; Song et al., 2022), or selected time intervals (Shang et al., 2021; Pase et al., 2020; Xin et al., 2022) (Bi et al., 2022), which ignored the actual travel characteristics. *k*-segmentation could only obtain continuous phases with similar temporal patterns (Chai et al., 2021). However, the pandemic-related policies significantly changed people's daily behavior, and different dates may have different patterns. The station clustering algorithms that simultaneously consider spatiotemporal features (Li et al., 2015; Galvani, Torti, & Menafoglio, 2020) could not cope with the situation of increasing number of stations. The long-term cycling patterns cannot be compared if the same stations are divided into different spatial clusters in different years. Therefore, we need to develop a method that can not only model a unified representation of cycling networks across different years, but also automatically group the dates according to the underlying cycling characteristics, which will help to analyze the spatiotemporal patterns more accurately.

In this study, we analyze year-to-year spatiotemporal changes in shared bicycle user behavior in the pre-, during and post-pandemic periods using five years of trip data in New York City (NYC). To the best of our knowledge, this is the first study to reveal the impact of the pandemic on the spatiotemporal mobility of bike-sharing over such a long period. The contributions of this study are threefold: 1) A framework for exploring the long-term change in spatiotemporal cycling patterns influenced by the pandemic is developed. 2) A spatiotemporal clustering algorithm is proposed to obtain a unified representation of cycling patterns across different years. 3) A five-year longitudinal analysis is conducted for NYC. The evidence-based results are insightful for providing suggestions on how to respond to the unprecedented event for different stakeholders.

The rest of the paper is organized as follows. Section 2 reviews related studies on the usage patterns of bike-sharing systems. Section 3 describes the datasets used. Section 4 presents the methodology in detail. Section 5 analyzes the experimental results. Section 6 discusses the research findings and policy implications. Section 7 draws conclusions.

2. Literature review

As bike-sharing becomes prevalent around the world, it has attracted considerable interest from researchers. Research on bike-sharing includes bike flow prediction (Zi et al., 2021; Seng et al., 2021), bike scheduling (Legros, 2019), electric fence planning (Zhang et al., 2019; Shi et al., 2023), and usage pattern analysis (Shi et al., 2018; Caulfield et al., 2017), etc. Much research has been done to understand the spatiotemporal patterns of shared bike usage (Zhou, 2015; Corcoran et al., 2014; Li et al., 2021; X Chen et al., 2022).

In 2020, the outbreak of the COVID-19 pandemic has significantly disrupted urban mobility, including bike sharing. Researchers have shown great interest in studying the impact of the pandemic on bike sharing. Hu et al. (2021) found that the proportion of commuting trips was significantly lower during the pandemic and the trend in bicycle use followed an "increase-decrease-rebound" pattern in Chicago. For the bike-sharing system in NYC, paper (Teixeira & Lopes, 2020; Wang & Noland, 2021) showed evidence of a modal shift from subway users to shared bike users, and found that the bike-sharing system was more resilient than the subway system. Pase et al. (2020) studied four weeks of workdays in March 2020, and found that Citi Bike's critical workers program played an important role in efforts to make cycling more

popular among commuters. Wealthier neighborhoods in Manhattan are better able to maintain social distance than poorer areas with a high percentage of low-wage workers. Xin et al. (2022) found that previous cycling hotspots in NYC were weakened or disappeared during the pandemic, the number of cycling communities decreased and the area of community coverage increased. For the London Cycle Hire (LCH) system, paper (Chibwe et al., 2021; Li et al., 2021) examined the impact of pandemic-related policies on bicycle use. The results showed that demand was most dispersed during the COVID-19 lockdown days. The UK's lockdown policy led to an immediate decrease in the LCH use, while the first lockdown ease had no statistically significant immediate impact. Song et al. (Song et al., 2022) investigated cycling patterns in Singapore before, during and after the local authorities imposed lockdown measures. They found a 150% increase in total ridership during the lockdown compared to the pre-pandemic winter period. Biking mobility graphs became more locally clustered and polycentric as the pandemic progressed. Using Pittsburgh as a case study, Qin & Karimi (2023) examined the changes in the spatiotemporal dynamics of bicycle use in a small-scale bike-sharing system. The results demonstrated the resilience and critical role of bike-sharing during the pandemic, consistent with previous observations in large-scale systems. Paper (Kim & Cho, 2022; Sung, 2023; Jiao et al., 2022) examined the causal relationships between bike-sharing and various influencing factors during the pandemic situation in Seoul, Korea, and verified that bike-sharing was a disease-resilient travel mode. Changes in the ridership and usage patterns before and during the pandemic have also been studied in other cities, such as Košice (Kubalák et al., 2021), Washington D.C. (Y Chen et al., 2022) and Barcelona (Bustamante et al., 2022). Padmanabhan et al. (Padmanabhan et al., 2021) used the least squares regression model to explore the marginal effects of COVID-19 cases on the number of shared bike trips in NYC, Boston, and Chicago.

The existing studies also found that the pandemic could change the places that people tend to visit. In Beijing, during the pandemic, people were less likely to ride shared bikes to places that were popular before (Shang et al., 2021). Reductions in mobility near subway stations, high-tech companies, and shopping plazas were higher than other POI categories due to the quarantine restrictions and "work from home" messages. Li et al. (2021) found that the proportion of home, park and grocery activities increased during the lockdown period in Zurich. In NYC, overall bicycle use decreased during the outbreak. However, residential areas, supermarkets, parks, and hospitals were less affected (Chai et al., 2021; Bi et al., 2022). In Seoul, bike rentals for recreation and exercise rather than transportation increased during the pandemic (Kim, 2021; Sung, 2023).

In Table 1, we list the literature that has studied the usage patterns of bike-sharing systems during pandemic periods. Specifically, we list the study periods and further specify whether they include pre-pandemic years (before 2020), the year of the pandemic outbreak (2020), and the post-pandemic years (after 2020). As shown in Table 1, most studies cover several months before and during the pandemic outbreak. Only a few studies cover the post-pandemic years. Although the studies in Chen et al. (2022); Sung (2023); Chibwe et al. (2021) cover more than 3 years, they used daily aggregated data to analyze the travel intensity and factors influencing bicycle use, which loses the detailed information of spatiotemporal mobility. Qin and Karimi (2023) investigated the changes in the spatiotemporal dynamics of shared bicycle use from 2019 to 2021 in Pittsburgh. The system scale is relatively small, with only 113 stations. The long-term impact of the pandemic on the spatiotemporal dynamics of large-scale bike-sharing systems has not been studied.

The analysis of the spatiotemporal characteristics is based on two aspects: the definition of the spatial unit and the temporal division. The spatial units were defined as stations or communities (Hu et al., 2021; Pase et al., 2020; Xin et al., 2022), postcode areas (Li et al., 2021; Bi et al., 2022), or grids (Shang et al., 2021). For the temporal division, a workday-weekend segmentation is commonly used for normal years (Corcoran et al., 2014; Wu et al., 2018). For the pandemic years,

Table 1

Studies on bike-sharing systems during the COVID-19 pandemic.

Research study	Location	System description	Study period	Include pre-pandemic years	Include year of pandemic outbreak	Include post-pandemic years
Hu et al. (2021)	Chicago	608 stations, over 5800 bikes	2019.3.11–2019.7.31 2020.2.1–2020.7.31	Yes	Yes	No
Teixeira and Lopes (2020)	NYC	890 stations, 14,500 bikes	2019.02–2019.03 2020.02–2020.03	Yes	Yes	No
Wang and Noland (2021)	NYC	about 1000 stations, 15,000 bikes	2019.01–2019.09 2020.01–2020.09	Yes	Yes	No
Pase et al. (2020)	NYC		2020.03	No	Yes	No
Xin et al. (2022)	NYC		2019.01–2019.04 2020.01–2020.04	Yes	Yes	No
Bi et al. (2022)	NYC	1027 stations, 14,500 bikes	2019.01–2020.06	Yes	Yes	No
Chibwe et al. (2021)	London	780 stations, 11,500 bikes	2012.01–2020.06	Yes	Yes	No
H Li et al. (2021)	London	761 stations	2019.01–2020.06	Yes	Yes	No
Song et al. (2022)	Singapore	14,000 bikes	2019.12–2020.07	Yes	Yes	No
Qin and Karimi (2023)	Pittsburgh	113 stations, over 500 bikes	2019.01–2021.12	Yes	Yes	Yes
Sung (2023)	Seoul	2500 stations, 37,500 bikes	2017.01–2021.12	Yes	Yes	Yes
Kim (2021)	Seoul	more than 1500 stations	2019.01–2020.12	Yes	Yes	No
Kim and Cho (2022)	Seoul		2018.01–2020.12	Yes	Yes	No
Jiao et al. (2022)	Seoul	2083 stations	2019.03–2020.06	Yes	Yes	No
Kubařák et al. (2021)	Košice	91 stations, 1000 bikes	2019.01–2019.12 2020.01–2020.12	Yes	Yes	No
Y Chen et al. (2022)	Washington D. C.	627 stations, around 5400 bikes	2019.01–2021.12	Yes	Yes	Yes
Bustamante et al. (2022)	Barcelona	500 stations, about 7000 bikes	2019.01–2020.12	Yes	Yes	No
Shang et al. (2021)	Beijing		2020.1.14–2020.3.10	No	Yes	No
Chai et al. (2021)	Beijing		2019.03–2020.03	Yes	Yes	No
Li et al. (2021)	Zurich	153 stations	2020.2.15–2020.4.14	No	Yes	No
Padmanabhan et al. (2021)	NYC, Boston, Chicago	NYC: nearly 900 stations; Chicago: 600 stations; Boston: 330 stations	2019.10–2020.05	Yes	Yes	No
Our study	NYC	2018: 762 stations; 2019: 824 stations; 2020: 994 stations; 2021: 1352 stations; 2022: 1683 stations	2018.01–2022.12	Yes	Yes	Yes

researchers divided the study period into different phases based on human experience (Hu et al., 2021; Li et al., 2021; Song et al., 2022), selected time intervals (Shang et al., 2021; Pase et al., 2020; Xin et al., 2022; Bi et al., 2022), or k-segmentation (Chai et al., 2021). Based on the spatiotemporal division, bicycle networks were constructed (Hu et al., 2021; Shang et al., 2021; Pase et al., 2020; Xin et al., 2022; Li et al., 2021; Bi et al., 2022). The networks were visualized and the associated network indicators were calculated to quantify the transformation of user behavior. Since we want to analyze the evolution of cycling over a long period, while the number of stations increases every year, a unified representation of the network structure is important. The number of communities or postcode areas is small, which does not support a detailed analysis between regions. The bidirectional flows of station-based networks or grid-based networks are difficult to visualize clearly.

In summary, more research is needed to fully understand how spatiotemporal patterns evolved in the pre-pandemic, pandemic, and post-pandemic periods. Current methods could not support the analysis of the long-term impact of the pandemic on cycling patterns for large-scale systems with increasing numbers of stations. Our study, listed in the last row of Table 1, aims to fill this gap and provide a starting point for future research.

3. Data

3.1. Data description

Our study focuses on Citi Bike in NYC, which is also the largest bike-sharing system in the United States. The system was launched in May 2013 with 322 stations and has been continuously expanded every year. We obtained Citi Bike trip data from their official website (<https://www.citibikenyc.com/system-data>).

The data records include trip information such as start time, end time, start station, end station, and user status. We collected trip data from two normal years (2018, 2019) and three years with COVID-19 (2020 to 2022) to investigate the cycling behavior influenced by COVID-19.

To help analyze the relationship between bike ridership and COVID-19, we obtained the data on confirmed cases in NYC from the website (<https://github.com/nychealth/coronavirus-data>). The first laboratory case was confirmed on February 29, 2020, marking the beginning of the COVID-19 outbreak in NYC. Since then, the number of infections has continued to rise. The U.S. government took a number of actions in response to the pandemic. President Trump declared a national emergency on March 13, 2020. The governor of New York declared a lockdown policy from March 22 until the end of May, followed by a four-phase reopening plan since June 2020 ([Timeline of COVID-19 in New York City, 2021](#)).

In addition, we used check-in data from Foursquare (Yang et al., 2014) to construct the POI dataset. Foursquare provides location-based social media services for people to record information about POIs. A POI that is frequently checked in by many users is popular and attractive. The distribution of POIs around a location indicates the relationship between the land-use characteristic of the location and potential human activities (Shi et al., 2019; Bao et al., 2017). The POI dataset contains 227,428 check-in records. Each record contains the user ID, venue ID, venue name, GPS coordinates of the check-in location, and timestamp. The venue categories are generalized to ten root categories, i.e., Arts & Entertainment, Travel & Transport, Nightlife Spot, which provides clues for inferring individuals' visiting purposes.

3.2. Data preprocessing

Bike trip data preprocessing. Since we want to compare users' cycling patterns over five years, the stations need to be presented in a consistent way. However, we noticed that the station information has changed since 2021. We first unified the station representation. A station is given the same ID for different years according to its name and location. We also removed some outliers. The records with trip duration less than 2 min and more than 6 h are deleted because such records may not reflect real mobility. Besides, we set the maximum cycling speed as 30 km/h to exclude trips that are too fast.

POI data preprocessing. The number of check-ins for the "Event" category is 0. The related locations of "Food" category exist in workplaces, residential areas, and entertainment areas, which cannot help us determine the trip purposes. Therefore, we removed these two categories. To analyze the visit purposes of bike stations, we defined the influence area of a station as a 250 m circular buffer zone, as in previous studies (Bao et al., 2017; Faghili-Imani & Eluru, 2016). For a station, we extracted all POIs within its influence area, and aggregated the corresponding check-in records according to the POI locations.

4. Method

This study proposes a data analysis framework that combines spatiotemporal clustering algorithm with visualization, network analysis, and important location identification. We propose a novel spatiotemporal clustering algorithm that can explore the evolutionary patterns of cycling over time. Compared with the previous spatiotemporal aggregating/clustering methods for bike-sharing systems (Pase et al., 2020;

Xin et al., 2022; Li et al., 2021; Li et al., 2015; Galvani, Torti, & Menafoglio, 2020), our method can incrementally add newly constructed stations to the existing station clusters to support the unified pattern comparison over long periods. Moreover, our method can automatically group the dates into different categories according to the underlying cycling characteristics, which can better reflect the complex change of patterns in response to pandemic under the government policies.

Fig. 1 illustrates the pipeline of our method. Based on the pre-processed datasets, a spatiotemporal clustering algorithm is performed to obtain the spatiotemporal clustered networks containing the latent cycling patterns. Then, we analyze the cycling patterns by using spatiotemporal visualization, network analysis, and important station cluster exploration, to answer the three research questions, respectively.

4.1. Spatiotemporal clustering

The spatiotemporal clustering algorithm can group the bike trip data into spatiotemporal clusters with similar cycling patterns, including four steps.

Step 1: Incremental station clustering. We consider spatial clusters as the basic spatial units, and propose an incremental station clustering algorithm to integrate the newly constructed station every year. We first use the k -means++ algorithm (Arthur & Vassilvitskii, 2007) to group the stations for year y . The input feature of each station is its latitude and longitude coordinates, and the Euclidean distance between station locations is used as the similarity criterion. Thus, we obtain the station clusters $SC^y = \{sc_j^y\} (1 \leq j \leq KS_y)$. KS_y denotes the

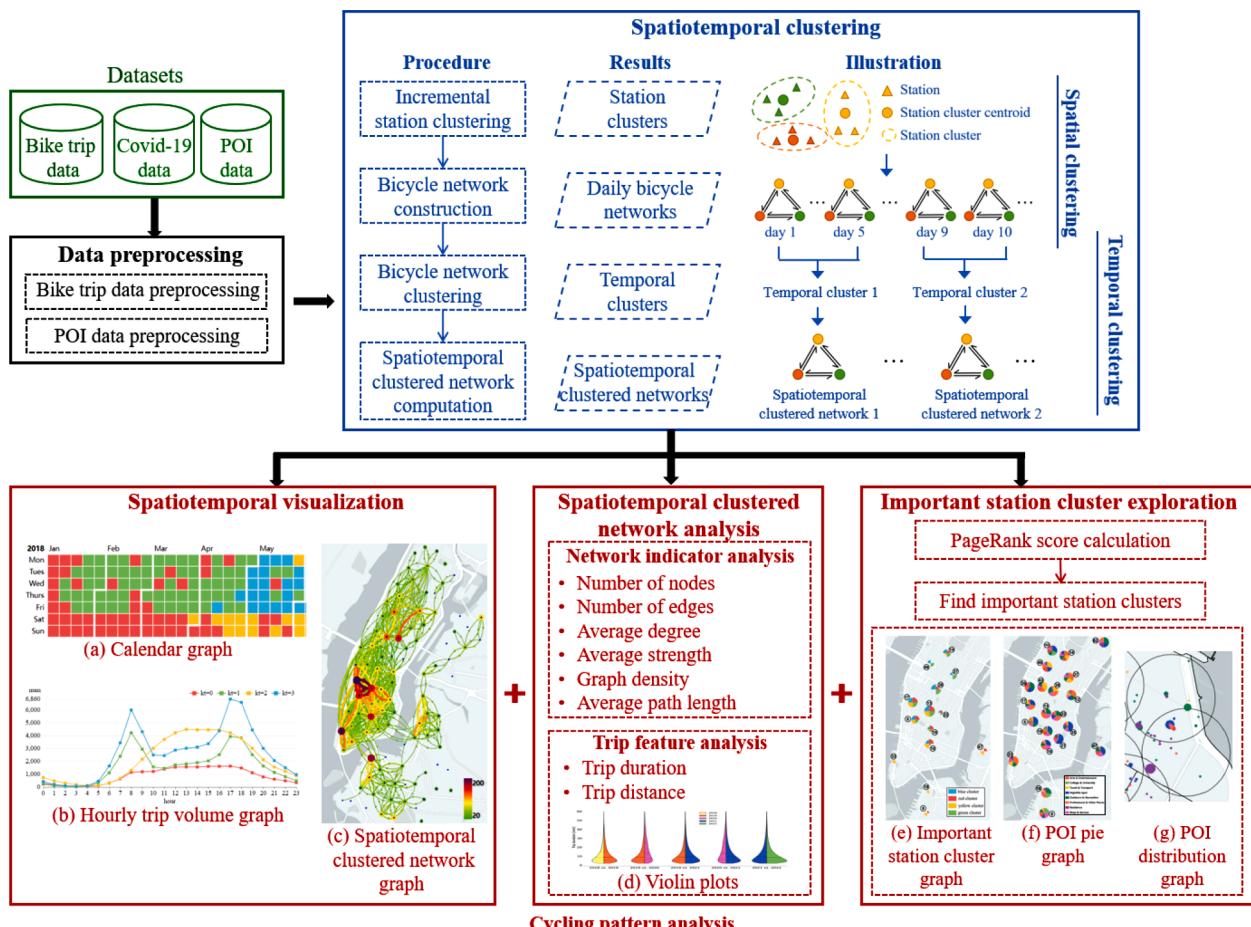


Fig. 1. Method pipeline.

number of station clusters in year y , while sc_j^y represents the j -th station cluster in year y . The centroid of a station cluster is calculated as the geographic center of all stations belonging to that cluster. For year $y + 1$, the station clusters that appeared in year y remain unchanged. For a new station i , the following sub-steps are performed iteratively to determine whether it belongs to the existing station clusters or a new station cluster. 1) Compute the distance between i and the centroid of the nearest existing station cluster sc_j . 2) For all existing station clusters, compute the longest distance between the stations and the centroids of subordinate station clusters, and denote it as d_{max} . If the distance is less than a threshold $\alpha \cdot d_{max}$, then i is assigned to station cluster sc_j . Otherwise, i becomes the centroid of a new station cluster. 3) Update the centroids of all station clusters.

Step 2: Daily bicycle network construction. Based on the station clusters, the daily bicycle networks are constructed. The daily bicycle network on date d in year y is defined as $G_d^y = (V^y, E^y, A_d^y)$. The node set V^y is equal to the station clusters SC^y identified in step 1. $E^y \subseteq V^y \times V^y$ denotes the edge set. A_d^y is the adjacency matrix, whose element $A_d^y(sc_i, sc_j)$ denotes the number of flows departing from station cluster sc_i and arriving at sc_j on date d in year y . The days in the same year are processed uniformly, and we obtain a sequence of daily bicycle networks for each year, which encodes the evolution of cycling patterns along the timeline.

Step 3: Bicycle network clustering. To divide the days in a year into different categories, we group the daily bicycle networks with similar features for each year. Since a daily bicycle network is uniquely characterized by its trip distribution, the magnitudes of the network edges are used as the features. Specifically, for a given day, the adjacency matrix A_d^y is transformed into a one-dimensional feature vector, and the feature vectors of different days are stacked up to form a high-dimensional feature matrix. We employ the t-SNE algorithm (Van der Maaten & Hinton, 2008), a nonlinear dimensionality reduction technique, to obtain a feature matrix in 2D format. t-SNE algorithm performs better on nonlinear data than other dimensionality reduction methods such as Principal Component Analysis (PCA) (Anowar et al., 2021). By applying the t-SNE algorithm to the feature matrix, the feature vector of a station cluster is mapped into a 2D space, where station clusters with similar features remain closer to each other and dissimilar station clusters are modeled as distant with high probability. Then, the k -means++ algorithm is used to cluster the feature vectors. After clustering, the daily bicycle networks (days) for each year are divided into KT categories. The days in the same category form a temporal cluster.

Step 4: Compute the spatiotemporal clustered network. The spatiotemporal clustered network represents the flow relationship for each temporal cluster. The node set and edge set of a spatiotemporal clustered network are equal to the daily bicycle network in that temporal cluster, while the adjacency matrix is the average adjacency matrix of the days belonging to that temporal cluster. The diagonal elements of the adjacency matrix denote the intra-cluster flows in the temporal cluster. Other elements convey the inter-cluster flows between the station clusters.

4.2. Spatiotemporal visualization

To explore the time-varying characteristics and spatial distribution of trips, we design three visual graphs to show the spatiotemporal clustering results.

The calendar graph (Fig. 1(a)) shows the spatiotemporal clustering results in the temporal domain. A day is represented by a grid in the graph. The color of the grid indicates the subordinate temporal cluster.

The hourly trip volume graph (Fig. 1(b)) shows the hourly trip numbers for different temporal clusters. For each temporal cluster, the days belonging to that cluster are obtained. Then, the average hourly rental numbers of these days are calculated and represented by a

polyline.

The spatiotemporal clustered network graph (Fig. 1(c)) shows the flow distribution of the spatiotemporal clustered network on the map. A blue dot represents the centroid of a station cluster, whose outer ring indicates the intra-cluster flows. A curve with an arrow indicates the inter-cluster flows between two station clusters. The number of intra-cluster or inter-cluster flows is double-encoded by line width and a gradient color scheme. A purple and thick outer ring or curve represents a large volume flow, while a green and thin outer ring or curve represents a small volume flow. To avoid visual clutter, users can define a threshold to filter out unimportant flows. Only flows with volumes greater than the threshold will be displayed on the map.

4.3. Spatiotemporal clustered network analysis

After obtaining the spatiotemporal clustered networks, we use network indicators (Li et al., 2021; Wu & Kim, 2020) to quantify the characteristics of these clustered networks. The six network indicators used are described below:

- Number of nodes (N_N) represents the number of station clusters with cycling flows, including inflows or outflows.
- Number of edges (N_E) represents the number of edges whose edge weights are greater than 0.
- Average degree (Avg_deg). The degree of a node refers to the number of edges connected to it, which is the sum of in-degree and out-degree. Avg_deg is the average value of degree per node, calculated as $(N_E \times 2)/N_N$, reflecting the connectivity of the network.
- Average strength (Avg_str). The strength of a node is defined as the sum of the weights of all edges connected to it. Avg_str is the average value of strength for all nodes in the network.
- Graph density (G_den) is the ratio of the edges present in a network to the maximum number of edges the network can contain. G_den is calculated as $N_E/(N_N \times (N_N - 1))$, which is also equal to $Avg_deg / ((N_N - 1) \times 2)$. It measures the sparsity and density of edges in a network.
- Average path length (Avg_pLen) is the average number of steps along the shortest paths for all pairs of network nodes (Watts & Strogatz, 1998). It measures the efficiency of information transport in a network. For a bike-sharing network, the accessibility between nodes is measured by the distance between nodes. We define Avg_pLen as:

$$Avg_pLen = \frac{1}{N_N \times (N_N - 1)} \sum_{i \neq j} d_{ij} \quad (1)$$

where d_{ij} is the shortest cycling distance from node (station cluster) i to node j . If node i cannot be reached from node j , then $d_{ij} = \infty$.

In addition, we also analyze the trip features (trip duration and trip distance) for different spatiotemporal clustered networks. The average trip duration and trip distance are calculated for each network. Violin plots (Fig. 1(d)), a combination of boxplot and kernel density estimation (Li et al., 2021), are used to compare the distributions of trip features for different spatiotemporal clustered networks. The solid horizontal lines represent the medians, while the dashed lines represent the quartiles. The kernel density curves depict the trip feature distributions of the networks. Since each spatiotemporal clustered network contains different dates, the corresponding statistical trip features are normalized by the number of dates, to avoid the influence of the number of dates on the statistical results.

4.4. Important station cluster exploration

Finally, we employ the PageRank algorithm to rank the importance of station clusters for different spatiotemporal clustered networks, and

visualize the POI attributes of the highly ranked station clusters.

Although originally used to measure the importance of web pages, the PageRank algorithm has been applied to network analysis in recent years due to its superior performance (Chen et al., 2022; Xu et al., 2017). Compared to other centrality metrics (e.g., degree and betweenness) for evaluating network nodes, PageRank collectively assesses the importance of nodes by counting the number and quality of links to them (Brin & Page, 1998). The PageRank score for a station cluster i is formulated as:

$$PR_i = \frac{1-d}{KS} + d \sum_{j \in B_i} \frac{PR_j}{L_j} \quad (2)$$

where d is a damping factor with a default value of 0.85. KS is the number of station clusters. B_i is the set of station clusters with bike flows from station cluster i , and L_j is the sum of trips out from station cluster j . The underlying assumption is that more bikes will drop off at more important station clusters. A station cluster with a large PageRank score reflects its higher importance.

For each spatiotemporal clustered network, we adopt the PageRank algorithm to rank the importance of station clusters, and find the top 10 important station clusters. Then, we use three graphs to understand the visit purposes of these station clusters.

The important station cluster graph (Fig. 1(e)) shows the location and importance of the important station clusters in different spatiotemporal clustered networks for one year. Each station cluster is represented by a pie chart and divided into four sectors corresponding to the spatiotemporal clustered networks. The radius of the sector encodes the importance of the station cluster. The first ranked station cluster has the largest sector radius, while the 10th ranked station cluster has the smallest sector radius. The POI pie graph (Fig. 1(f)) shows the proportions of different POI categories for all important station clusters over five years. A station cluster is also represented by a pie chart. The division of the pie chart is related to the number of check-in records in the influence area of all stations in a station cluster. The POI distribution graph (Fig. 1(g)) shows the geographic distribution of POIs in a station cluster. The POIs in the 250 m buffer around the stations are shown. A dot represents a POI category. The position and color of the dot indicate the location and category of the POI, respectively. The size of the dot is proportional to the number of check-ins at that location.

5. Results

5.1. Statistical and station clustering results

Fig. 2 shows the monthly rental numbers for five years. We can see that the rental numbers increase from 2018 to 2019. Meanwhile, in April 2020, the rental number decreased dramatically and then gradually increased. In September 2020, the rental number reached the same level as in 2019. The rental numbers are relatively higher in 2021 and 2022, except for February 2021.

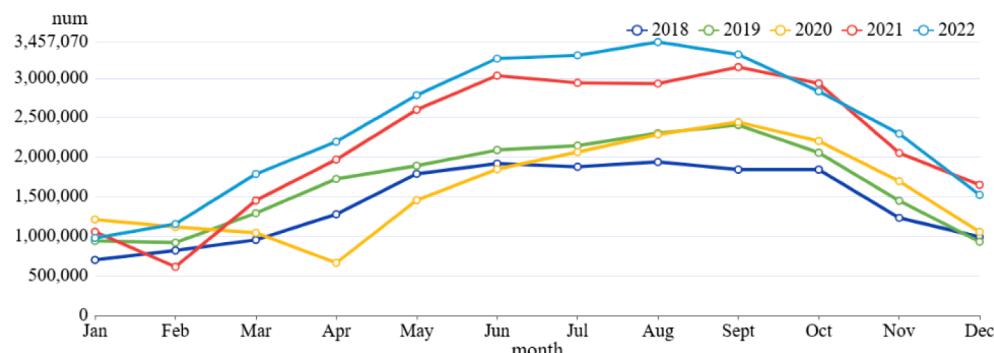


Fig. 2. The monthly rental numbers for five years.

Fig. 3 shows the daily number of confirmed COVID-19 cases in NYC. We find that the pandemic has three peaks, from mid-March 2020 to April 2020, from December 2020 to March 2021, and from December 2021 to January 2022.

Fig. 4 shows the incremental station clustering results for 2020 and 2021. We set α as 0.9 and $KS_{2018}=100$ in our experiments. Fig. 4(a) shows the stations in 2021. The gray dots represent the existing stations in 2020, while the blue and red dots represent the newly added stations in 2021. The blue dots indicate the stations merged into the station clusters of 2020, while the red dots indicate the stations identified as new station clusters in 2021. We can see that the stations of 2020 close to the stations of 2021 are merged into existing station clusters, while the stations of 2021 far from the stations of 2020 form new station clusters. Fig. 4(b) and Fig. 4(c) show the station clustering results of 2020 and 2021, respectively. The stations in a station cluster are rendered by the same color. We finally obtain 100, 104, 113, 144, and 162 station clusters from 2018 to 2022, respectively.

5.2. Spatiotemporal cycling pattern analysis

We have compared the spatiotemporal clustering results for different KT , and find that the latent patterns are clear when $KT=4$. We discuss the detailed results below.

Fig. 5 shows the spatiotemporal clustering results for 2018. As seen from the calendar graph in Fig. 5(a), the dates are divided into four temporal clusters: warm workdays (blue cluster), cold workdays (green cluster), warm weekends (yellow cluster), and bad weather days (red cluster). The dates in the green and blue clusters contain most workdays, and the hourly trip demand shows morning and evening peaks (Fig. 5(b)). The magnitude of the blue cluster is larger than that of the green cluster, indicating that people ride more on warmer days. The dates in the yellow cluster include warmer weekends and some holidays, such as Memorial Day (May 28), Independence Day (July 4), and Labor Day (September 3). The flow numbers during noon and afternoon are relatively higher. The red cluster includes cold weekends and some workdays with inclement weather, and the hourly trip volumes are relatively low.

We further observe the spatiotemporal clustered networks for different temporal clusters in Fig. 5(c)-Fig. 5(f). To highlight important flows, the flows with magnitudes less than 20 are not shown. The networks for the green and blue clusters show multicentric structures Fig. 5(c), Fig. 5(d). The dominant centers include station clusters 13, 6, 17, 38, 37 and 14. We will analyze the attributes of important station clusters in Section 5.4. The spatiotemporal clustered network for the red cluster (Fig. 5(e)) has relatively smaller flow magnitudes. The primary flows occur between clusters 6, 17 and 13. The spatiotemporal clustered network for the yellow cluster (Fig. 5(f)) has a different appearance than the workday-related networks, reflecting the leisure and entertainment movements of citizens. Compared to Fig. 5(c) and Fig. 5(d), trips in Midtown Manhattan are significantly reduced, and trips around Central

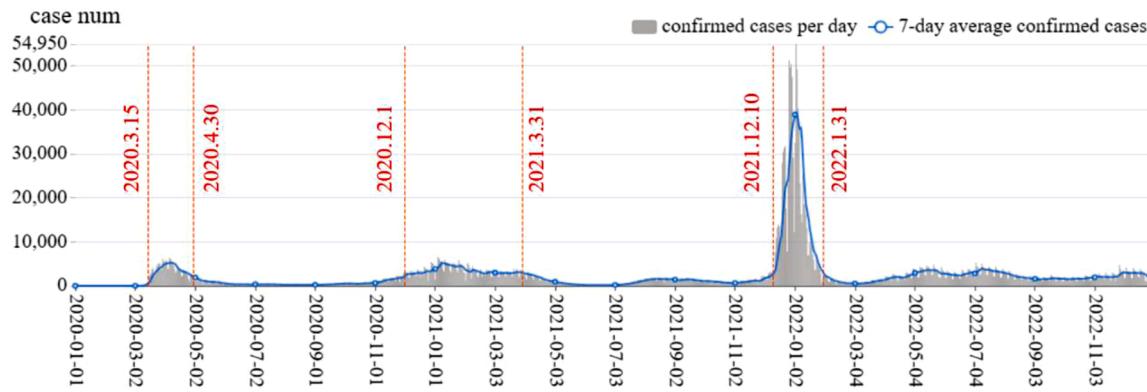


Fig. 3. The daily number of confirmed COVID-19 cases in NYC.

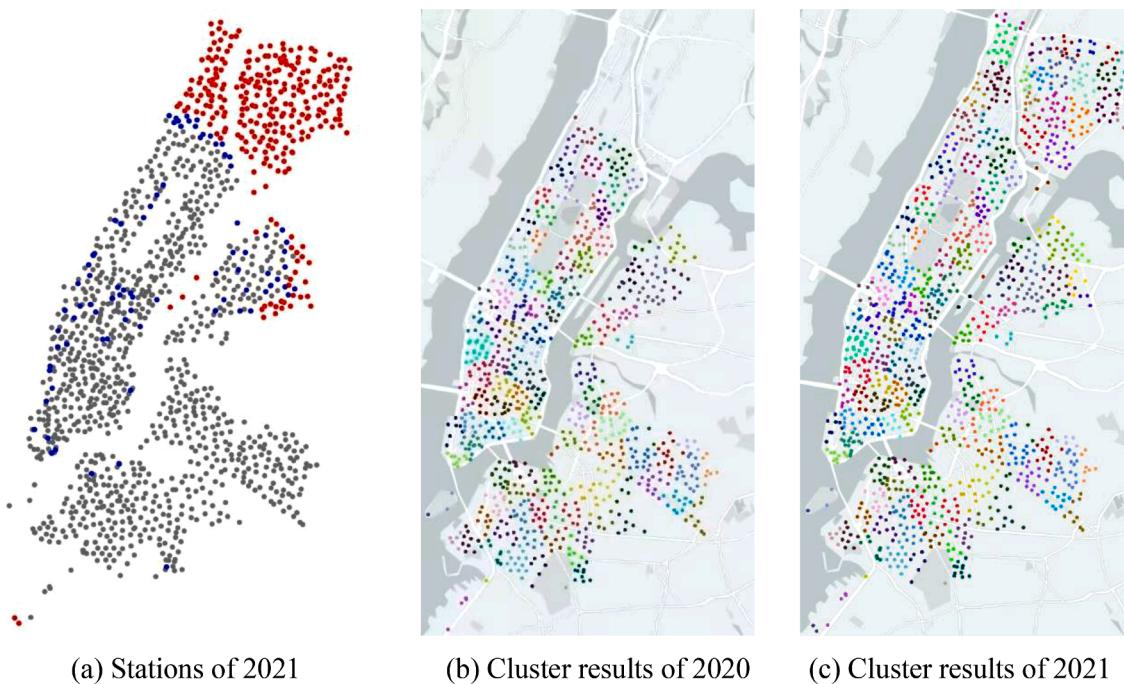


Fig. 4. The incremental station clustering results for 2020 and 2021.

Park are dense. There is a popular trip whose origin and destination are in different directions from Central Park, as shown in Fig. 5(f). There are also more trips between Lower Manhattan and Brooklyn, suggesting that people prefer to travel between the two boroughs on weekends.

The temporal pattern in 2019 is similar to that in 2018, as seen in Fig. 6(a) and Fig. 6(b). The days are divided into four categories: cold workdays (green cluster), warm workdays (blue cluster), cold weekends (red cluster), and warm weekends (yellow cluster). The spatial patterns of the different temporal clusters Fig. 6(c)-Fig. 6(f) are also similar to those in 2018 Fig. 5(c)-Fig. 5(f).

Fig. 7 shows the spatiotemporal clustering results for 2020, the year of the pandemic outbreak. As can be seen in Fig. 7(a), it has a different division of dates compared to the normal years (2018, 2019). The green cluster contains the workdays before the pandemic, which show morning and evening peaks (Fig. 7(b)). Although the first laboratory case was confirmed on February 29, 2020, the cycling pattern in early March was not affected. The red and blue clusters correspond to the pandemic outbreak (lockdown) and stabilization phases, respectively. The red cluster mainly includes days from mid-March to mid-May. Due to the declaration of a national emergency on March 13, 2020 and the state stay-at-home order on March 22, 2020, residents' daily travel is

restricted. The red cluster has the lowest number of hourly rentals. The blue cluster includes days from mid-May through October. With the reopening of low-risk recreational activities on May 15, 2020 and the declaration of the Phase 1 reopening on June 8, 2020 (Timeline of COVID-19 in New York City, 2021), the number of rentals increases in the pandemic stabilization phase, as shown by the blue polyline in Fig. 7 (b). The morning peaks in the red and blue clusters almost disappear due to the work-from-home policy. The yellow cluster contains pre-pandemic weekends and most days in the last two months of 2020. The flow magnitude of the yellow cluster lies between the red and blue clusters.

We further observe the corresponding spatiotemporal clustered networks in Fig. 7(c) to Fig. 7(f). The network of the green cluster shows a polycentric structure, similar to the workday-related networks in 2018 and 2019. For the red cluster, the trips between station clusters are greatly reduced, and the internal flows of station clusters are more significant (Fig. 7(d)). Due to the lockdown policy, most citizens stay at home or only travel short distances. The trips in the blue cluster (Fig. 7(e)) are distributed throughout the city, indicating the recovery of human activity. The spatiotemporal clustered network is dominated by internal flows and long-distance flows. The multicentric structure of the

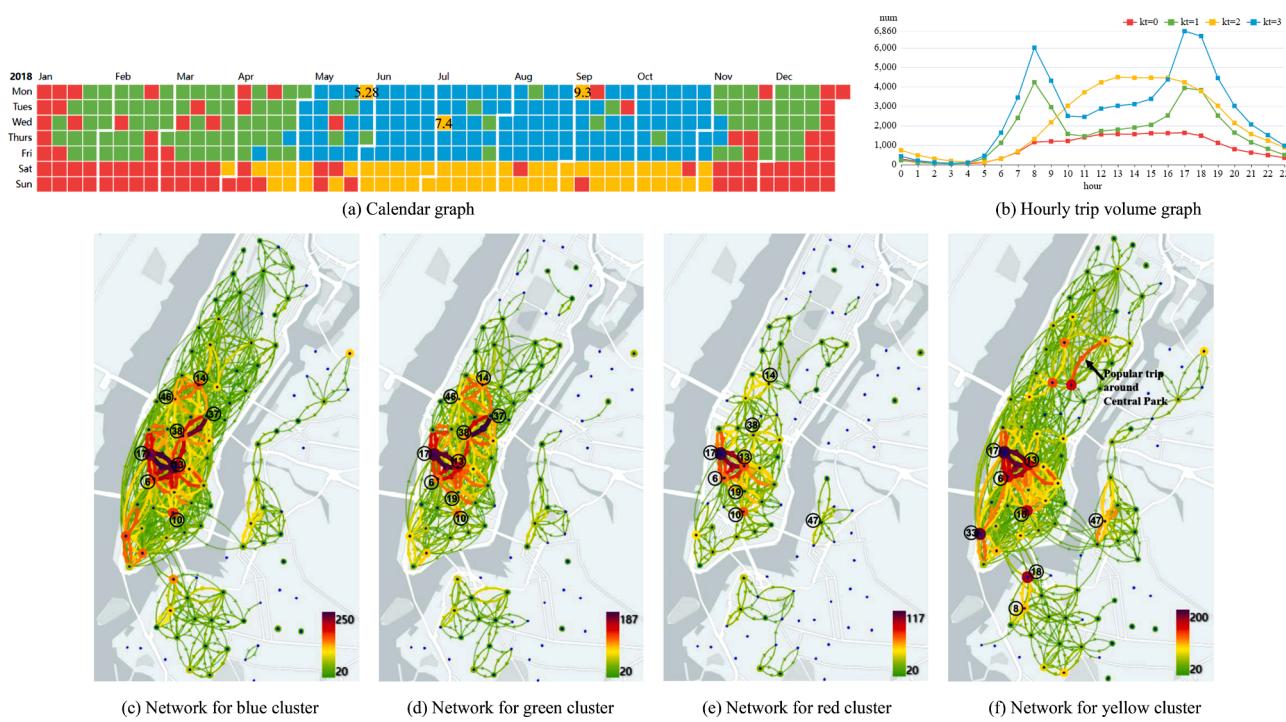


Fig. 5. The spatiotemporal clustering results for 2018 ($KT=4$).

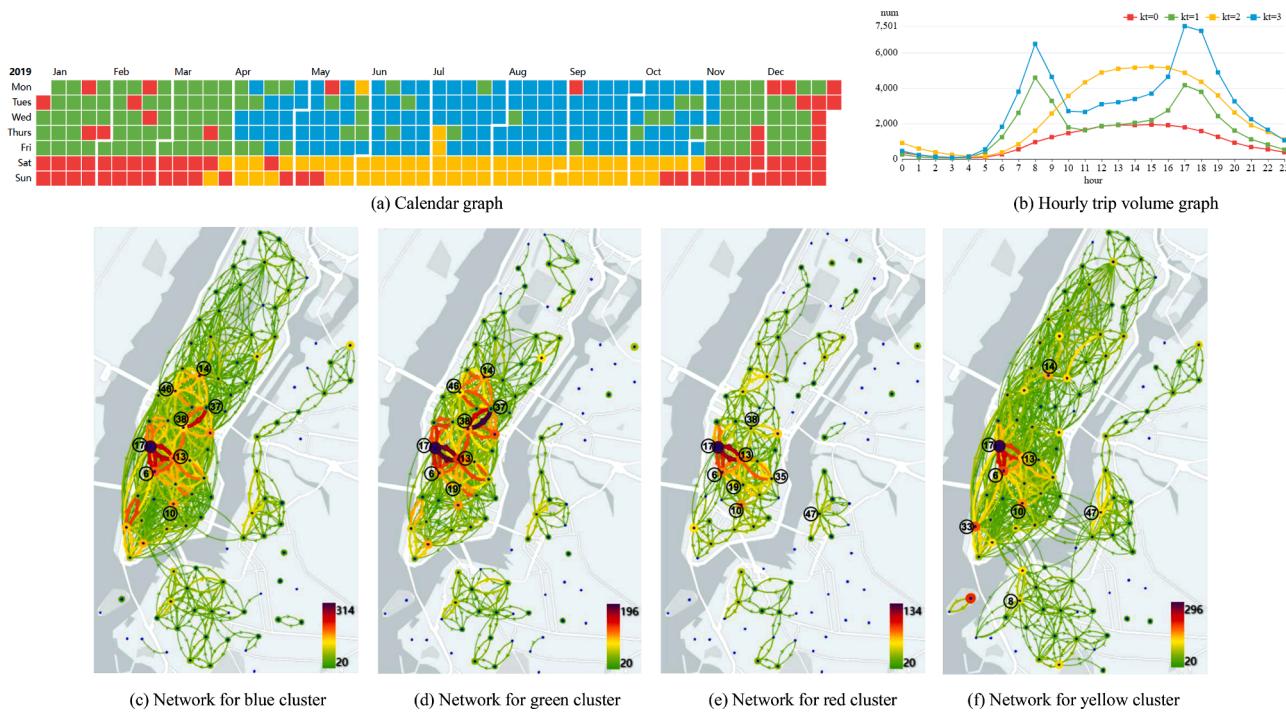


Fig. 6. The spatiotemporal clustering results for 2019 ($KT=4$).

network is not evident. In addition, we find many trips between Lower Manhattan and Brooklyn, indicating the increasing inter-borough bike-sharing traffic during the pandemic stabilization phase. For the yellow cluster (Fig. 7(f)), the flows in Upper Manhattan are reduced compared to the blue cluster. The flow interactions between Lower Manhattan and Brooklyn also decrease.

Fig. 8 and Fig. 9 show the spatiotemporal clustering results for 2021 and 2022, respectively. The workday-weekend pattern disappears after the pandemic outbreak in 2020. However, this pattern reappears in

2021 and 2022 (Fig. 8(a), Fig. 9(a)). The blue and green clusters correspond to workdays with morning and evening peaks (Fig. 8(b), Fig. 9(b)). The yellow cluster corresponds to warm weekends. The number of hourly rentals of the red cluster is the lowest (Fig. 8(b), Fig. 9(b)). Observing the number of confirmed cases in Fig. 3, we find that the red clusters in 2021 and 2022 correspond to the peaks of confirmed cases, and the weather is relatively cold.

Fig. 8(c)-(F) and Fig. 9(c)-(F) show the spatiotemporally clustered networks for 2021 and 2022, respectively. As seen in Fig. 8(c) and Fig. 9

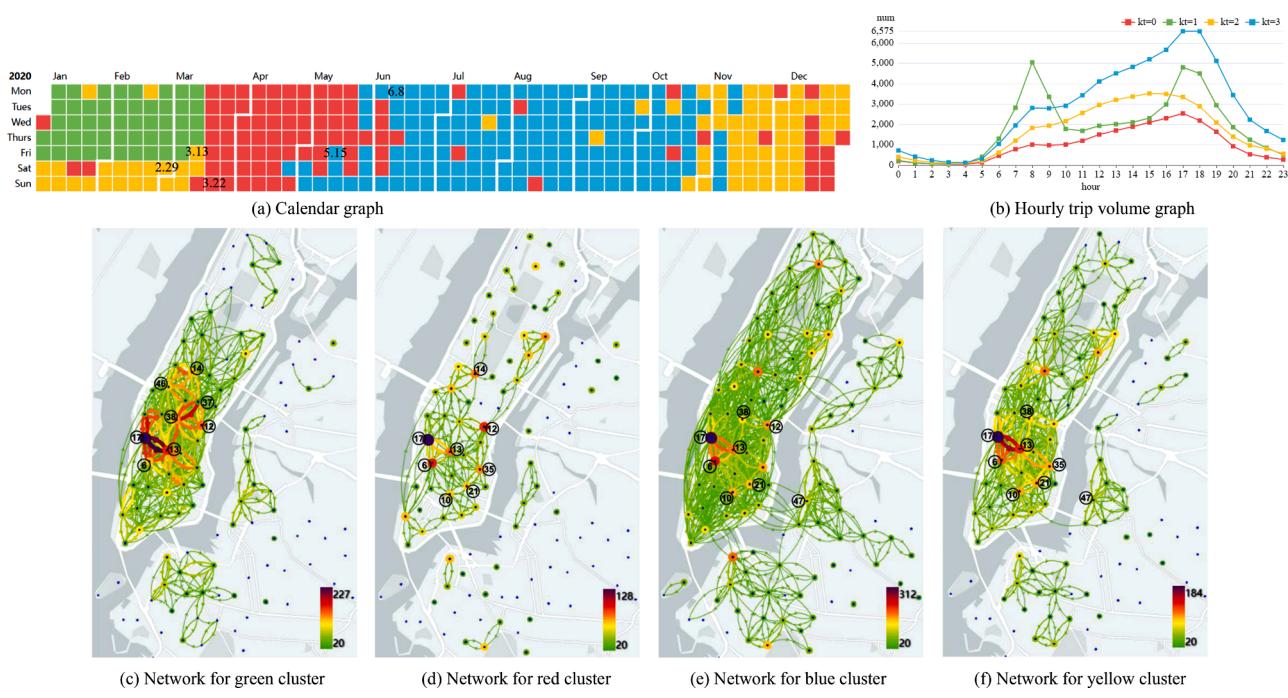


Fig. 7. The spatiotemporal clustering results for 2020 ($KT=4$).

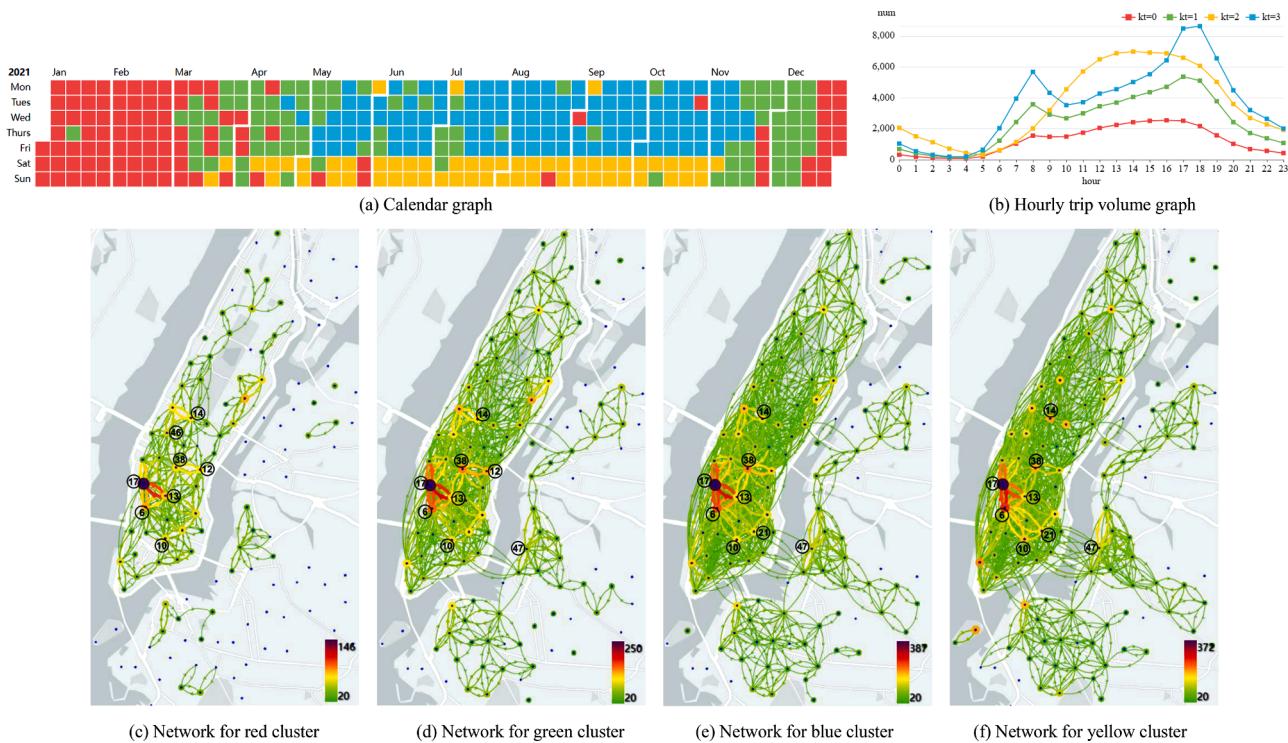


Fig. 8. The spatiotemporal clustering results for 2021 ($KT=4$).

(c), the red clusters contain many internal flows and the citywide mobility is weak. As shown in Fig. 8(d)-(F) and Fig. 9(d)-(F), the green, blue, and yellow clusters have more citywide movements. The interborough movement between Manhattan, Brooklyn and Queens is stronger in the blue and yellow clusters. The weekday-related clusters (green and blue clusters) also show multicentric structures.

5.3. Spatiotemporal clustered network analysis

5.3.1. Network indicators analysis

Table 2 shows the statistical indicators of all spatiotemporal clustered networks. We recognize the similarities and differences between different years.

- (1) The values of number of edges (N_E) are similar in 2018 and 2019. The value of N_E for the green cluster in 2020 is similar to

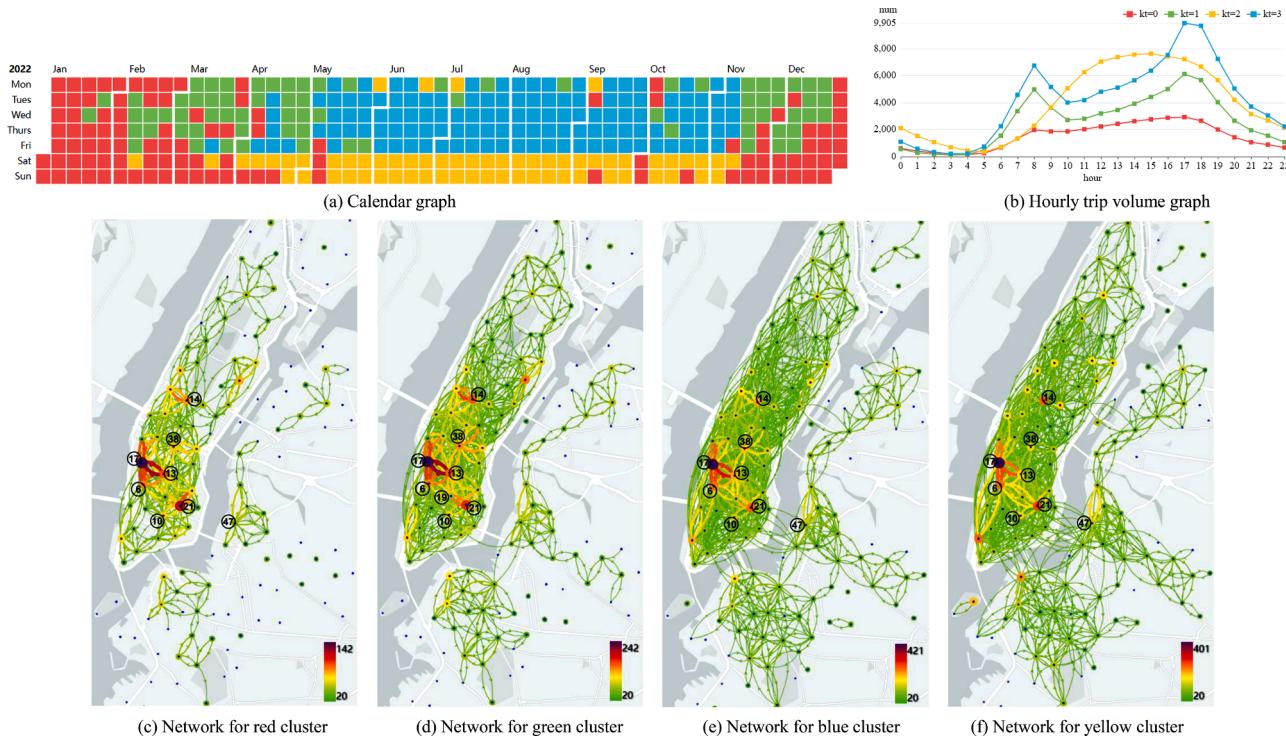


Fig. 9. The spatiotemporal clustering results for 2022 ($KT=4$).

Table 2
Statistical indicators of all spatiotemporal clustered networks.

Year	Cluster	N_N	N_E	Avg.deg	Avg.str	G_{den}	Avg.pLen
2018	Red	100	7759	155.18	399.52	0.784	6.097
	Green	100	8086	161.72	744.31	0.817	6.095
	Yellow	100	8426	168.52	987.65	0.851	6.095
	Blue	100	8806	176.12	1223.50	0.889	6.094
2019	Red	102	7672	150.43	438.47	0.745	6.187
	Green	102	8292	162.59	770.68	0.805	6.170
	Yellow	102	8763	171.82	1132.37	0.851	6.143
	Blue	102	8815	172.84	1300.24	0.856	6.157
2020	Red	111	10,439	188.09	399.92	0.855	6.249
	Green	109	8491	155.80	787.52	0.721	6.257
	Yellow	111	10,468	188.61	682.07	0.857	6.302
	Blue	111	11,618	209.33	1151.14	0.952	6.296
2021	Red	142	13,463	189.62	388.26	0.672	7.836
	Green	142	14,884	209.63	784.76	0.743	7.835
	Yellow	142	15,536	218.82	1136.82	0.776	7.826
	Blue	142	16,011	225.51	1171.57	0.7997	7.827
2022	Red	162	16,337	201.69	421.02	0.626	9.007
	Green	162	16,913	208.80	747.10	0.648	9.002
	Yellow	162	19,082	235.58	1093.38	0.732	9
	Blue	162	19,983	246.70	1170.98	0.766	8.999

the green cluster in 2019 because the pandemic hadn't started in that period. The values of N_E of other clusters increase significantly in 2020, 2021 and 2022. The percentage of edge increase is higher than the percentage of node increase, even for the red cluster of 2020. This indicates that more regions are connected after the pandemic outbreak, and people ride to places they didn't reach before to avoid infection.

(2) The values of average degree (Avg.deg) are similar in 2018 and 2019. Except for the green cluster in 2020, the values of Avg.deg of other clusters are larger in 2020, 2021, and 2022. The values of Avg.deg of the yellow and blue clusters in 2022 are the largest, indicating high connectivity of the regions in the post-pandemic period.

- (3) For the red, yellow and blue clusters in 2020, although the values of Avg.deg increase compared to 2019, the values of average strength (Avg.str) decrease significantly, indicating that the onset of the pandemic has greatly reduced travel frequency. For 2021 and 2022, the values of Avg.str gradually increase to reach the level before the pandemic outbreak. We also find another interesting phenomenon. Although the confirmed cases in early 2022 had reached the peak level since the pandemic outbreak (Fig. 3), the value of Avg.str of the red cluster was higher in 2022 than in 2020 and 2021. This suggests that as the pandemic progressed, people learned to live with the virus and traveled a lot even during the severe pandemic.
- (4) The values of graph density (G_{den}) for the red, yellow and blue clusters in 2020 are relatively high, mainly because the increase

of edges is faster than the increase of nodes. The blue cluster of 2020, which represents the pandemic stabilization phase, even reaches 0.95, indicating the extensive movement of people in the city. However, the value of G_{den} decreases significantly in 2021 and 2022, due to a significant increase in the number of nodes.

- (5) As the number of nodes increased from 2018 to 2022, the value of $Avg.pLen$ gradually increased. Most of the newly added stations are located far from the city center, resulting in a longer average accessible distance between nodes.

5.3.2. Trip feature analysis

Trip features are critical indicators to reflect the characteristics of human travel. As can be seen in Fig. 10(a), the average trip durations of the yellow clusters (warm weekends) are longer than those of the blue clusters (warm workdays), except for 2020. However, their trip distances are similar (Fig. 10(b)). This indicates that the cycling distance of leisure trips and commuting trips are similar. People may ride slower or take a break during leisure trips. With the pandemic outbreak, the blue and red clusters of 2020 have the largest average trip duration and average trip distance (Fig. 10(a), Fig. 10(b)). Compared to 2019, the average duration for the blue and red clusters increased by 5.57 min and 6.57 min, respectively, while the average distance increased by 0.3 km and 0.45 km, respectively. This suggests that people ride public bikes for longer durations and longer distances after the pandemic outbreak. In 2021 and 2022, the average values of durations and distances decrease and are slightly higher than those in 2018 and 2019.

We further compare the differences in trip durations for different spatiotemporal clustered networks using violin plots. We have some conclusions. (1) The green cluster (Fig. 11(a)), which mainly represents cold workdays, has similar curve shapes across different years. The median of 2021 is slightly higher. (2) For the red, blue and yellow clusters (Fig. 11(b)-Fig. 11(d)), the curve shapes for 2018, 2019, 2021 and 2022 are similar, with gradually increasing trip counts. (3) For the year 2020, the red cluster (Fig. 11(b)) and the blue cluster (Fig. 11(c)) correspond to the pandemic outbreak and stabilization phases. The median of these two clusters in 2020 is significantly higher than the medians of the other years. Their kernel density curve is thinner and taller, indicating a smaller number of trips and a larger proportion of long-duration trips. The yellow cluster contains days of cold weekends, November and December during the pandemic. The median of trip duration is relatively low (Fig. 11(d)).

Fig. 12 compares the trip distance distributions for different spatiotemporal clustered networks. Since the range of trip distance is from 0 km to 10 km, the difference in the median of trip distance is less noticeable. The trend of trip duration and trip distance is generally consistent. The distributions and medians for the green clusters are similar over five years (Fig. 12(a)). For the red and blue clusters in 2020 (Fig. 12(b), Fig. 12(c)), the medians of trip distance are large, with a significantly different distribution than in the other years. The thinner and taller curve indicates a decrease in the number of trips and an increase in long-distance trips after the pandemic outbreak. For the yellow

cluster (Fig. 12(d)), unlike the medians for trip duration, the medians for trip distance are similar over five years, indicating that the distribution of trip distance varies little from year to year.

5.4. Measuring the properties of important station clusters

Table 3 shows the top 10 important station clusters for different spatiotemporal clustered networks. The locations and importance of these station clusters are visualized in Fig. 13. To help better understand the flow relationship between different station clusters, we also mark the top 8 station clusters in Fig. 5(c)-(F) to Fig. 9(c)-(F). Fig. 14 shows the category distributions of the POIs for all important station clusters. We divide the identified important station clusters into four types based on their characteristics.

The first type of station clusters appears in the top 10 list in different spatiotemporal clustered networks of different years. Station clusters 13, 17 and 6 are critical, as shown in Table 3 and Fig. 13. They are highly connected, and form a triangle representing one of the most popular regions in Manhattan (Fig. 5(c)-(F) to Fig. 9(c)-(F)). As seen from the POI distribution graph in Fig. 15(a), station cluster 13 is located around Union Square, which includes Union Square Park, subway stations, and many shops and office buildings. The surroundings of station clusters 17 and 6 (Fig. 14) include nightlife spots, office buildings, shopping malls, and recreational facilities. The associated movements can be caused by various travel purposes. In addition, station clusters 14, 19 and 38 with mixed land-use functions are also important in most spatiotemporal clustered networks. The largest POI proportion of cluster 38 is "Travel & Transport", which includes Penn Station (Fig. 15(b)).

The second type of station clusters is work-related. Station clusters 37 and 46 show great attractiveness in the spatiotemporal clustered networks related to work, i.e., the green and blue clusters in 2018 and 2019, and the green cluster in 2020 (Table 3). Due to the work-from-home policy, the two station clusters are missing from the top 10 list of the red, blue and yellow clusters in 2020. The largest POI proportion is "Travel & Transport", which contains transportation hubs such as Grand Central Terminal and Port Authority Bus Terminal (Fig. 16(a), Fig. 16(b)).

The third type of station clusters is related to leisure and entertainment. Station clusters 47 and 10 show high importance in the red and yellow clusters in 2018 and 2019. Except for the green cluster in 2020, station clusters 47, 10 and 21 are attractive in the clusters from 2020 to 2022 (Table 3, Fig. 13). Station cluster 47 is located in the northern part of Brooklyn (Fig. 5(f)), which is highly connected to the adjacent station clusters and can be considered an essential zone for the surrounding areas. Station clusters 10 and 21 are located in Lower Manhattan (Fig. 9(f)). The top 2 POI types of the three station clusters are "Nightlife Spots" and "Shop & Service" (Fig. 14). We speculate that they provide leisure and daily life services to residents during both normal and pandemic periods.

The fourth type of station clusters is important during the pandemic outbreak. Station clusters 12 and 35 rank 4th and 6th in the red cluster

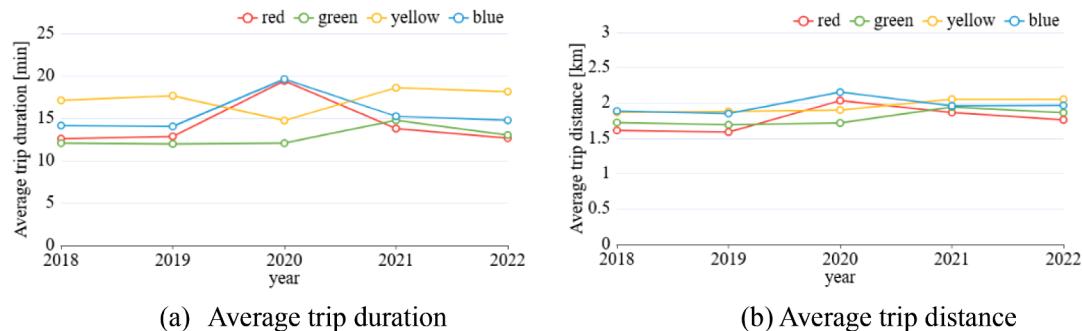


Fig. 10. Average trip duration and average trip distance for different spatiotemporal clustered networks.

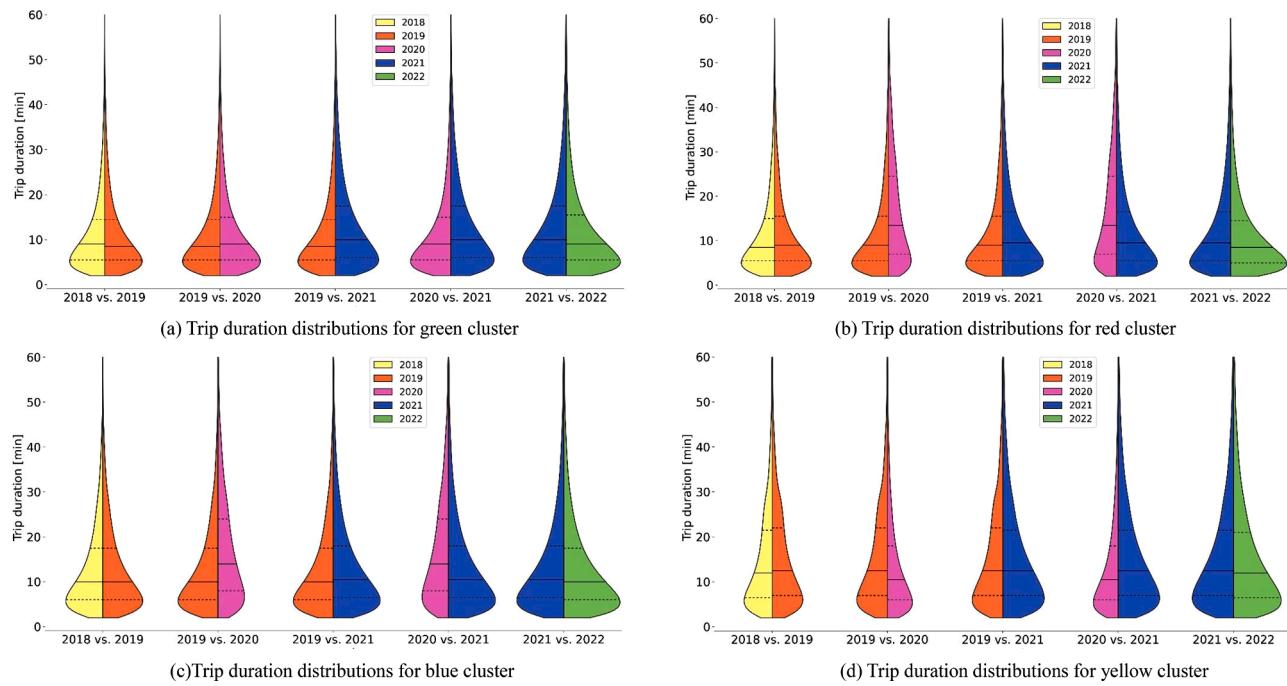


Fig. 11. The comparison of trip duration distributions for different spatiotemporal clustered networks.

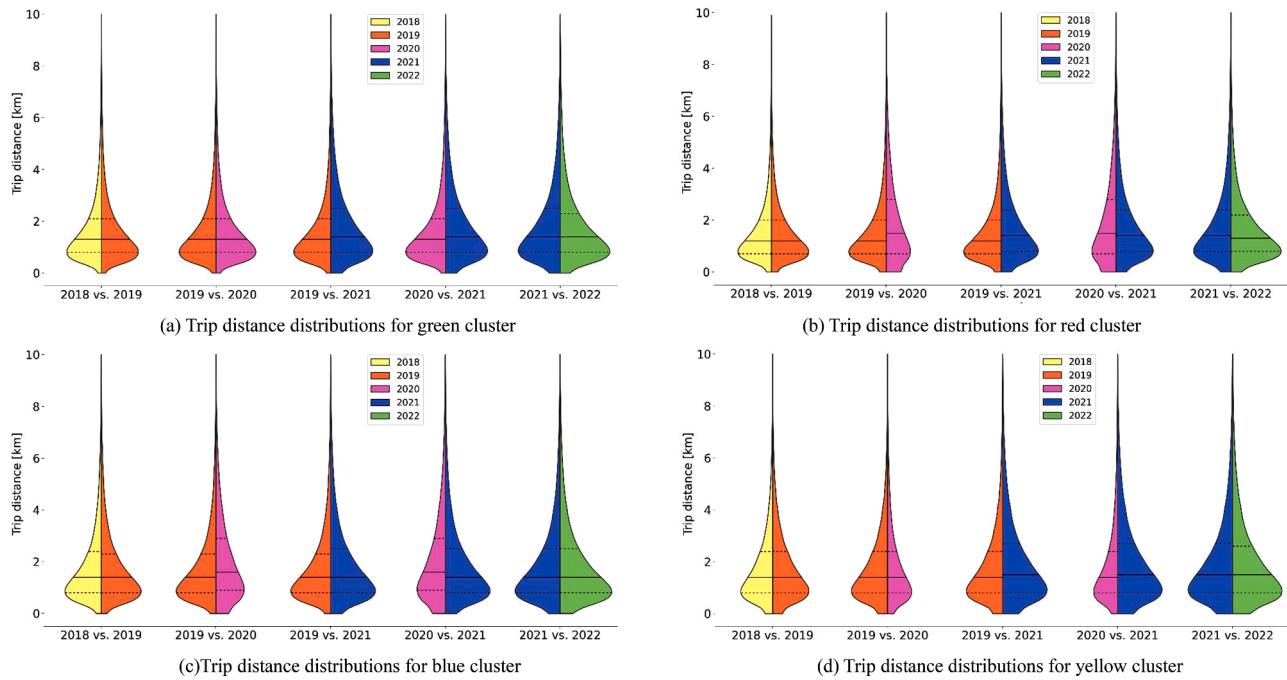


Fig. 12. The comparison of trip distance distribution for different spatiotemporal clustered networks.

of 2020 (Table 3). They are also critical in the blue and yellow clusters in 2020. These station clusters contain medical centers, such as NYU Langone Medical Center, Tisch Hospital, and Mount Sinai Beth Israel, as shown in Fig. 17(a) and Fig. 17(b).

6. Discussion

6.1. Result analysis and policy recommendation

This study examines spatiotemporal changes in cycling patterns in NYC over a five-year period. Key findings are discussed below.

First, according to the spatiotemporal clustering results, the five years can be divided into three periods: the pre-pandemic period (2018, 2019), the pandemic outbreak period (2020), and the post-pandemic period (2021, 2022). The pre-pandemic period had a workday-weekend pattern. The spatial distribution of trips was different for commuting and recreation. For the pandemic outbreak period, the workday-weekend pattern disappeared, and the dates were divided according to the evolution of the pandemic. During the lockdown phase (red cluster in 2020), the trip volume was low, and the self-loop ratio was the highest of all spatiotemporal clustered networks. People reduced unnecessary travel and made daily purchases close to home. In

Table 3

The top 10 important station clusters for all spatiotemporal clustered networks.

Year	Cluster	1	2	3	4	5	6	7	8	9	10
2018	red	13	17	10	6	19	47	38	14	22	8
	green	13	38	17	37	14	46	6	19	22	10
	blue	13	38	17	6	14	10	46	37	19	22
	yellow	13	17	10	6	47	33	8	18	14	19
2019	red	13	17	6	10	47	19	35	38	8	21
	green	13	38	17	37	14	46	6	19	43	12
	blue	13	17	6	38	37	10	46	14	19	33
	yellow	17	6	10	47	13	33	8	14	19	35
2020	red	17	13	6	12	10	35	14	21	38	47
	green	13	17	38	37	6	14	12	46	43	19
	blue	17	6	13	10	47	21	12	38	14	35
	yellow	13	17	6	10	21	47	38	35	14	19
2021	red	17	13	38	6	10	14	46	12	21	62
	green	17	38	6	13	10	14	47	12	21	162
	blue	17	6	38	13	10	47	14	21	46	12
	yellow	17	6	10	47	13	38	21	14	39	162
2022	red	17	13	21	38	6	10	14	47	46	8
	green	17	13	38	14	21	6	10	19	46	43
	blue	17	13	6	38	21	14	47	10	19	33
	yellow	17	6	21	47	10	13	14	38	19	33

addition, the red and blue clusters of 2020 have more long-duration and long-distance trips. This suggests that people tended to substitute cycling for other travel modes to reduce the risk of infection, similar to cities such as Chicago (Hu et al., 2021), Zurich (Li et al., 2021) and Beijing (Shang et al., 2021). The cycling pattern in the post-pandemic period was similar to the pre-pandemic period. We can find workday-weekend patterns except for the periods when the pandemic was sever. The trip duration/distance distributions were similar to those in the pre-pandemic period, with increased trip volume. In addition, we find that the confirmed cases in early 2022 were the highest since the pandemic outbreak. However, the corresponding travel intensity was higher than the red cluster in 2020 and 2021. This confirms that after a one-year adjustment period, people have learned to live with the virus and the daily cycling returned to the pre-pandemic status. Previous studies revealed that the use of shared bikes in 2020 followed an “increase-decrease-rebound” pattern (Hu et al., 2021). Looking at a five-year period, the macroscopic pattern can also be concluded as an “increase-decrease-rebound” pattern. Cycling increased before the pandemic, decreased in 2020, and rebounded in 2021 and 2022. The results of this study increase public knowledge about the system. It can also help operators and urban transportation planners to better understand the citywide flow structures of the bike-sharing system, which can help them to estimate user demand in different regions at different times.

Second, we find that the pandemic-related networks have larger values of edge number and average degree. Even for the lockdown phase (red cluster) in 2020, the growth rate of edge number is much higher than that of node number. Xin et al. (2022) found that the edge number decreased from March 2020 to April 2020, and assumed that the riding connections in the bicycle network were reduced with the development of the pandemic. By analyzing a longer period with accurate classification dates, we believe that the area of riding connections became more extensive during the lockdown phase. People may ride bikes to places they didn't reach before or travel by other traffic modes. A similar conclusion was drawn for the docked bike networks during the lockdown period in Zurich, Switzerland (Li et al., 2021). The results confirm that bike-sharing services have the potential to facilitate a disease-resilient transport system. City authorities should formulate subsidy policies for bike-sharing companies to support the stable operation of the system, as societies may have to coexist with COVID in the future.

Third, bike-sharing companies should pay attention to the identified key station clusters. Several station clusters are important in all years, pandemic or not, such as station clusters 13, 17, 6, 14, 19 and 38. The

areas around these station clusters have mixed land-use functions. Station clusters 47, 10 and 21 are among the top 10 important station clusters even in the pandemic outbreak phase. These station clusters have a high proportion of “Nightlife Spots” and “Shop & Service” POIs, and provide leisure and daily life services to residents. Maintaining the availability of bicycles in relevant areas is important to support people's daily mobility. During the pandemic, some station clusters that contain medical centers were heavily used, such as station clusters 12 and 35. Disinfecting and supplying bikes to stations in these clusters is particularly important. In addition, in the event of another pandemic in the future, bike-sharing operators can set up more stations in the medical-related station clusters and build bike lanes to make it easier for citizens to get to the hospital. Locations near major transportation hubs, such as station clusters 37 and 46, which include Grand Central Terminal and the Port Authority Bus Terminal, are attractive in workday-related clusters, but are less important in the pandemic year due to the work-from-home policies. Bike scheduling during peak hours in workdays is particularly important for these station clusters.

Fourth, we find many inter-borough trips between Manhattan and Brooklyn in 2021 and 2022. The New York Department of Transportation converted a highway into a bike lane for the Brooklyn Bridge in 2021. The results confirm that well-constructed bike lanes on the bridges, including the Manhattan Bridge, Brooklyn Bridge, and Williamsburg Bridge, can encourage travelers to ride and reduce traffic congestion.

6.2. Limitations of the study

In this study, we only analyze the cycling pattern, which cannot summarize human mobility. Multiple travel modes should be considered together to fully understand urban mobility. In addition, due to different local government policies, different cities have their own characteristics in responding to the pandemic. If data are available, comparing cycling patterns in other cities may help to understand the global differences in the pre- and post-pandemic periods. In addition, we roughly infer the visit purposes of station clusters from the proportion and distribution of POIs. Predicting visit purposes at the trajectory level is more accurate, but requires information about the user profile.

7. Conclusion

Bike sharing is an important option during the COVID-19 pandemic due to its low risk of contagion. The objective of this paper is to reveal the long-term changes in cycling patterns affected by the pandemic at

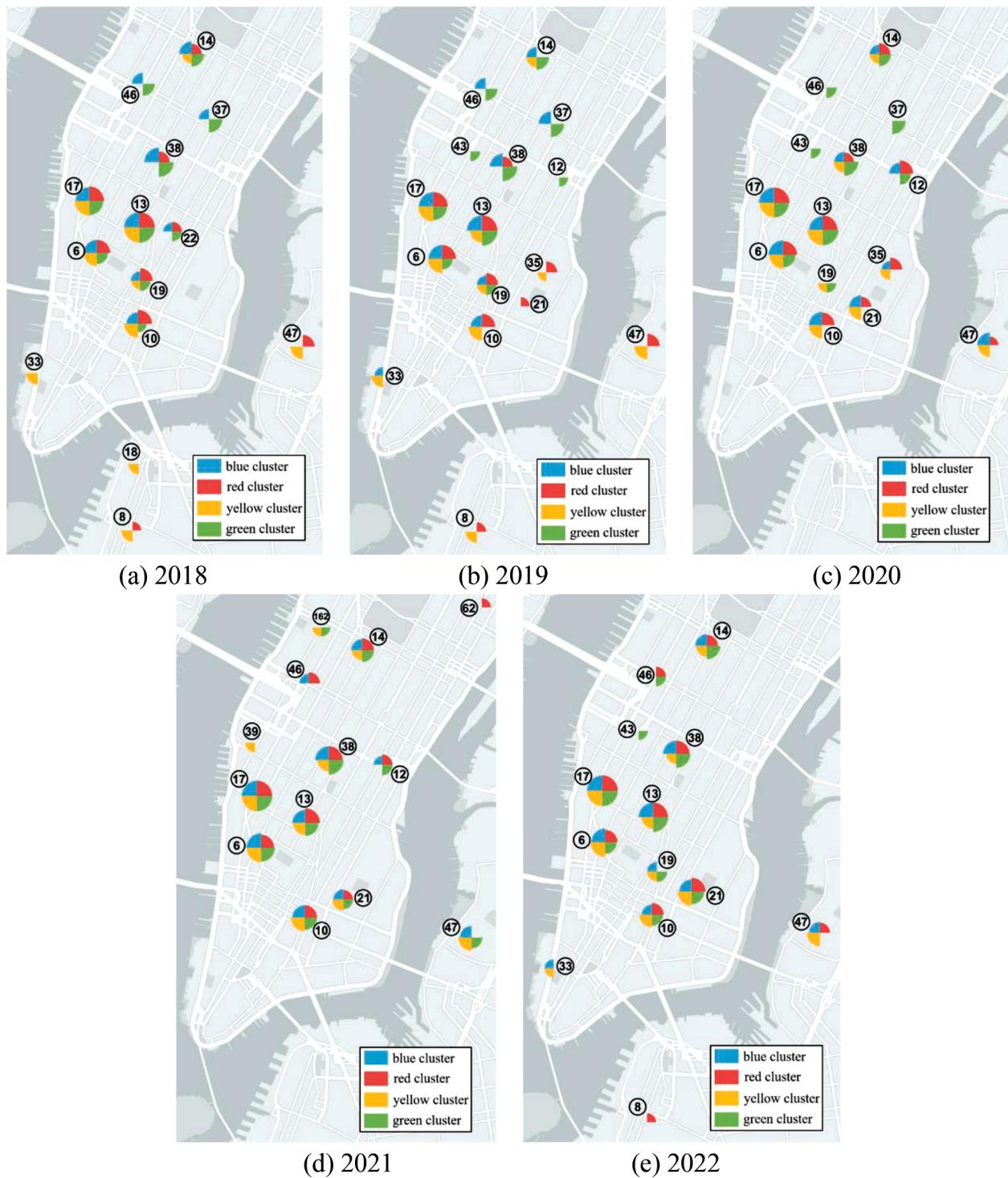


Fig. 13. The important station cluster graphs for different years.

the city level. Using the bike-sharing dataset in NYC, the evolutionary patterns, network statistical characteristics, and the role of different spatial locations in the pre-, during, and post-pandemic periods are clearly demonstrated. Based on our analysis, we find that the bike-sharing system plays an important role in daily travel in both normal and pandemic years, and draw the following conclusions:

- (1) Based on the explored spatiotemporal patterns, the study period can be divided into three sub-periods: the pre-pandemic period, the pandemic outbreak period, and the post-pandemic period.

The bike usage pattern follows an “increase-decrease-rebound” pattern.

- (2) The pre-pandemic period has a workday-weekend pattern. The bicycle networks show multicentric structures on workdays, and reflect the leisure movements of citizens on weekends. The warm workdays and warm weekends have similar average trip distances, but the latter has longer average trip durations, suggesting that cycling distances are similar for different trip purposes, but people may ride slower during leisure trips.

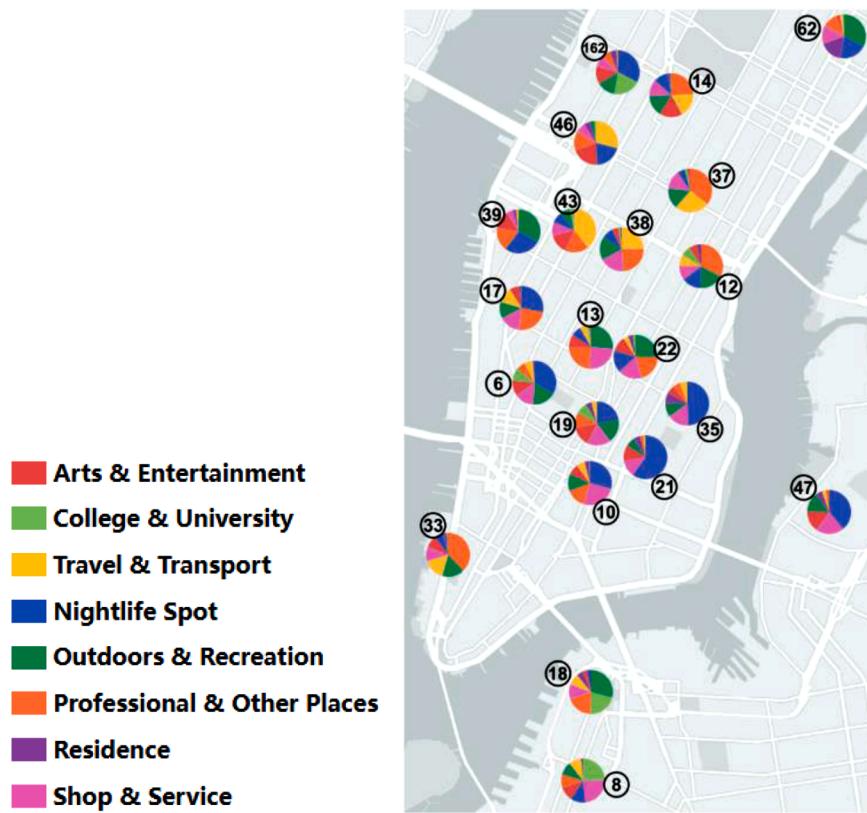


Fig. 14. The POI pie graph.

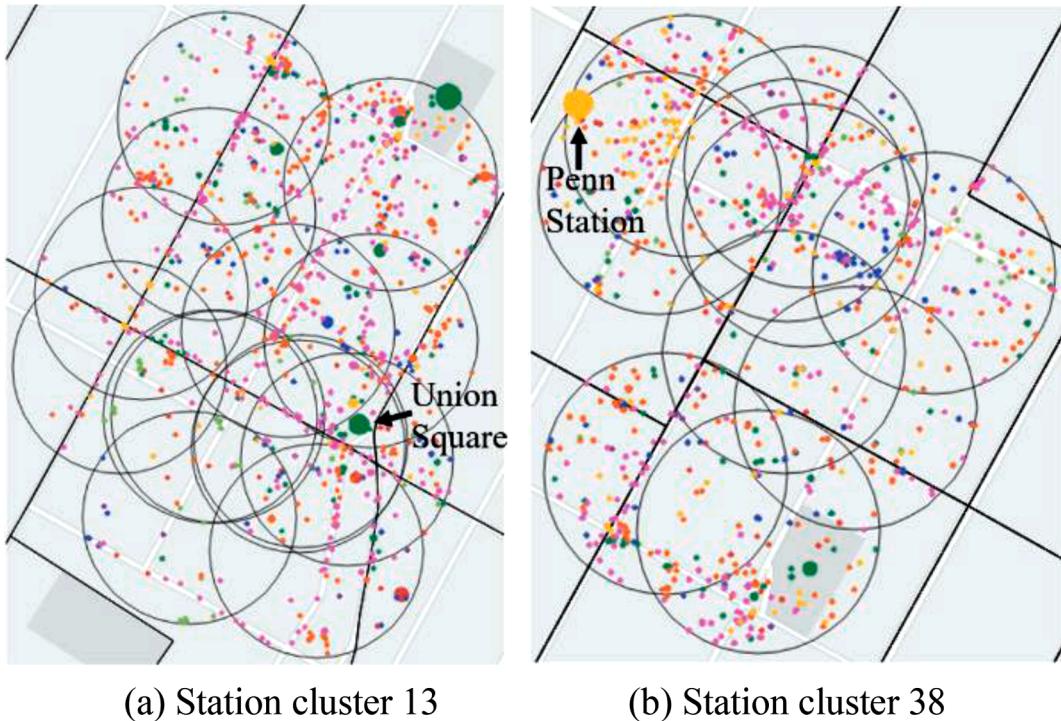


Fig. 15. The POI distribution graph of important station clusters.

(3) The spatiotemporal pattern in 2020 is different. The workday-weekend pattern disappears after the pandemic outbreak. The dates can be divided into the pre-pandemic phase, the pandemic outbreak phase, and the pandemic stabilization phase. During the

latter two phases, trips are long in duration and distance, but low in volume. The multicentric structures of the network are not obvious and the traffic near transportation hubs is reduced. In the pandemic outbreak phase, the citywide mobility is weak and the

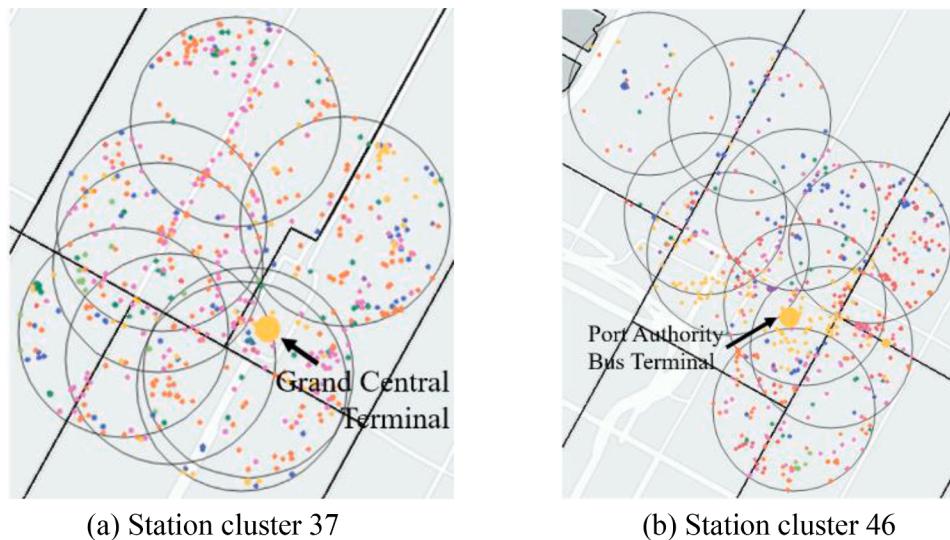


Fig. 16. The POI distribution graph of work-related important station clusters.

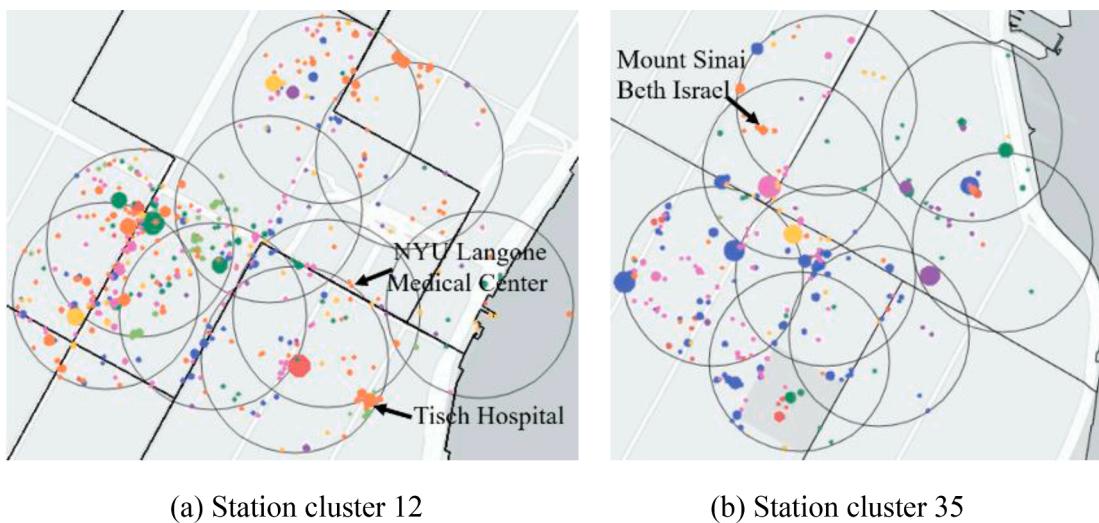


Fig. 17. The POI distribution graph of pandemic-related important station clusters.

network contains significant internal flows. However, the network connectivity is better, indicating that people ride to places they didn't reach before. The network corresponding to the pandemic stabilization phase has the highest graph density, indicating the extensive movement of people in the city during this phase. Bicycle mobility in 2020 suggests that bike-sharing is a disease-resilient travel mode.

- (4) In the post-pandemic period, the cycling patterns are similar to the pre-pandemic period. The workday-weekend pattern reappears, and the days can be divided into four types: warm workdays, cold workdays, warm weekends, and days when confirmed cases increase. The coverage range of cycling and the number of trips continue to increase, and the average values of duration and distance decrease compared to 2020. People traveled a lot even when confirmed cases were at their peak. The results confirm that as the pandemic progresses, people have learned to live with the virus. The frequency and range of cycling is even higher than before the pandemic. As the demand for bikes continues to grow, bicycle companies can consider building more stations, and the number of rentals will increase as a result.

(5) We have identified several types of key station clusters. In the post-pandemic period, bike-sharing operators should disinfect the slots and bikes in these stations in a timely manner to ensure the safety of users. They should also use the appropriate bike scheduling strategies according to the usage characteristics of different types of station clusters.

Our study complements and extends existing research. The results provide a new perspective for understanding the pandemic impact on shared bike riding and offer valuable insights for bicycle operators and policymakers. Our approach can be applied to other cities if the bike-sharing dataset is available. The approach can also be used to analyze other travel modes if the corresponding origin-destination datasets are available. In the future, we plan to collect different types of human mobility data to comprehensively analyze the impact of the pandemic on the urban transportation system.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

the work reported in this paper.

Data availability

The data are public. I have written data links in the manuscript.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 61903109), and the Zhejiang Provincial Natural Science Foundation of China (No. LTGS23F030004).

References

- Anowar, F., Sadaoui, S., & Selim, B. (2021). Conceptual and empirical comparison of dimensionality reduction algorithms (pca, kpca, lda, mds, svd, lle, isomap, le, ica, t-sne). *Computer Science Review*, 40, Article 100378.
- Arthur, D., & Vassilvitskii, S. (2007). K-means++: The advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on discrete algorithms* (pp. 1027–1035).
- Bao, J., Xu, C., Liu, P., et al. (2017). Exploring bikesharing travel patterns and trip purposes using smart card data and online point of interests. *Networks and Spatial Economics*, 17(4), 1231–1253.
- Bergantino, A. S., Intini, M., & Tangari, L. (2021). Influencing factors for potential bike-sharing users: An empirical analysis during the COVID-19 pandemic. *Research in Transportation Economics*, 86, Article 101028.
- Bi, H., Ye, Z., Zhang, Y., et al. (2022). A long-term perspective on the COVID-19: The bike sharing system resilience under the epidemic environment. *Journal of Transport & Health*, 26, Article 101460.
- Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN Systems*, 30(1–7), 107–117.
- Bustamante, X., Federo, R., & Fernández-i-Marín, X. (2022). Riding the wave: Predicting the use of the bike-sharing system in Barcelona before and during COVID-19. *Sustainable Cities and Society*, 83, Article 103929.
- Caulfield, B., O'Mahony, M., Brazil, W., et al. (2017). Examining usage patterns of a bike-sharing scheme in a medium sized city. *Transportation Research Part A: Policy and Practice*, 100, 152–161.
- Chai, X., Guo, X., Xiao, J., et al. (2021). Analysis of spatiotemporal mobility of shared-bike usage during COVID-19 pandemic in Beijing. *Transactions in GIS*, 25(6), 2866–2887.
- Chen, Y., Sun, X., Deveci, M., et al. (2022a). The impact of the COVID-19 pandemic on the behaviour of bike sharing users. *Sustainable Cities and Society*, 84, Article 104003.
- Chen, X., Huang, K., & Jiang, H. (2022b). Detecting changes in the spatiotemporal pattern of bike sharing: A change-point topic model. *IEEE Transactions on Intelligent Transportation Systems*, 23(10), 18361–18377.
- Chen, W., Chen, X., Cheng, L., et al. (2022c). Delineating borders of urban activity zones with free-floating bike sharing spatial interaction network. *Journal of Transport Geography*, 104, Article 103442.
- Chibwe, J., Heydari, S., Imani, A. F., et al. (2021). An exploratory analysis of the trend in the demand for the London bike-sharing system: From London Olympics to Covid-19 pandemic. *Sustainable Cities and Society*, 69, Article 102871.
- Corcoran, J., Li, T., Rohde, D., et al. (2014a). Spatiotemporal patterns of a Public Bicycle Sharing Program: The effect of weather and calendar events. *Journal of Transport Geography*, 41, 292–305.
- Corcoran, J., Li, T., Rohde, D., et al. (2014b). Spatio-temporal patterns of a public bicycle sharing program: The effect of weather and calendar events. *Journal of Transport Geography*, 41, 292–305.
- DeMaio, P. (2009). Bike-sharing: History, impacts, models of provision, and future. *Journal of Public Transportation*, 12(4), 3.
- Faghih-Imani, A., & Eluru, N. (2016). Incorporating the impact of spatiotemporal interactions on bicycle sharing system demand: A case study of New York Citi bike system. *Journal of Transport Geography*, 54, 218–227.
- Fishman, E. (2016). Bikeshare: A review of recent literature. *Transport Reviews*, 36(1), 92–113.
- Galvani, M., Torti, A., Menafoglio, A., et al. (2020). A novel spatio-temporal clustering technique to study the bike sharing system in Lyon. *EDBT/ICDT Workshops*. Copenhagen, Denmark.
- Gkiotsalitis, K., & Cats, O. (2021). Public transport planning adaption under the COVID-19 pandemic crisis: Literature review of research needs and directions. *Transport Reviews*, 41(3), 374–392.
- Hu, T., Wang, S., She, B., et al. (2021a). Human mobility data in the COVID-19 pandemic: Characteristics, applications, and challenges. *International Journal of Digital Earth*, 14(9), 1126–1147.
- Hu, S., Xiong, C., Liu, Z., et al. (2021b). Examining spatiotemporal changing patterns of bike-sharing usage during COVID-19 pandemic. *Journal of Transport Geography*, 91, Article 102997.
- Jiao, J., Lee, H. K., & Choi, S. J. (2022). Impacts of COVID-19 on bike-sharing usages in Seoul, South Korea. *Cities*, 130, Article 103849. London, England.
- Jobe, J., & Griffin, G. P. (2021). Bike share responses to COVID-19. *Transportation Research Interdisciplinary Perspectives*, 10, Article 100353.
- Kim, M., & Cho, G. H. (2022). Examining the causal relationship between bike-share and public transit in response to the COVID-19 pandemic. *Cities*, 131, Article 104024. London, England.
- Kim, K. (2021). Impact of CoViD-19 on usage patterns of a bike-sharing system: Case study of Seoul. *Journal of Transportation Engineering, Part A: Systems*, 147(10), Article 05021006.
- Kubařák, S., Kalašová, A., & Hájnik, A. (2021). The bike-sharing system in Slovakia and the impact of COVID-19 on this shared mobility service in a selected city. *Sustainability*, 13(12), 6544.
- Legros, B. (2019). Dynamic repositioning strategy in a bike-sharing system; how to prioritize and how to rebalance a bike station. *European Journal of Operational Research*, 272(2), 740–753.
- Li, Y., Zheng, Y., Zhang, H., et al. (2015). Traffic prediction in a bike-sharing system[C]. In *Proceedings of the 23rd SIGSPATIAL international conference on advances in geographic information systems* (pp. 1–10).
- Li, A., Zhao, P., Haitao, H., et al. (2021a). How did micro-mobility change in response to COVID-19 pandemic? A case study based on spatial-temporal-semantic analytics. *Computers, Environment and Urban Systems*, 90, Article 101703.
- Li, S., Zhuang, C., Tan, Z., et al. (2021b). Inferring the trip purposes and uncovering spatio-temporal activity patterns from dockless shared bike dataset in Shenzhen, China. *Journal of Transport Geography*, 91, Article 102974.
- Li, H., Zhang, Y., Zhu, M., et al. (2021c). Impacts of COVID-19 on the usage of public bicycle share in London. *Transportation Research Part A: Policy and Practice*, 150, 140–155.
- Mehdizadeh Dastjerdi, A., & Morency, C. (2022). Bike-sharing demand prediction at community level under COVID-19 using deep learning. *Sensors*, 22(3), 1060.
- Nikiforidis, A., Afantopoulou, G., & Stamelou, A. (2020). Assessing the impact of COVID-19 on bike-sharing usage: The case of Thessaloniki, Greece. *Sustainability*, 12(19), 8215.
- Padmanabhan, V., Pennetsa, P., Li, X., et al. (2021). COVID-19 effects on shared-biking in New York, Boston, and Chicago. *Transportation Research Interdisciplinary Perspectives*, 9, Article 100282.
- Pase, F., Chiariotti, F., Zanella, A., et al. (2020). Bike sharing and urban mobility in a post-pandemic world. *IEEE access : Practical Innovations, Open Solutions*, 8, 187291–187306.
- Qin, Y., & Karimi, H. A. (2023). Evolvement patterns of usage in a medium-sized bike-sharing system during the COVID-19 pandemic. *Sustainable Cities and Society*, Article 104669.
- Seng, D., Lv, F., Liang, Z., et al. (2021). Forecasting traffic flows in irregular regions with multi-graph convolutional network and gated recurrent unit. *Frontiers of Information Technology & Electronic Engineering*, 22(9), 1179–1193.
- Shang, W. L., Chen, J., Bi, H., et al. (2021). Impacts of COVID-19 pandemic on user behaviors and environmental benefits of bike sharing: A big-data analysis. *Applied Energy*, 285, Article 116429.
- Shi, X., Yu, Z., Chen, J., et al. (2018). The visual analysis of flow pattern for public bicycle system. *Journal of Visual Languages & Computing*, 45, 51–60.
- Shi, X., Lv, F., Seng, D., et al. (2019). Visual exploration of mobility dynamics based on multi-source mobility datasets and POI information. *Journal of Visualization*, 22(6), 1209–1223.
- Shi, X., Liang, Z., & Seng, D. (2023). Electric fences for dockless bike-sharing systems: an electric fence-planning framework for a dockless bike-sharing system based on a land Parcel subdivision and regional coverage maximization. *IEEE Intelligent Transportation Systems Magazine*, 15(2), 58–69.
- Song, J., Zhang, L., Qin, Z., et al. (2022). Spatiotemporal evolving patterns of bike-share mobility networks and their associations with land-use conditions before and after the COVID-19 outbreak. *Physica A: Statistical Mechanics and its Applications*, 592, Article 126819.
- Sung, H. (2023). Causal impacts of the COVID-19 pandemic on daily ridership of public bicycle sharing in Seoul. *Sustainable Cities and Society*, 89, Article 104344.
- Teixeira, J. F., & Lopes, M. (2020). The link between bike sharing and subway use during the COVID-19 pandemic: The case-study of New York's Citi Bike. *Transportation Research Interdisciplinary Perspectives*, 6, Article 100166.
- Teixeira, J. F., Silva, C., & Sá, F. M. (2021). The motivations for using bike sharing during the COVID-19 pandemic: Insights from Lisbon. *Transportation Research. Part F, Traffic Psychology and Behaviour*, 82, 378.
- Timeline of COVID-19 in New York City, 2021, A Historical Timeline of COVID-19 in New York City, Available at: <https://www.investopedia.com/historical-timeline-of-covid-19-in-new-york-city-5071986> (Accessed 2 April 2022).
- Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(11), 2579–2605.
- Wang, H., & Noland, R. B. (2021). Bikeshare and subway ridership changes during the COVID-19 pandemic in New York City. *Transport Policy*, 106, 262–270.
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 440–442.
- Wu, C., & Kim, I. (2020). Analyzing the structural properties of bike-sharing networks: Evidence from the United States, Canada, and China. *Transportation Research Part A: Policy and Practice*, 140, 52–71.
- Wu, J., Wang, L., & Li, W. (2018). Usage patterns and impact factors of public bicycle systems: Comparison between city center and suburban district in Shenzhen. *Journal of Urban Planning and Development*, 144(3), Article 04018027.
- Xin, R., Ai, T., Ding, L., et al. (2022). Impact of the COVID-19 pandemic on urban human mobility-A multiscale geospatial network analysis using New York bike-sharing data. *Cities*, 126, Article 103677. London, England.
- Xu, J., Li, A., Li, D., et al. (2017). Difference of urban development in China from the perspective of passenger transport around Spring Festival. *Applied Geography*, 87, 85–96.

- Yang, D., Zhang, D., Zheng, V. W., et al. (2014). Modeling user activity preference by leveraging user spatial temporal characteristics in LBSNs. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 45(1), 129–142.
- Yang, H., Guo, Z., Zhai, G., et al. (2022). Exploring the Spatially Heterogeneous Effects of the Built Environment on Bike-Sharing Usage during the COVID-19 Pandemic. *Journal of Advanced Transportation*, 2022, Article 7772401.
- Zhang, Y., Lin, D., & Mi, Z. (2019). Electric fence planning for dockless bike-sharing services. *Journal of Cleaner Production*, 206, 383–393.
- Zhou, X. (2015). Understanding spatiotemporal patterns of biking behavior by analyzing massive bike sharing data in Chicago. *PLOS One*, 10(10), Article E0137922.
- Zi, W., Xiong, W., Chen, H., et al. (2021). TAGCN: Station-level demand prediction for bike-sharing system via a temporal attention graph convolution network. *Information Sciences*, 561, 274–285.