sustainability

Article

Revealing Daily Mobility Pattern Disparities of Monomodal and Multimodal Travelers through a Multi-Layer Cluster Analysis: Insights from a Combined Big Dataset

Jingyao Zhao [1,]*, Fan Zhang [2], Lei Gao [2], Chunhai Han [2] and Xiongxiong Chen [2]

1 School of Transportation, Southeast University, Nanjing 211186, China
2 College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China; billbronte@nuaa.edu.cn (F.Z.); glzjy@nuaa.edu.cn (L.G.); hanchunhai@nuaa.edu.cn (C.H.); xiongxiongchen@nuaa.edu.cn (X.C.)
* Correspondence: zjyyaoyao@126.com or 230219398@seu.edu.cn

Abstract: More detailed and precise mobility patterns are needed for policies to reduce monomodal automotive dependency and promote multimodality in travel behaviors. Yet, empirical evidence from an integrated view of a complete door-to-door trip mode chain with daily mobility for pattern identification is still lacking. As an improvement and a solution on this issue, a multi-layer cluster model was designed and proposed for distinguishing 20 mobility pattern clusters, including six monomodal traveler groups, two non-transit multimodal traveler groups, and 12 transit multimodal based on big data mining. Statistical analysis with seven indicator measurements and a spatial distribution analysis with the Kernel density GIS maps of travelers' residential location were carried out to reveal significant disparities across pattern clusters concerning spatial, social, and trip characteristics, based on which more precise and target policies for each group were discussed. This research may help provide more detailed information in establishing traveler mobility pattern profiles and solutions in filling the planning–implementation gap from the perspective of planners, policymakers, and travelers.

Keywords: mobility patterns; monomodal; multimodal; cluster analysis; big data mining

Citation: Zhao, J.; Zhang, F.; Gao, L.; Han, C.; Chen, X. Revealing Daily Mobility Pattern Disparities of Monomodal and Multimodal Travelers through a Multi-Layer Cluster Analysis: Insights from a Combined Big Dataset. *Sustainability* **2024**, *16*, 3811. https://doi.org/10.3390/su16093811

Academic Editor: Socrates Basbas

Received: 20 March 2024
Revised: 27 April 2024
Accepted: 30 April 2024
Published: 1 May 2024

1. Introduction

Human mobility is a complex sociotechnical phenomenon facilitated by transportation networks [1]. With the rapid development of urban society, human mobility is facing challenges related to heavy car use, such as traffic congestion, noise, and air pollution. Encouraging the combined use of different travel modes to reduce the dependency of monomodal car use is considered an essential and promising solution, which is also described as increasing multimodality in human mobility [2–4].

Generally, multimodality, or multimodal mobility, is defined as using several modes of transport in a given period of time at the individual level [5,6] or using several modes to complete a trip at the trip level [7]. With massive investment on multimodal transportation infrastructures, there is a general upward trend of multimodality observed in practice, mainly related to the combination use of public transport, car, and active mode [8–10]. However, monomodal car travel still dominates in big cities or metropolises [3]. Thus, a key challenge for transportation planners and policymakers is how to motivate traveler transfer from monomodal car mobility to multimodal mobility, which has been considered a primary objective in travel behavior studies in recent years [2,11].

From a planner or policymaker's perspective, transportation planning needs to be implemented based on the statistical knowledge of mobility demands. Groups or patterns are used for capturing disparities of mobility behavior, so as to better support policymaking for targeted people. Therefore, understanding the nature and differences of monomodal and multimodal mobility patterns is considered as a foundation to help better facilitate the required mode shift towards a more sustainable future [12]. However, compared to monomodal travelers who simply use a single mode irrespective of travel context, multimodal mobility can be regarded as a reflection of the complex choice behavior towards different modes under the context of multimodal transportation systems [13]. It is important albeit more difficult to capture multimodal mobility patterns for understanding behavior differences [14].

The existing literature had given clues to uncover the mobility patterns at an individual level through travel diary surveys. The results varied across studies based on the overall mobility culture of a particular city, such as its car dependency or cycle-friendliness [15]. For example, Diana and Mokhtarian discovered one monomodal car user group, one car-dominant but multimodal user group, and two highly multimodal user groups in France, while one monomodal car user group and three car-dominant but multimodal user groups in the USA [16]. Ton et al. identified one monomodal car user group, one monomodal bicycle user group, one car and walk and bicycle user group, as well as one public transport + user group in the Netherlands [17]. Their findings indicate that to better understand behavior differences and to better implement transportation plannings and policies, it is important to establish local-specific mobility pattern profiles as foundations.

However, there are also arguments about the lower resolution of distinguishing mobility patterns at an individual level. With few groups or patterns, travel behavior disparities discussed in those studies provide general ideas about "fit all" policy implications. It is still not precise enough to prioritize multimodal infrastructure investment properly or to target transport policies efficiently for mode-shifting behavior based on those studies. Also, the survey design considered only primary mode of travel and underrepresented the connecting mode, which may lead to misunderstanding in multimodal transportation modeling, integrated multimodal transport network design and infrastructure planning [14]. For example, nearly every transit trip made involves a mode beyond the transit itself. Planning policies like neighborhood design, parking supply, and interchange pricing programs need more knowledge about the role of "first and last mile" in mobility pattern studies [18,19].

Therefore, there is a growing body of literature aiming to identify multimodality at trip level to fill the knowledge gap. However, to emphasize the connections between different modes during one trip requires more detailed information on trip segments [20]. To capture such features, most research studies have employed a stage-based travel diary survey design at trip level. The results showed that multimodal behavior was highly dependent on how many trip segments were considered or how the mode options were provided in the survey. Generally, the more stage and mode were considered, the more complex multimodal mobility was found to be [1]. For instance, considering three trip stages (main stage, access stage, and egress stage) and one main mode, Yang et al. classified metro commuters into seven detailed groups in Nanjing, China [21]. Considering four main modes and four trip stages, including possible transfer stage, Krygsman and Dijst distinguished more various mode combinations [22].

The above findings have provided an extended knowledge of more detailed mobility patterns. However, two major limitations still need to be addressed. Firstly, even though recent travel diary surveys captured trip stages, the self-reported method often led to undercounts of active modes [18,23]. It may bring imprecise results of multimodal behaviors and an incomplete understanding of mobility patterns [24]. Secondly, connections between mobility patterns obtained at individual level and at trip level are neglected in the existing literature. It is hard to tell how and to what extent the travel modes used in every trip stage become components of individual mobility patterns. Especially in developing countries where transit services are insufficient, the connection stage to transit may play important roles as the main stage in influencing individual travel behaviors.

To solve this, high-resolution and high-quality behavioral data are needed to provide an in-depth view of mobility patterns considering both trip level and individual level multimodalities. Technological innovations in mobile devices and transportation infrastructure have revolutionized travel surveys by utilizing big data like GPS-based data, smartcard-based data, and mobile data positioning records [25]. Such data have enabled researchers to track massive travel trajectories within any period, therefore meeting the need to provide a depth of insight into mobility patterns [26]. Based on this, our research mainly focuses on three aspects. First, to propose an integrated framework for monomodal and multimodal daily mobility pattern identification through big data mining, considering all modes presented in the traveler's complete travel trajectory within a day. Second, to examine and evaluate the possible characteristics, travel behavior disparities, and spatial distributional differences from a comparable perspective across all monomodal and multimodal pattern groups. Third, to provide a more precise view of targeted policies for multimodal transport infrastructure planning and reducing automotive dependency towards more sustainable travel behaviors.

The rest of this paper is organized as follows. Section 2 provides the theoretical framework of this study based on a literature review. Section 3 presents the data source, data mining method, and a multi-cluster analysis model designed for this research. Section 4 displays the clustering results of identifying monomodal and multimodal mobility patterns. Based on previous sections, Section 5 explores spatial and behavioral disparities among all identified mobility pattern clusters through statistical and spatial distribution analyses. Sections 6 and 7 discuss possible policy implications and summarize the main conclusions and limitations.

2. Theoretical Framework

2.1. Understanding Mobility Patterns: Characteristics and Disparities

Characterizing a mobility pattern is critical for understanding the dynamics of travel behaviors and have been a hot topic in recent decades. Normally, different local patterns are driven from travel diary surveys [27]. Most surveys and empirical studies are conducted in developed countries. According to the main focus points of their results, related works can be divided into three categories: (1) mobility pattern only related to mode use, (2) mobility pattern related to sociodemographic features, and (3) mobility pattern related to trip features.

Using statistical analysis, mobility pattern can be divided simply according to travel mode use during the survey time period. Mobility pattern research works at individual and trip levels can be found using this method. For example, using data from the German mobility panel and the mobility status in Germany, Nobis discovered that travelers can be divided into seven groups: Car monomodal (43.0%), Bike monomodal (3.4%), Public transport monomodal (4.9%), Car+Bike multimodal (27.7%), Car+Public transport multimodal (10.5%), Bike+Public transport multimodal (2.6%), and Car+Bike+Public transport multimodal (7.9%) [5]. Based on the Nanjing dataset, Yang et al. divided metro commuters into seven groups: Walk–Metro–Walk (44.0%), Walk–Metro–Bus (10.4%), Bike–Metro–Walk (7.7%), Bike– Metro–Bus (3.4%), Bus–Metro–Walk (18.1%), Bus–Metro–Bus (12.1%), and Car–Metro–Walk (4.3%) [21]. The above research studies provide the basic impressions of monomodal mobility concerning car and active mode use, as well as multimodal mobility concerning different combination uses of car, public transport, and active mode both at individual level and at one trip level.

There are more research works considering classifying mobility patterns beyond mode use. Sociodemographic features, which include objective attributes like age, gender, income, etc., along with subjective attributes like attitudes and perceptions, are most commonly considered in related works. For example, considering the frequency of car, bicycle, train, and BTM (bus, tram, or metro) use, as well as sociodemographic and mode perception variables, Molin et al. discovered five multimodal groups with significant sociodemographic differences in the Netherlands: Car multimodal, Bike multimodal, Bike + Car, Car mostly, and PT multimodal. Young and low-income travelers were more commonly seen in the public transport multimodal group; high-income travelers dominated in Car mostly and Car multimodal groups, while more elderly travelers were found to dominate in the

Bike+Car group [12]. The gender and income influencing effects in mobility pattern choices had also been confirmed by Pani et al. [28] and Buehler and Hamre [6]. More results can be found in mobility pattern distinctions for specific populations, like adolescents or young people [29], and people with large households [30], or when considering certain trip stages like the access stage [31–33].

The above works provide extended information about who and why they belong to monomodal or multimodal groups. However, as sociodemographic differences are discussed rather than travel behavior disparities, the aforementioned research studies failed to answer questions of how the planning or policies can be targeted to more sustainable travel behaviors. Research studies emphasizing the role of trip features in classifying mobility patterns have recently become the most popular. Diana and Mokhtarian, Ton et al., and Schneider et al. introduced trip numbers of each mode (or trip chain complexity) into mobility pattern classifications to evaluate mobility differences in travel frequency and mode [16,17,34]. Kroesen considered trip purpose and frequency, and have discovered that travel frequency differed more significantly than travel purpose between five mobility patterns [35,36]. An et al. examined multimodality disparities across trip purposes, trip distance, and numbers of trip stages, and discovered that trip distance was a more correlated variable contributing to average higher levels of individual multimodality if a trip had at least three stages or was a leisure trip [37]. Yin and Leurent provide a view at integrating trip level mobility with the individual level. Trip distance, daily travel time, trip departure time, and trip mode were used to classify 15 types at the trip level, based on which, six mobility pattern groups were then classified at day level [1].

The above survey-driven research works indicate that mobility pattern is better and more precise to be studied in multi-dimensional ways than simply predefined by mode used. However, even though the number of trip segment is confirmed to have significant disparities across mobility patterns, the role of the mode used for every trip segment has not been evaluated in mobility pattern studies. The door-to-door mode features may provide new evidence for more detailed profiles. Such diverse trip details are also considered as big advantages in travel trajectory study utilizing big data-driven methods, which will be emphasized in this study.

2.2. Identifying Mobility Patterns: Indicators and Methods

When considering beyond the travel mode itself, we can see that the existing literature highlighted regression analysis and cluster analysis in mobility pattern studies. both methods are considered to have great ability to deal with multiple indicators.

Generally, regression analysis is considered as a supervised method that is mainly used to identified determents of mobility patterns when the classifications are predefined. We can refer to the research of Kuhnimhof et al. [9], Klinger [15], Buehler and Hamre [38], and Mao et al. [39] as examples. Their research works provided valuable clues about indicators that can affect mobility or mode choices, mainly including sociodemographic indicators (from both objective and subjective aspects), built environment or spatial indicators and trip or journey characteristic indicators [40,41].

De Wittle et al. have comprehensively reviewed the indicators of determent. Sociodemographic indicators have shown to shape travelers' situations and social interactions, including age, gender, education, occupation, income, household, and car ownership in the objective aspect, as well as experiences, familiarity, habits, perceptions, and attitude in the subjective aspect. The built environment or spatial indicators characterize the unique environment of travelers' daily trips, including density, diversity, public transport availability, and parking. The trip or journey characteristic indicators include trip distance, travel time, cost, and trip number or complexity [42]. Their findings regarding indicators provide understandings about different dimensions of mobility patterns, which can be considered as foundations for mobility pattern classification.

Cluster analysis, however, is considered as an unsupervised method that can be used for finding homogeneous groups within multivariate data [43], which is the most commonly used method for identifying mobility patterns. According to specific means used in cluster analysis, existing research works can be divided into latent class cluster analysis and K-means cluster analysis. The former is normally seen in survey-driven mobility pattern studies, such as those conducted by Kroesen [35], Molin et al. [12], and Ton et al. [17]. It is considered as a model-based clustering technique with the benefit of using statistical criteria to determine the optimal number of classes. Each individual has a probability to belong to each class, based on its characteristics considered [44]. The structure part of latent class analysis allows covariates to predict the class membership of individuals, while the measurement part allows the latent class analysis to explain associations between indicators [45]. However, there are also limitations. For example, the computational complexity of this method often makes it unsuitable for large datasets. In addition, if the initial choice is not an accurate reflection of the internal structure of the data, it may lead to suboptimal results [46].

The K-means cluster is a more powerful method when dealing with massive data, which is becoming popularly used in big data-driven transportation pattern classification studies. We can see the typical applications in air passenger grouping [47], tourist pattern grouping [48], or travel purpose classification [49] based on traffic big data. Despite the obvious advantages, one of the inherent challenges is the need to specify the number of K clusters in advance. The algorithm is also sensitive to the placement of the initial center point, which may lead to convergence to the local optimal solution. To alleviate this problem, it is recommended to use different center point initialization for multiple iterations [47].

Compared to regression analysis, cluster analysis tends to deal with only core indicators that can affect pattern choice significantly. Seven indicators have been considered as highly emphasized ones in the literature. Age, income, access to public transport, trip purpose, trip distance per mode, travel time, and number of trips per mode are also taken into consideration as indicators in this study.

3. Data and Methodology

3.1. Extracting Daily Trip Features: Big Data Processing and Mining

Trajectory data like GPS-base data and smartphone location data are known for their large sample size and continuous behavior observations, which allows researchers to identify travelers' temporal and spatial regularities hidden in datasets [50,51]. However, because of complex data structures and privacy protection, such data also require careful processing procedures before being usable for trip analysis [52]. The most commonly used data are mobile phone GPS or signaling data for travel behavior analysis (see [53,54] as examples) and smartcard data for capturing transit modal behavior separately. There is also an increasing trend for the combined use of enormous datasets, especially in Chinese cities where the massive population requires a more considerable amount of data and has shown excellent efficiency in improving data precision (see [55,56] as examples).

In this research, we have achieved access to both mobile phone signaling data and public transportation smartcard origins and destination station data for rail and bus transport modes in Nanjing, China, and can set up a combined dataset including seven statistical data sources and three dynamic big data sources; the dataset structures are presented in Table 1. To capture monomodal and multimodal travelers' daily mobility patterns and travel behaviors, we extracted one traveler's trip origin and destination (OD) paths, door-to-door travel modes, and trip purposes within a day from the combined big dataset. The data processing progress can be divided into five steps, as follows:

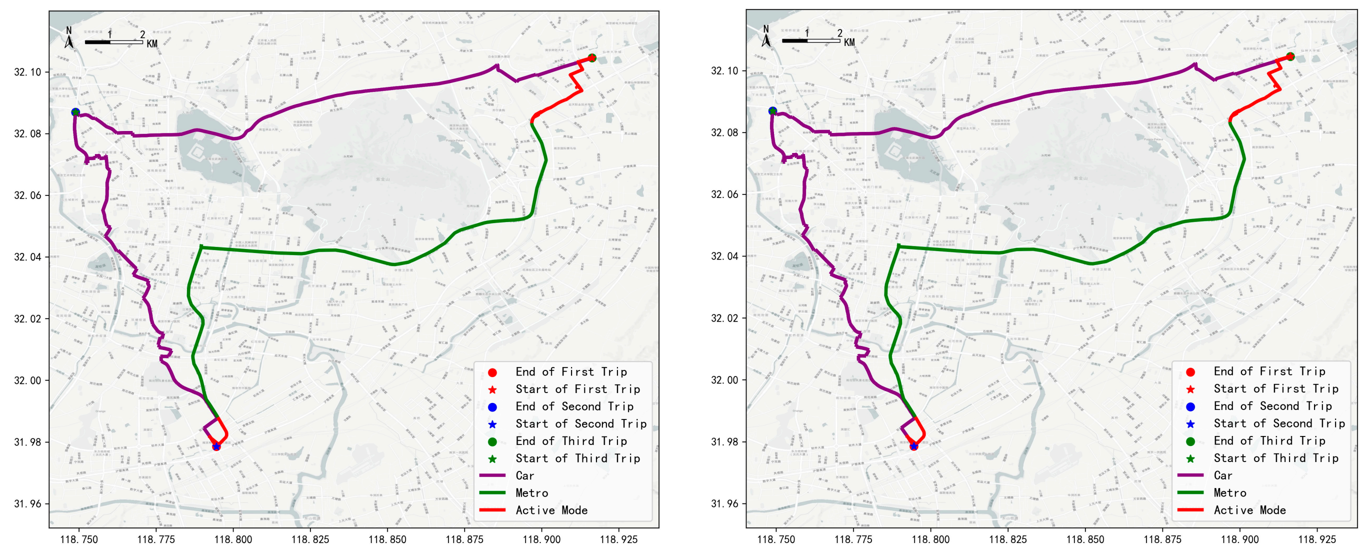
(a) The data washing step. Five kinds of abnormal data were washed before our dataset could be used for further analysis. Excluding conventional dirty data, ping-pong switching data, and drift data, we have followed Ding et al.'s data processing framework for the mobile phone signaling dataset [57]. Also, incomplete or systematically misrecorded data were excluded from the public transportation smartcard dataset. After this step, these two big datasets can provide mutual authentication evidence about individual transit use with the same time-stamp and stop location for getting on and off.

(b) The map matching step. The mobile signaling data use the location coordinates of the base station to approximately replace the actual coordinates of the user's travel; it is necessary to match the user's movement trajectory recorded by the base station with the road network and convert it into the user's travel trajectory. Considering the amount of data and processing efficiency, we use a combination of geometry-based and topology-based methods for processing. Firstly, a regional road topology network was established. Then, the recorded base station points were matched to the nearest road network nodes by the point-to-point neighbor matching method. Finally, the shortest path between continuous multiple nodes on the topology road network was formed to present the traveler's real-world travel trajectory.

(c) Trip segment step. After map matching, all travel trajectories were split into specific trip segments with characteristics like trip originals and destinations, travel purposes, and door-to-door travel modes. According to local Points of Interest (POI) features of Nanjing city, we have considered four purposes: maintaining trip, commuting trip, leisure trip, and return trip. According to existing research, we have identified trajectory stopping points as origins and destinations of each trip based on the density cluster method and stopping time and have assigned four specific trip purposes according to the weight of each POI type within the trip origin and destination buffer areas [58–60].

(d) Door-to-door modal split and trajectory correction step. According to the existing literature, metro, bus, car, and active modes are the most used travel modes throughout the day and were selected in our research. We established specific classification rules based on short-term velocity and acceleration to split the active mode and the others first. Then, we used public transportation smartcard data to check for metro or bus trips and modified the trip trajectory according to metro and bus lines if needed.

(e) Daily trip chain extraction step. After all the travel trajectories in the combined big dataset were processed, the daily trip chain of all individuals can be extracted from the track records marked with travel purposes and door-to-door travel modes. The typical extraction results with three trips are presented in Figure 1.



(**a**) Travel purpose of each trip (**b**) Door-to-door travel mode of each trip

Figure 1. Travel purposes and door-to-door modes of an individual with three trips.

Table 1. The combined big dataset used in this study and data formats.

Type	Name	Source	Content	Original Data Volume
Statistical dataset	Distribution data of communication base stations	Mobile operator	Includes base station ID number, base station longitude, and base station latitude	163,721 records
	Urban POI distribution data	Amap (name of a Chinese map operate)	Includes 14 categories of Points of Interests (POIs), including transportation facilities, leisure and entertainment, companies and enterprises, healthcare, commercial residential areas, tourist attractions, automotive related, life services, science and education culture, shopping and consumption, sports and fitness, hotel accommodation, financial institutions, and catering and food	322,542 records
	Urban road network data	OpenStreetMap	Includes OpenStreetMap ID (OSMID), one-way traffic, number of lanes, road number (such as S205), road name, road grade, road length, and road geometry	30,377.53 km
	Urban bus network data	Amap (name of a Chinese map operate)	Includes route name, route geometry, station name, station longitude, station latitude, and station ID	713 lines (10,743.64 km) with 6964 stations
	Urban metro network data	Amap (name of a Chinese map operate)	Includes route name, route geometry, station name, station longitude, station latitude, and station ID	11 lines (427 km) with 198 stations
	Resident population distribution data of the community	Human Resources and Social Security Department	Includes community name, geometry, zoning, population, street area, population density, the population aged 0–14, the population aged 15–59, the population aged 60 and above, the population aged 65 and above	Records of 906 communities
	Residential housing price distribution data	Human Resources and Social Security Department	Includes residential housing estate name, average price, zoning, latitude and longitude	4369 records
Dynamic big dataset	Mobile phone signaling data	Mobile operator	Includes user ID, work and residence information, and travel information (travel attribute information, travel trajectory). The travel attribute information includes start and end base stations, start times, end times, and subway travel information (departure and arrival stations, entry and exit times, and route station information)	41,292 individuals, 4,913,488 trajectory points
	Origin and destination station data of urban bus card swiping	Bus Operation Department	Includes record ID number, route name, vehicle number, cost, time, boarding point latitude and longitude, and alighting point latitude and longitude	1,171,137 records
	Origin and destination station data of urban rail card swiping	Rail Operation Department	Includes record ID, entry and exit time, name of starting and ending stations, start and end stations' longitude and latitude, travel time, and distance	1,216,136 records

Note: All the above data strictly follow the data privacy limitation and no personal information is included.

3.2. Classifying Daily Mobility Pattern: A Multi-Layer Cluster Analysis

After multiple door-to-door trip features have been extracted from individuals' complete trajectory, cluster analysis can be used to classify each individual into different mobility patterns. Before this step, however, a statistical analysis of all the mode combinations is needed to obtain the first impression about daily mobility pattern. When considering modes used in every segment of a door-to-door trip, there are 8042 different types of mode combinations recognized from their daily trajectories. Therefore, the K-means cluster method is utilized to identify typical mobility pattern classifications.

Considering the massive combinations, a multi-layer structure is beneficial for studying different features according to the existing literature [1]. Based on this information, we have proposed a multi-layer cluster analysis model with one classification layer and three cluster layers. In the first layer, travelers in the dataset are classified into three classes according to the modal types used in individual daily trips: monomodal travelers, non-transit multimodal travelers, and transit multimodal travelers. In the second, third, and fourth layers, a K-means clustering method is used within every traveler class, considering different indicators according to the features of the class. The overall multi-layer cluster model framework is presented in Figure 2.

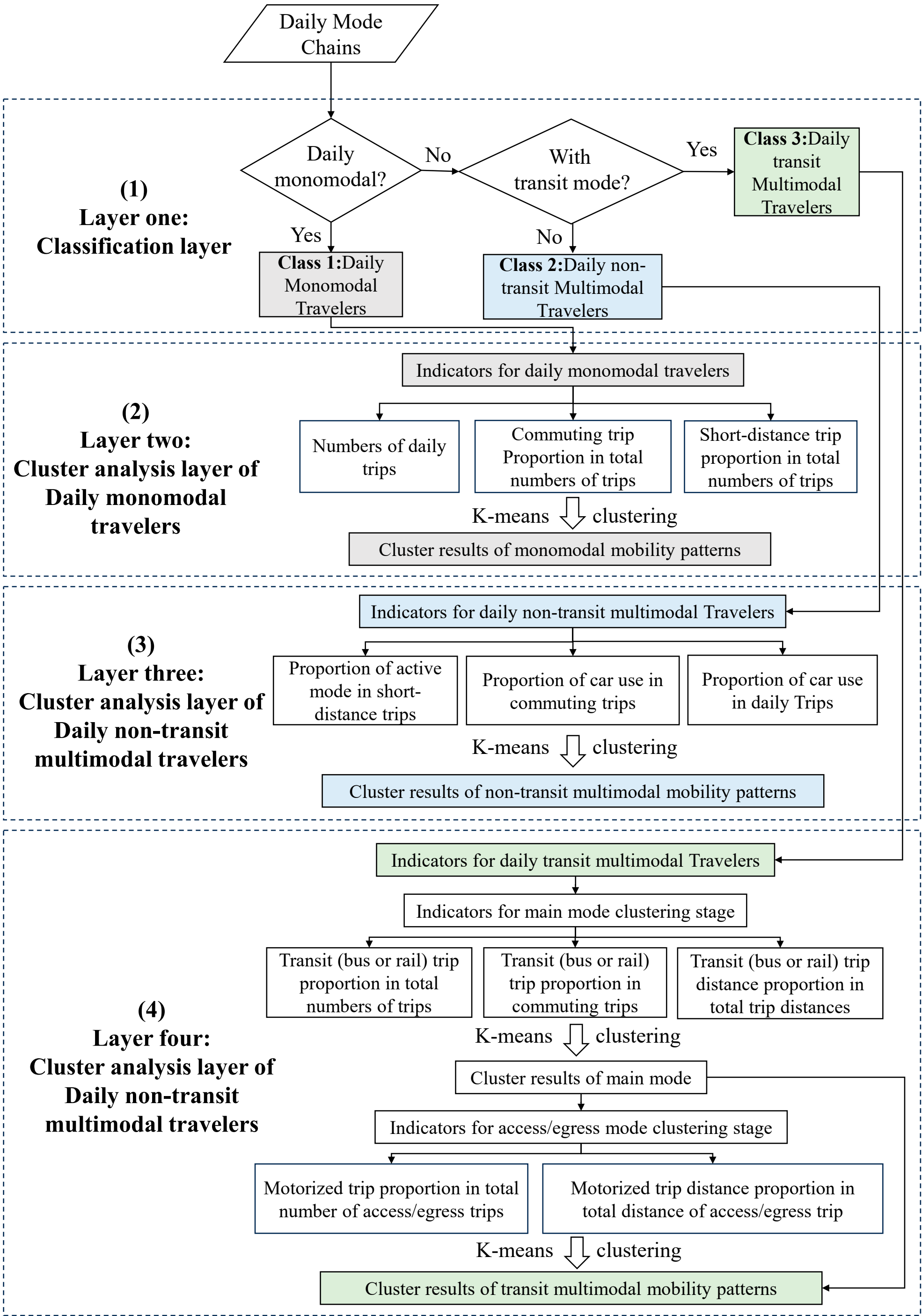


Figure 2. Multi-layer cluster analysis model framework.

Indicators are also important for cluster analysis and the following behavior analysis. We have considered seven indicators that are highly emphasized in the literature. it includes two sociodemographic indicators (age and income), one spatial indicator (access to public transport), and four trip characteristic indicators (trip purpose, trip distance per mode, travel time per mode, and number of trips per mode). However, because of data privacy, we are not authorized to access travelers' profiles like age or income. To fix the lock-in situation, we are inspired by Sulikova and Brand's research in which they introduce neighborhood sociodemographic distribution characteristics into spatial indicators and found the propensity of like-minded people to self-select into neighborhoods with similar incomes and cultural backgrounds [61]. Following this method, we have considered residential house pricing and community elderly status as indicators that can reflect both spatial and sociodemographic profiles to some extent. However, those non-intuitive indicators are more typically used in behavior analysis than in cluster analysis.

Therefore, we have considered three trip characteristic indicators used for defining daily mobility patterns according to the existing literature: trip purpose, distance per mode, and number of trips per mode [45]. Cross-indicator design is often used in cluster analysis to reduce data dimensions, such as number of trips per mode for different travel purposes [36]. Considering access and egress trip stages, short-distance trips are also emphasized in cross-indicator design. With massive data, the wide-range continuous indicators of distance per mode and number of trips per mode are not suitable to be directly input into a cluster analysis. We have replaced them with proportion indicators ranging between 0 and 100. After cross-indicator design, 11 specific indicators are considered in the following cluster analysis process.

For the daily monomodal travelers, with their uniform use of a single travel mode like active mode or automobile, we have considered their number of daily trips, commuting trip proportion in total numbers of trips, and short-distance trip proportion in total number of trips as three cluster indicators.

For daily multimodal travelers with non-transit modes, we have considered the proportion of active mode in short-distance trips, the proportion of car use in commuting trips, and the means of car trip proportion as three cluster indicators.

For daily multimodal travelers with transit modes, we consider a two-stage method of the main mode clustering and then the access/egress mode clustering in sequence. In the main mode clustering stage, we consider three indicators: transit (bus or rail) trip proportion in total number of trips, transit (bus or rail) trip proportion in total number of commuting trips, and transit (bus or rail) trip distance proportion in full trip distances. In the access/egress mode clustering stage, we consider two indicators: motorized connection trip proportion in total number of access/egress trips and total distance of access/egress trip.

4. Cluster Results of Daily Mobility Patterns

Tables 2–4 present the clustering results from the multi-layer cluster analysis model and their indicator frequencies using the Nanjing dataset. There are twenty identified clusters of daily mobility patterns, including six clusters in the monomodal traveler class, two clusters in non-transit multimodal traveler class and twelve clusters in transit multimodal traveler class. When considering and capturing the main mode's access and egress connection stage through big data mining, we discovered a much higher proportion of multimodal travelers in the Nanjing dataset (85.6% of total travelers obtained).

Table 2. Cluster results and indicator frequency of monomodal traveler group.

Mobility Pattern Clusters	Traveler Labels	Proportion in Monomodal Traveler Class	Average Number of Trips per Day	Means of Commuting Trip Proportion	Means of Short-Distance Trip Proportion
exclusive active mode user with low flexibility	LAM	15.94%	2.7	65.34%	93.93%
exclusive active mode user with high flexibility	HAM	21.38%	2.7	15.11%	92.87%
exclusive car user with low flexibility and low intensity	LLC	19.61%	2.6	58.21%	95.57%
exclusive car user with low flexibility and high intensity	LHC	11.85%	2.5	57.19%	26.87%
exclusive car user with high flexibility and low intensity	HLC	19.7%	2.7	19.62%	96.35%
exclusive car user with high flexibility and high intensity	HHC	11.51%	2.6	23.84%	27.65%

Table 3. Cluster results and indicator frequency of non-transit multimodal traveler group.

Mobility Pattern Clusters	Traveler Labels	Proportion in Non-Transit Multimodal Traveler Class	Proportion of Active Mode in Short-Distance Trip	Proportion of Car Use in Commuting Trip	Means of Car Trip Proportion
highly car-dominant users	HC+AM	56.85%	52.91%	86.24%	87.18%
moderately car-dominant users	MC+AM	43.15%	78.86%	38.21%	58.03%

To verify the results' effectiveness for further analysis, we checked the total modal share with the 2022 Nanjing household travel survey results. When considering trip main mode only, our results have shown high consistency with automobile use, while a 3% higher proportion with transit use and 3% lower of active mode than the self-reported survey, which validates the accuracy of trip stage segments and modal split for individuals' daily travel trajectory in our dataset. Specifically, our results showed 14.4% of monomodal travelers, 53.4% of non-transit multimodal travelers, and 31.85% of transit multimodal travelers. These results indicate that, with combinations of active mode and automobile use on a daily scale becoming the most common in Nanjing, it is crucial to understand mobility patterns in this multimodal class for potential mode shift, which is often neglected in the existing literature. Meanwhile, with twelve mobility pattern clusters identified and a second large proportion, the results also show that transit multimodal travelers have much more complicated characteristics that may attach to more precise policy implications. The specific monomodal and multimodal mobility pattern clusters are described in detail in the following.

Table 4. Cluster results and indicator frequency of transit multimodal traveler group.

Mobility Pattern Clusters	Traveler Labels	Proportion in Transit Multimodal Traveler Class	Means of Rail Trip Proportion in Total Number of Trips	Means of Rail Trip Proportion in Commuting Trips	Means of Rail Trip Distance Proportion in Total Trip Distances	Means of Bus Trip Proportion in Total Number of Trips	Means of Bus Trip Proportion in Commuting Trips	Means of Bus Trip Distance Proportion in Total Trip Distances	Means of Motorized Trip Proportion in the Total Number of Access/ Egress Trips	Means of Motorized Trip Distance Proportion in Total Distance of Access/Egress Trip
less frequent bus users with low motorized connections	LBLM	33.28%	0.46%	0.04%	0.23%	24.46%	5.96%	10.14%	0.75%	2.41%
moderately bus-dominant users with low motorized connections	MBLM	22.82%	1.17%	0.36%	0.96%	51.12%	30.93%	40.89%	1.38%	2.46%
highly bus-dominant users with low motorized connections	HBLM	9.43%	0.54%	0.10%	0.33%	77.95%	41.37%	89.52%	0.51%	3.11%
less frequent rail users with low motorized connections	LRLM	2.48%	24.57%	7.66%	19.69%	4.22%	1.23%	1.62%	15.98%	60.14%
less frequent rail users with moderately motorized connections	LRMM	8.00%	32.40%	6.77%	20.53%	2.59%	0.79%	1.00%	46.93%	89.29%
rail users with highly motorized connections	LRHM	1.22%	34.89%	13.38%	17.60%	0%	0%	0%	94.78%	99.57%
moderately rail-dominant users with low motorized connections	MRLM	5.35%	50.05%	25.52%	56.34%	6.95%	4.36%	3.01%	37.98%	80.70%
moderately rail-dominant users with moderately motorized connections	MRMM	4.06%	55.11%	28.68%	58.86%	4.19%	3.24%	2.11%	63.85%	93.70%
moderately rail-dominant users with highly motorized connections	MRHM	1.43%	51.57%	29.44%	45.26%	1.02%	0.63%	0.82%	90.27%	98.59%
highly rail-dominant users with low motorized connections	HRLM	2.52%	71.71%	34.15%	91.54%	0.98%	0.12%	0.28%	34.81%	69.32%
highly rail-dominant users with moderately motorized connections	HRMM	6.35%	82.28%	39.69%	95.13%	1.08%	0.12%	0.34%	64.57%	93.84%
highly rail-dominant users with highly motorized connections	HRHM	3.04%	82.12%	62.28%	95.67%	0.27%	0%	0.05%	90.35%	98.69%

4.1. The Monomodal Class with Six Mobility Patterns

Six mobility patterns are identified in the monomodal class: exclusive active mode user with low flexibility (LAM), exclusive active mode user with high flexibility (HAM), exclusive car user with low flexibility and low intensity (LLC), exclusive car user with low flexibility and high intensity (LHC), exclusive car user with high flexibility and low intensity (HLC), and exclusive car user with high flexibility and high intensity (HHC).

With similar numbers of daily trips and trip distances, significant differences can be observed from two exclusive active mode clusters by indicating the commuting trip proportion. The LAM group has a much higher commuting trip proportion than the HAM group, which reflects more fixed routinary and less flexible travel tendencies in the LAM group. Among the four exclusive car users, the LLC group mainly engages in short-distance commuting trips; the LHC group mainly engages in long-distance commuting trips; the HLC group mainly engages in short-distance -maintaining or leisure trips, while the HHC group mainly engages in long-distance maintaining or leisure trips.

These four groups reflect different car travel habits; the LHC group has the lowest travel frequency but the highest dependency on car use, while the HLC group reflects the most unreasonable use of cars with the highest travel frequency and short-distance trip proportion.

4.2. The Non-Transit Multimodal Class with Two Mobility Patterns

Two mobility patterns are identified in the non-transit multimodal class: highly car-dominant users (HC+AM) and moderately car-dominant users (MC+AM). Both groups use the car for more than half of their daily trips and active mode for the rest of the day. Specifically, the HC+AM group uses automobiles for 87.18% and especially 86.24% of commuting trips, much higher than the same indicators in the MC+AM group. Higher car usage is also discovered in the HC+AM group for short-distance trips. Only 52.91% of short-distance trips are conducted by active mode, indicating more irrational car use in the HC+AM group than in the MC+AM group. In other words, the MC+AM group uses active mode more frequently on short-distance and commuting trips.

4.3. Transit Multimodal Class with Twelve Mobility Patterns

Considering both rail and bus travel modes and their access as well as egress trip modes, there are twelve mobility pattern clusters in the transit multimodal class, including three bus multimodal user groups and nine rail multimodal user groups: less frequent bus users with low motorized connections (LBLM), moderately bus-dominant users with low motorized connections(MBLM), highly bus-dominant users with low motorized connections (HBLM), less frequent rail users with low motorized connections (LRLM), less frequent rail users with moderately motorized connections (LRMM), less frequent rail users with highly motorized connections (LRHM), moderately rail-dominant users with low motorized connections (MRLM), moderately rail-dominant users with moderately motorized connections (MRMM), moderately rail-dominant users with highly motorized connections (MRHM), highly rail-dominant users with low motorized connections (HRLM), highly rail-dominant users with moderately motorized connections (HRMM), and highly rail-dominant users with highly motorized connections connections (HRHM). From the total proportion distribution, three bus multimodal user groups account for 65% of transit multimodal travelers, and nine rail multimodal user groups account for the remaining 35%.

All twelve groups show combinations of at least three travel modes. Still, only low motorized connections are identified in three bus user groups, indicating the absolute tendency to use active mode for bus trips’ access and egress stage. The LBLM, MBLM, and HBLM groups mainly differ in their dependency on bus travel. An upward trend can be seen in all three indicators of bus trip proportion in total number of trips, bus trip proportion in total number of commuting trips, and bus trip distance proportion in the three groups. Decreased trends can be seen in the bus trip proportion of commuting trips compared to daily trips within all three multimodal bus groups. However, a much more significant decrease occurred in the HBLM group than the other two groups, indicating that the increasing usage of buses was mainly conducted for leisure or maintaining trips rather than commuting. Meanwhile, a higher dominant status of bus trip distance than bus trip numbers on the daily level in the HBLM group also reveals longer travel distances by bus with active mode connections in this group.

Rail travelers choose more diverse access and egress modes than bus travelers. The LRLM, MRLM, and HRLM groups mainly use active mode for access and egress with a deepened dependency on daily rail travel, while bus or automobile access and egress trips are more common in the LRMM, MRMM, and HRMM groups. More than 90% of travelers in the LRHM, MRHM, and HRHM groups use buses or automobiles for their connection stages to rail trips. We can see increasing trends of all three indicators in rail multimodal user groups with similar motorized connection levels like the LRLM, MRLM, and HRLM; while such indicators are nearly at the same level in groups with equal rail use frequency, it indicates that the all three indicators play essential roles in classifying rail multimodal mobility patterns. The same phenomenon of a decline in proportion between rail usage in daily trips and commuting trips is also discovered in all nine rail multimodal traveler groups. However, more significant decreases are obtained in the HRLM group than in the LRLM and MRLM groups. This indicates that higher rail usages in groups with low motorized connections are mainly conducted in leisure or maintaining trips. The same trend is also discovered in groups with moderately motorized connections. However, we can see a similar drop of around 20% between rail usage on daily trips and commuting trips in the LRHM, MRHM, and HRHM groups. This indicates that a higher use of rail with highly motorized connections concerns all travel purposes, including commuting trips.

5. Spatial and Behavioral Disparities among Mobility Pattern Clusters

The calculation results of indicators for all twenty identified mobility pattern clusters are presented in Table 5. Also, with the advantages of big data mining, we hope to discover and exhibit the possible spatial differences among different travel classes and mobility pattern clusters more visually. A spatial distribution analysis of travelers' residential location was then carried out and is presented in Figures 3–7 to provide supplemental evidence with statistical analysis.

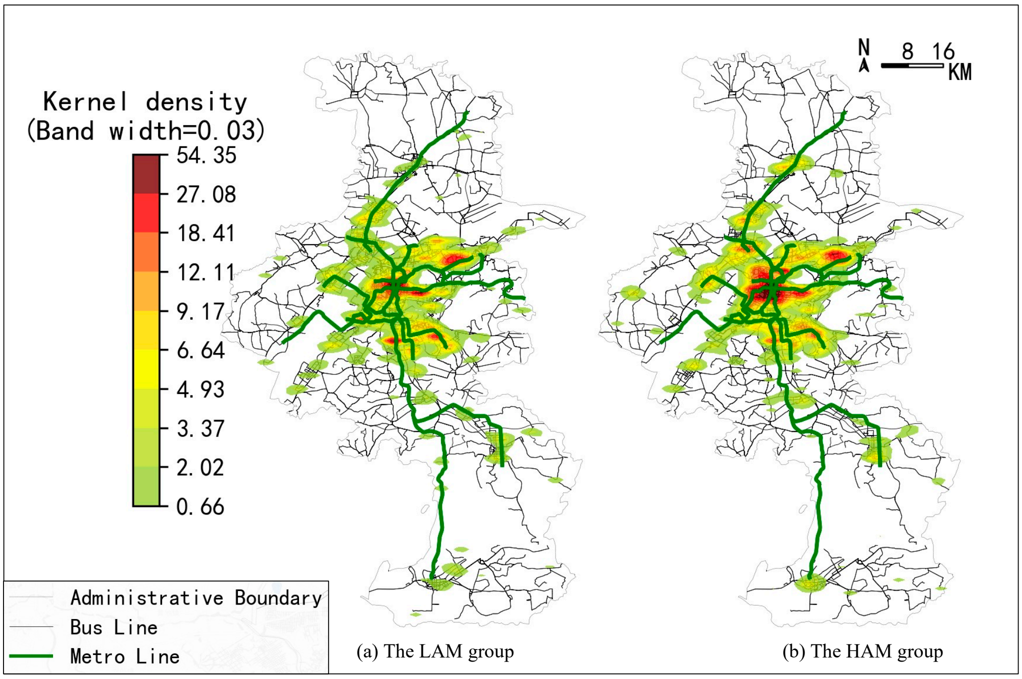


Figure 3. Spatial distributions of two exclusive active mode user groups.

Table 5. Indicator estimation results of twenty typical mobility pattern clusters.

Typical Mobility Pattern Cluster		Spatial and Social Indicators				Trip Characteristics' Indicators		
		Accessibility to the Transit Network on the Home End within a 5 min Walk	Accessibility to the Transit Network on the Work End within a 5 min Walk	Level of Residential House Price	Level of Community Elderly Status	Average Number of Trips per Day	Average Trip Distance	Average Travel Time per Trip (min)
exclusive active mode user with low flexibility	LAM	70.3%	66.3%	above average	low	2.7	2.46	10
exclusive active mode user with high flexibility	HAM	75.5%	74.8%	above average	high	2.7	2.14	13
exclusive car user with low flexibility and low intensity	LLC	72.5%	66.7%	below average	medium	2.6	15.99	21
exclusive car user with low flexibility and high intensity	LHC	69.7%	57.7%	below average	medium	2.5	54.08	39
exclusive car user with high flexibility and low intensity	HLC	76.5%	72.7%	below average	medium	2.7	15.68	21
exclusive car user with high flexibility and high intensity	HHC	70.9%	69.7%	below average	high	2.6	57.13	44
highly car-dominant users	HC+AM	76.0%	73.7%	average	medium	3.9	25.96	33
moderately car-dominant users	MC+AM	74.0%	72.4%	average	medium	3.7	11.58	25
less frequent bus users with low motorized connections	LBLM	84.6%	83.9%	average	high	5.2	16.42	26
moderately bus-dominant users with low motorized connections	MBLM	85.1%	85.7%	above average	high	3.9	10.59	22
highly bus-dominant users with low motorized connection	HBLM	87.6%	88.1%	above average	high	3.2	5.96	17
less frequent rail users with low motorized connections	LRLM	79.9%	82.1%	above average	medium	5.1	15.41	32
less frequent rail users with moderate motorized connections	LRMM	79.7%	80.0%	above average	medium	4.5	25.83	40
rail users with highly motorized connections	LRHM	77.0%	74.8%	below average	medium	4.3	32.95	40

Table 5. *Cont.*

Typical Mobility Pattern Cluster		Spatial and Social Indicators				Trip Characteristics' Indicators		
		Accessibility to the Transit Network on the Home End within a 5 min Walk	**Accessibility to the Transit Network on the Work End within a 5 min Walk**	**Level of Residential House Price**	**Level of Community Elderly Status**	**Average Number of Trips per Day**	**Average Trip Distance**	**Average Travel Time per Trip (min)**
moderately rail-dominant users with low motorized connections	MRLM	81.2%	82.7%	above average	medium	3.9	16.69	37
moderately rail-dominant users with moderately motorized connections	MRMM	82.6%	81.0%	average	medium	3.6	23.81	41
moderately rail-dominant users with highly motorized connections	MRHM	71.1%	76.1%	average	low	3.5	26.25	37
highly rail-dominant users with low motorized connections	HRLM	83.5%	84.6%	above average	medium	3.0	14.32	37
highly rail-dominant users with moderately motorized connections	HRMM	82.1%	86.1%	average	medium	2.8	24.59	47
highly rail-dominant users with highly motorized connections	HRHM	81.1%	80.8%	average	medium	2.5	26.02	45

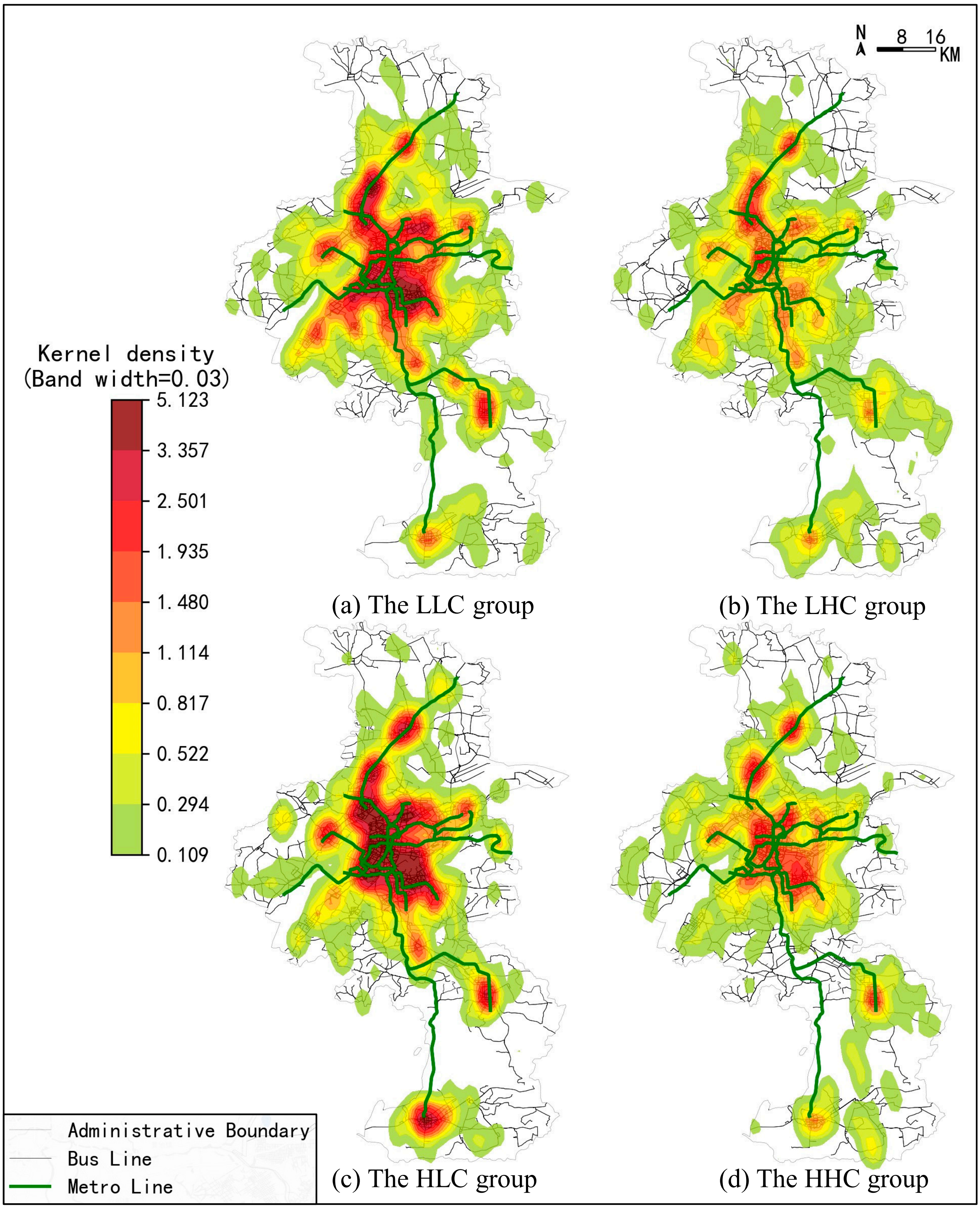


Figure 4. Spatial distributions of four exclusive car user groups.

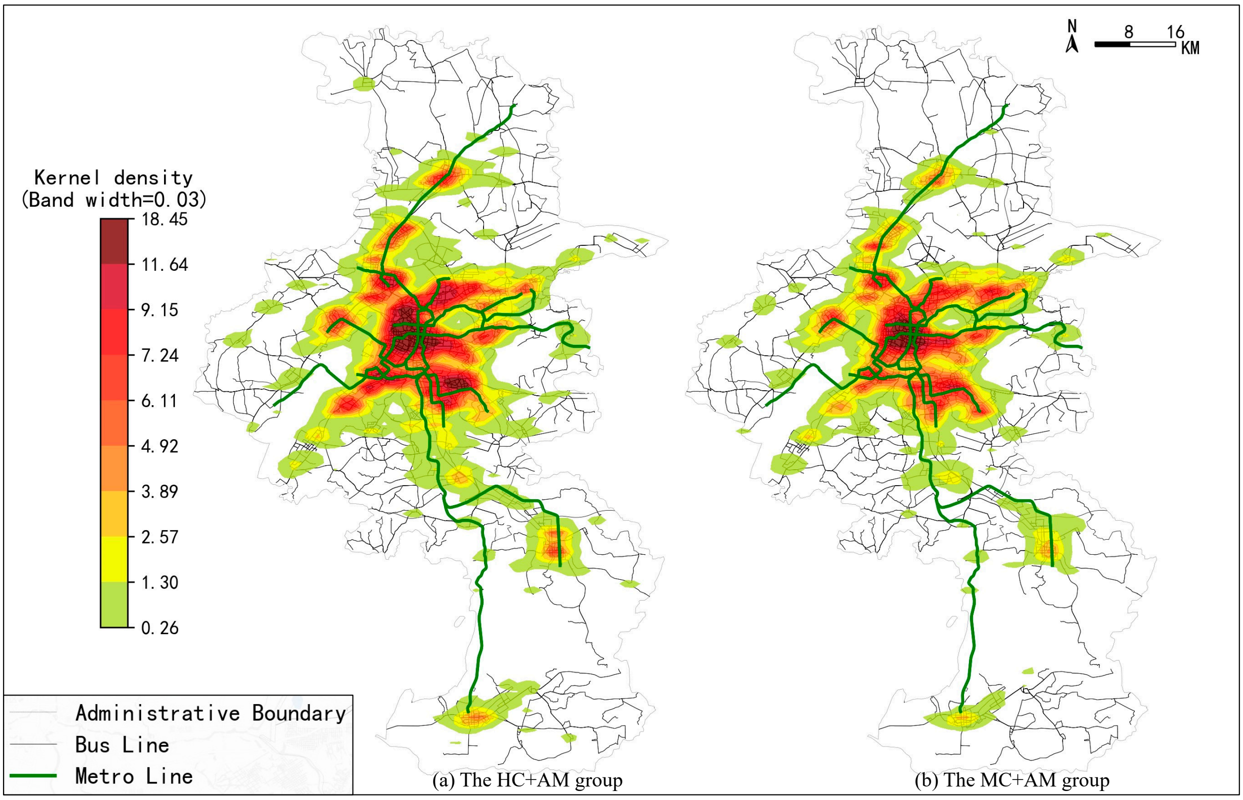


Figure 5. Spatial distributions of two non-transit multimodal user groups.

5.1. Statistical Analysis of Disparities among Mobility Pattern Clusters

Significant differences across the identified mobility pattern clusters are found according to the results in Table 5. Among spatial and social environment indicators, accessibility to transit networks at home is more regularly distributed among mobility patterns than at the work end. Higher levels of accessibility to transit networks are discovered in all twelve transit multimodal mobility patterns except the MRHM group. We can see the highest level of transit accessibility from both home and work ends in the LBLM, MBLM, and HBLM groups, with an incremental tendency in accessibility and bus dependency level. The same tendency can also be seen in rail multimodal traveler groups. Taking the LRLM, MRLM, and HRLM groups as examples, we can see an increasing trend in home end transit accessibility (79.9%, 81.2%, 84.5%) along with the growing proportions of rail trips. It indicates that transit infrastructure improvement will benefit daily bus and rail use, especially in residential areas.

Meanwhile, a decreasing trend of home end transit accessibility is obtained among rail multimodal users with the same level of rail use frequency, which may explain the upward level of motorized connections. There is no apparent tendency about accessibility indicators that can be seen from mobility patterns in monomodal and non-transit multimodal classes. However, among all identified clusters, the LHC group has the lowest level of transit accessibility (69.7% on the home end, 57.5% on the work end) and relatively longer trip distance, which may explain their high dependency on cars.

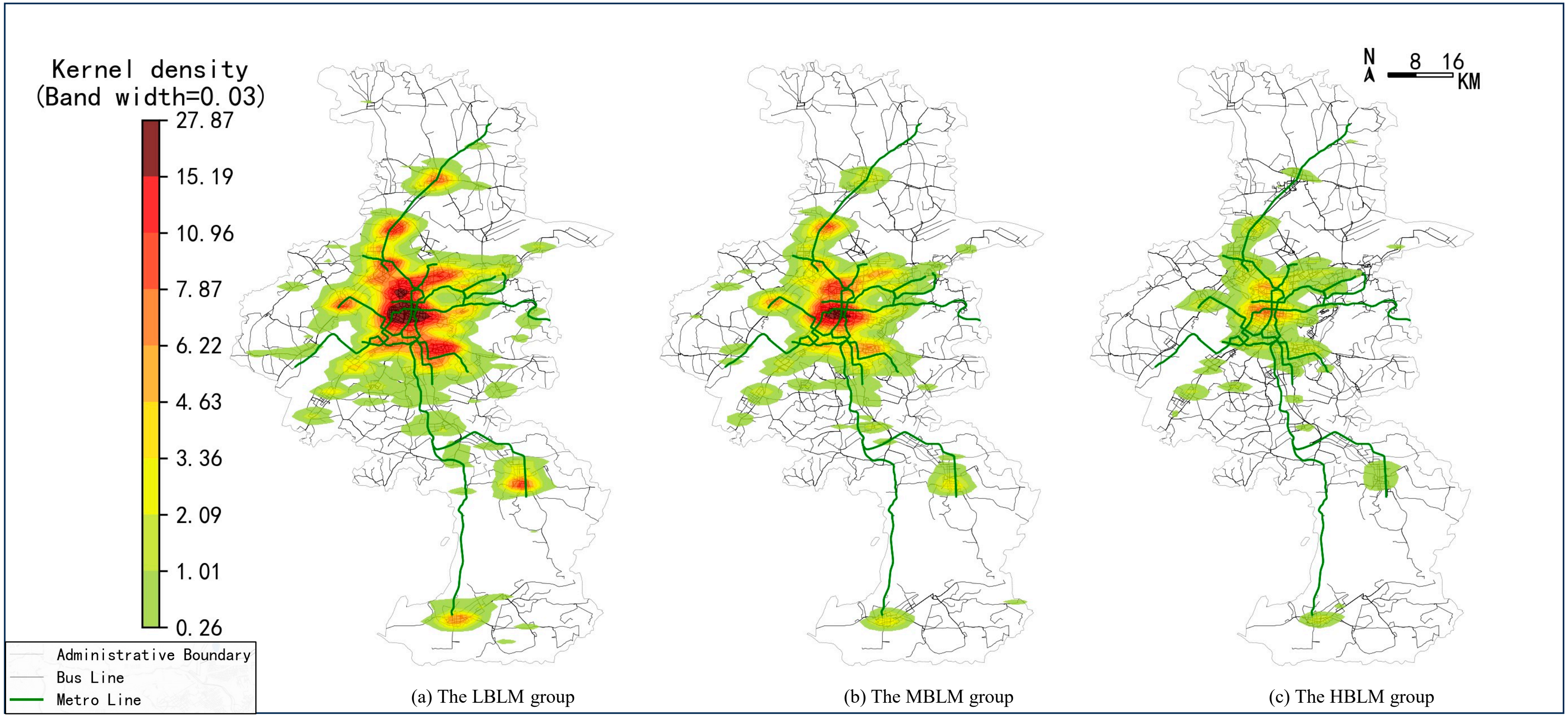


Figure 6. Spatial distributions of three bus multimodal user groups.

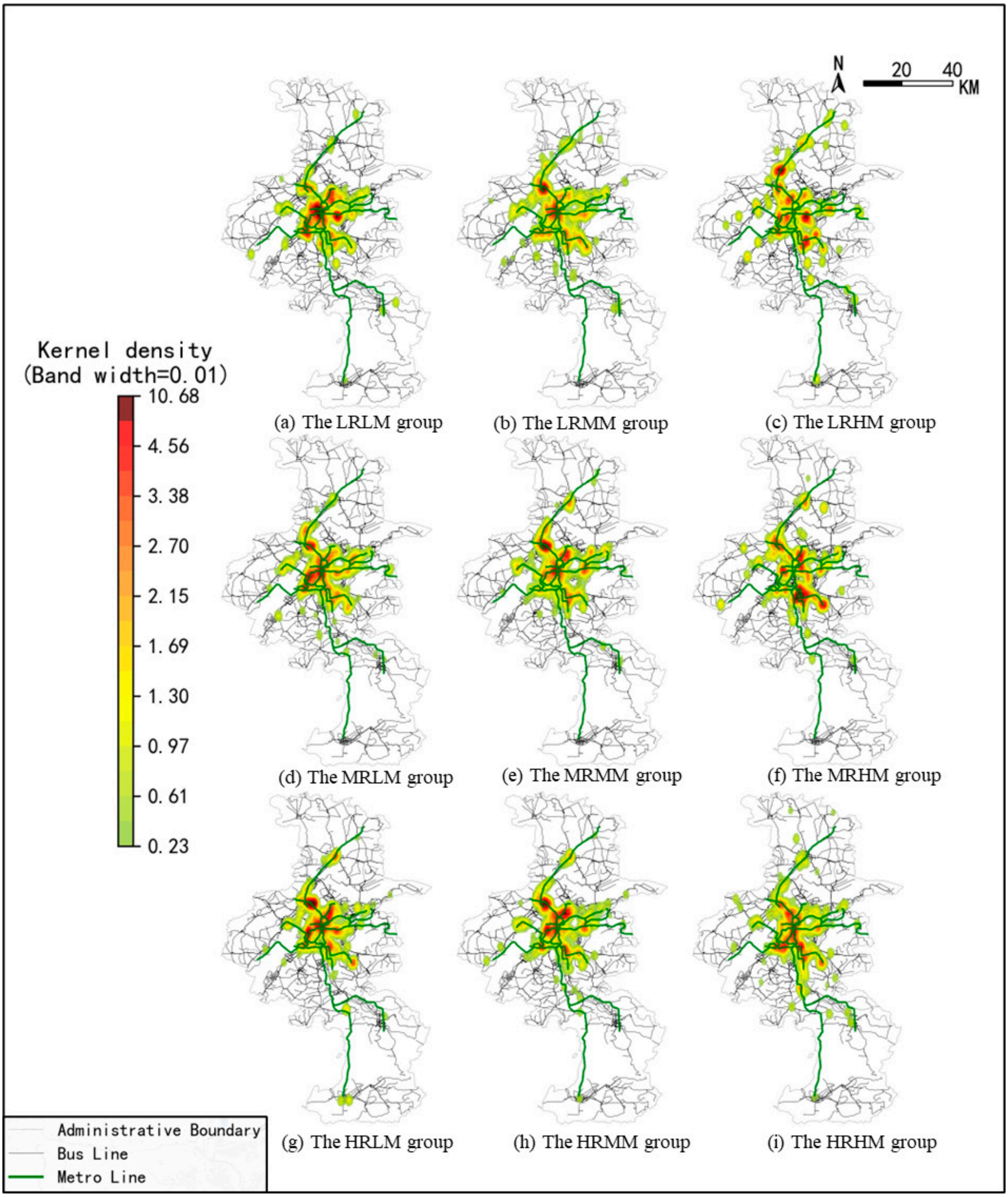


Figure 7. Spatial distributions of nine rail multimodal user groups.

The level of residential house prices also points to an essential hint that exclusive car users are more dispersed in suburban areas of Nanjing, with lower-than-average levels of house prices discovered in all four exclusive car user groups of LLC, LHC, HLC, and HHC. Consistent results can be discovered in rail multimodal user groups with highly motorized connections; only below-average and average levels of residential house pricing are found in the LRHM, MRHM, and HRHM groups. With lower accessibility to transit networks in their residential areas, this result may indicate unwilling dependencies on car use for both daily travel and access trips in Nanjing city.

According to the average proportion of residents above 65 years old in Nanjing city (nearly 15%), we give value labels of high (20% and above), medium (15–20%), and low (below 15%) to the identified groups according to their residential community average proportion of residents above 65 years old. The HAM and HHC groups are found to have higher levels of elderly status, which, to some extent, reflects the high flexibility described in such groups. The three groups of bus multimodal travelers also tend to live in elderly communities, while the LAM user group and one group of rail multimodal users (MRHM) are found to have lower elderly status, with similar levels of poor accessibility to transit networks but average or above average house pricing. It gives a clue that those two groups may live in new residential areas with distances to transit networks in the main city or city center, which indicates an urgent need for special transit services in those areas.

Disparities can also be observed from three indicators of trip characteristics. All six mobility pattern clusters in the monomodal traveler class tend to have simpler daily trip chains than non-transit multimodal and transit multimodal travelers according to the indicator of an average number of trips per day. This is consistent with Schneider et al.'s (2020) and Yin and Leurent's (2022) research that a higher level of multimodality may relate to a more complex trip chain [1,34]. Also, we can see decreased trends of trip chain complexity with the uprising proportion of transit use within three bus multimodal user groups as well as within nine rail multimodal user groups. There are also decreased trends of trip chain complexity within groups having similar proportions of transit use but with an upward frequency of motorized connections. This indicates that a higher dependency on transit use with motorized connections may reduce daily travel demand. A possible reason for this may be that more transfers have to be made using the combination of transit and motorized access or egress mode during one trip.

The distribution of average trip distances and travel time among different mobility pattern clusters have revealed less regularity but also important evidence. There is no doubt that the shortest trip distance and travel time are discovered in two exclusive active mode user groups, LAM and HAM. With similar trip numbers per day, the HAM group, however, has a shorter trip distance but a slightly longer trip duration than the LAM group. With a much higher level of community elderly status, it may indicate that the improvement of active mode infrastructures like bike lanes may be needed around the HAM group residential community. In exclusive car user groups, the LHC and HHC groups have nearly four times longer trip distances but only two times longer travel time than the LLC and HLC groups, indicating less-congested travel situations for trips of the LHC and HHC groups, whereby the mode shift for such groups may be more difficult than the LLC and HLC groups.

Significant differences in trip distance and travel time can also be seen in two non-transit multimodal groups with combinations of active mode and car. With a higher proportion of car use, the HC+AM group has a longer trip distance on average than the MC+AM group, which is reasonable. However, with nearly half of daily trips conducted by active mode, the MC+AM group has slightly shorter travel distances than the exclusive car user groups of LLC and HLC, but significantly longer trip durations than those two groups; such time-consuming active mode trips should be examined for rationality.

A decreased trip distance and travel time are discovered with the increasing use frequency of buses in the LBLM, MBLM, and HBLM groups. There is no regular tendency that can be seen from rail multimodal users with similar levels of motorized connections. However, increasing trends along with higher usage of motorized connection modes are discovered in rail multimodal user groups with similar rail use frequencies. The LRHM, MRHM, and HRHM groups are found to have the longest average distance, indicating that under similar rail usage frequency, the highly motorized access and egress stages of rail trips may be more responsible for longer door-to-door trip distances than on-rail stages. The average trip time tells another story: the group with moderately motorized connections tends to have the longest travel time (see the HRLM, HRMM, and HRLM group as examples). Evidence from the spatial distribution of such groups is needed for further detail.

5.2. Spatial Distribution Analysis of Mobility Pattern Clusters

On the basis of the statistical analysis of the identified mobility pattern clusters, we have also evaluated the spatial distribution differences among cluster patterns. Figures 3–7 present the Kernel density GIS maps of travelers' residential location distributions from all 20 mobility pattern clusters identified in monomodal, non-transit multimodal, and transit multimodal classes, respectively. Significant disparities can be seen by comparing such maps. Generally speaking, four groups of exclusive car users have the most dispersed distributing pattern than others, with travelers' residential locations covering the majority of areas of Nanjing. Dispersed distributions can also be observed in two non-transit multimodal user groups of HC+AM and MC+AM, with the second largest coverage area. Two exclusive active mode user groups and three bus multimodal user groups present a beaded distribution with multiple spatial clusters located in both Nanjing's main city center and suburban centers. However, nine rail multimodal user groups have shown more concentrated distributions, with most users located alongside the rail networks in Nanjing's main city center.

More detailed spatial differences can be distinguished from the perspective of each group. From Figure 3, the LAM group is mainly located in the main city center, the Xianlin subcenter in the Qixia district, as well as the Jiulonghu and Daxuecheng subcenters in the Jiangning district. The HAM group is mainly located in the main city center, the Xianlin subcenter, and the Qiaobei subcenter in the Jiangbei district. Compared to LAM, the HAM group has an obviously larger concentration area in the main city, with a shorter average trip distance but longer travel time obtained in this group; the southern areas and the northwestern areas of the main city require extra attention for active mode infrastructures. With a higher proportion of commuting trips, the Jiulonghu and Daxuecheng subcenters in the Jiangning district should focus more on commuting continuity using active mode of transport.

With exclusive car users located everywhere, those dispersed distributions may be responsible for the below-average level of residential house prices discovered in the LLC, LHC, HLC, and HHC groups. Spatial clusters are obtained both in the main city center and three sub-city centers of Liuhe, Lishui, and Gaochun for the HLC group, as presented in Figure 4, which may explain the short distance but time-consuming travel experience. In contrast, unique spatial concentrations are discovered in Banqiao for the LLC and LHC groups, indicating that commuting car trips are highly emphasized in such areas. The eastern main city (Chengdong) is discovered as a more concentrated center in the HLC and HHC group than the other two groups, which reveals leisure and maintaining travel demand in such areas. Meanwhile, the LHC and HHC groups require special attention in planning policies concerning possible mode shift behaviors with relatively smaller concentration areas in the main city but significantly longer-distance car trips.

The HC+AM and MC+AM groups have similar spatial distribution (Figure 5), with only slight differences in the Xianlin subcenters emphasized in the MC+AM group, and the Banqiao area is emphasized in the HC+AM group; both areas have poor accessibility to transit networks.

Along with the increasing usage frequency of buses, a smaller trend of concentration areas is discovered (Figure 6) for the LBLM, MBLM, and HBLM groups, which are consistent with a higher trend of transit accessibility shown in Table 5. Significant differences can also be distinguished from nine rail multimodal traveler groups (Figure 7). Along with the increasing rail usage proportion under similar low motorized connection levels, we can see a dispersed trend of travelers' residential spatial center distributions.

With more distant residential centers like Qiaobei, Cheng-dong, and Daxuecheng discovered in MRLM and HRLM groups, evidence of the longer trip distance obtained in such groups has been confirmed. However, using highly motorized modes like automobiles and buses for rail trip connection, the LRHM, MRHM, and HRHM groups have outlined centrally concentrated spatial distribution trends for both residential centers and outside clusters. The statistical results driven from Table 5 support the finding that the HRHM group has relatively shorter door-to-door trip distance than the MRHM and LRHM group, indicating that appropriate distances rather than remote distances from home to rail network can formulate higher usage of rail with motorized connections.

All three groups of LRMM, MRMM, and HRMM have shown concentration effects alongside rail routes, with relatively larger residential areas rather than multiple spatial clusters. With a higher possibility of combining more travel modes and relatively long connections, such complex distribution patterns may explain the most time-consuming trips obtained in these groups. Comparisons between groups with low-frequency or moderately dominant rail usage proportion but different levels of motorized connections have shown a more dispersed distribution in both residential centers and outside clusters.

The LRHM and MRHM groups have captured more remote residential areas in Liuhe sub-cities and southern Yuhuatai and Jiangning districts, where rail networks are the most sparse. On the other hand, the HRLM, HRMM, and HRHM groups see more concentrated distributions in the main city center but more spread-out distributions in suburban areas along with the increasing usage of rail. This reveals that improving integrated rail and bus network services may promote the usage frequency of both rail trips and bus–rail connections.

6. Discussion

Generally, when considering the mode used in every segment of a door-to-door trip, 85.6% of travelers in the Nanjing dataset are multimodal. Compared to the multimodal travelers identified at individual level through the daily travel diary survey (i.e., 53.6% in Germany [5], 60% in the Netherlands [34], 8.7% in Chongqin, China [62]), our results show a higher resolution in capturing multimodal behavior. Especially with huge distinctions between multimodal proportions obtained in the Chinese context, it may cause serious bias in implementing planning or policy when trip-level multimodal mobility is neglected.

Also, compared to other studies where active mode is only considered in one or two clusters [16,17], the role of active mode in an individual's daily mobility pattern is highly emphasized in our research. With nearly every mobility pattern describing different usage levels of active mode, an individual's daily mobility pattern may serve the specific need of a psychological analysis of the active mode [63]. Considering door-to-door trip characteristics at the daily level, three bus multimodal groups and nine metro multimodal groups are identified. Compared to the multimodal mobility pattern identified only at trip level (i.e., Yang et al. [21]), our results help to provide a more comprehensive view of transit multimodal behavior. The disparity details between each group is of great importance for providing needed information in multimodal transportation network planning [64,65] and carbon footprint analysis [66,67] in the era of big data analysis.

Based on such information, more precise and targeted planning policies can be inferred from the statistical and spatial analyses of the 20 identified monomodal and multimodal mobility pattern clusters.

6.1. Policy Implications for Exclusive Active Mode Travelers

Exclusive active mode users are the most sustainable compared to other groups; however, their travel experiences are often less emphasized from planning or policymaking perspectives. According to our results, spatial gathering locations of the HAM group need to pay more attention to complete networks for active mode trips, especially in the Laochengnan area, the southern subcenter of Nanjing Jiangnan's main city. Those areas are the most concentrated places of elderly populations; however, with narrow or missing non-motorized lanes and sidewalks, improvement for the active mode travel experience is in urgent need. With a higher proportion of commuting trips, the concentration centers of the LAM users should examine the active mode networks for connectivity and fast accessibility during peak hours. Fast bike lanes may be needed in some high-flow areas in Xianlin and Jiangning city subcenters. Those implications may help to provide detailed evidence about facility planning in related areas.

6.2. Policy Implications for Exclusive Car Traveler and Non-Transit Multimodal Travelers

With different levels and features of car use, exclusive car users and non-transit multimodal travelers can be considered as targeted groups for promoting mode shift behaviors. In particular, with the irrational use of cars for short-distance trips identified, the concentration areas for travelers in the HLC and HC+AM groups may require a special push-and-pull policy design for reducing car use. For example, the core of the Yuhua District and the southern Hexi District are concentration areas for HLC and HC+AM groups. Especially in concentration areas with already relatively good rail services, such as the core area of the Yuhua District and the southern Hexi District, higher parking prices or a lower supply of parking spaces than other areas is needed. The areas in the Nanjing Parking Planning, however, are treated as regular areas with sufficient parking supply and normal parking prices. With the highest dependency currently on long-distance car travel but which has the worst transit accessibility, new rail route planning or higher rail network density is needed in concentration centers of the LHC group for possible mode shifts.

6.3. Policy Implications for Transit Multimodal Travelers

A transit multimodal traveler is considered to more easily change their patterns over time [35]. With a detailed pattern distinguished, our results may help to better avoid the possible unsustainable transferring behaviors. With a higher elderly population status and higher daily trip frequency, travelers in the three bus multimodal groups of LBLM, MBLM, and HBLM will benefit more from the renovation of bus interiors for aging people in addition to continuing to improve the accessibility of bus networks.

There are also unreasonable behaviors seen in the nine rail multimodal user groups, which may lead to a converse mode shift to car use if precise solutions are not provided. With the highest number of daily trips and longest average door-to-door trip distance, the LRLM users may be more likely to lose interest in rail use and change to other modes. Planning policies for providing mixed land use around rail stations is needed, especially for suburban spatial clusters of LRLM users. With most time-consuming door-to-door trips, the band distribution features alongside rail networks of the LRMM, MRMM, and HRMM groups have indicated that improvements of bus–rail transfers as well as shuttle buses in residential areas may be beneficial for such users. The most significant drop between rail trip proportions in daily trips and commuting trips is discovered in the HRLM group; with a higher use of rails in daily leisure and maintaining trips, travelers in this group may be more concerned about a comfortable on-rail environment than other groups. The lowest transit network accessibility is identified in the MRHM group, indicating that a targeted improvement of transit network density according to the MRHM user spatial distributions would be of great use.

Despite the high-resolution results, we also recognize the limitations of the big data mining method and multi-layer cluster analysis used in this study. First, the short-term velocity and acceleration we used for door-to-door modal split are difficult when distinguishing e-bike users. This means that the e-bike used by an individual is either to be split into automobile mode or into active mode according to specific speed features in our study. However, e-bike users have specific needs for travel, such as more parking space. A specific distinguishment of e-bikes may be beneficial in future works. Second, we used a K-means cluster method with several iterations in each layer to determine center point initialization; however, the time-consuming process may be less-efficient when more indicators are introduced. An improvement of the K-means cluster method may be needed to deal with massive data and indicators. Third, the selection of indicators was based on the knowledge of the existing literature; since the mobility patterns are considered more locally used than comparatively used, a model algorithm between local travel behavior and indicator variables may be beneficial for selecting more distinctive indicators.

7. Conclusions

This work contributes to a more comprehensive understanding of monomodal and multimodal mobility patterns, as well as to more detailed evidence of spatial and behavioral disparities across patterns from an integrated view of the complete door-to-door trip mode chain with the daily mode chain. Our knowledge of monomodal and multimodal mobility has been enhanced in several ways.

First, using combined big datasets, including mobile phone signaling data and public transportation IC card OD data, we have revealed the effectiveness and advantage of big data mining in capturing complete daily travel trajectories, including door-to-door trip mode, as a study basis, which may break the survey design barriers for collecting comprehensive, high-resolution and high-quality behavioral data needed in multimodality research.

Second, the multi-layer cluster analysis model framework we designed for this research may provide new evidence about identifying typical monomodal and multimodal mobility patterns from dealing with large-scale data. The model structure of one classification layer and three cluster analysis layers under specific indicators have enabled us to consider trip characteristics like trip purpose, trip distance, and door-to-door trip mode into daily modal mobility pattern identifications. The model results have shown great superiority in catching multimodal travel behaviors, with 85.6% of total travelers obtained as multimodal in the Nanjing dataset.

Third, the 20 mobility pattern clusters identified under this framework, including six types of monomodal traveler groups, two types of non-transit multimodal traveler groups, and 12 types of transit multimodal traveler groups, have provided more detailed information in establishing traveler profiles as well as distinguishing travel behavior, and may be beneficial in filling the planning–implementation gap from the perspective of planners, policymakers, and travelers.

Last, we have examined possible disparities among all 20 identified clusters through both statistical analysis and spatial distribution analysis with the Kernel density GIS maps of traveler residential location. Significant differences between detailed clusters have been obtained from both analyses. The two exclusive active mode users of LAM and HAM mainly differed in spatial concentration location and community elderly status, leading to different travel experiences. While dispersed spatial distributions of four exclusive car user groups and two non-transit multimodal user groups give common features like relatively poor accessibility to transit networks and lower house price status, a significant difference is still obtained for trip distances and car use dependency. With beaded spatial distribution, the three bus multimodal user groups have shown a decreased tendency of trip distance along with the increase in bus use frequency in mainly leisure or maintaining trips. More evidence can be found from the nine rail multimodal user groups concerning all indicators of spatial, social, and trip characteristics, which help to provide mutually corroborating views for possible policy deductions.

Author Contributions: Conceptualization, J.Z. and L.G.; methodology, J.Z. and F.Z.; software, J.Z. and F.Z.; validation, F.Z., C.H. and X.C. and J.Z.; formal analysis, C.H.; data curation, X.C.; writing—original draft preparation, J.Z.; writing—review and editing, J.Z. and L.G.; visualization, F.Z.; supervision, L.G.; project administration, J.Z.; funding acquisition, J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (52072066) and Jiangsu Province Science Fund for Distinguished Young Scholars (BK20200014).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The data are not publicly available due to confidentiality reasons.

Conflicts of 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.

References

1. Yin, B.; Leurent, F. What Are the Multimodal Patterns of Individual Mobility at the Day Level in the Paris Region? A Two-Stage Data-Driven Approach Based on the 2018 Household Travel Survey. *Transportation* **2022**, *50*, 1497–1526. [CrossRef]
2. Timmer, S.; Merfeld, K.; Henkel, S. Exploring Motivations for Multimodal Commuting: A Hierarchical Means-End Chain Analysis. *Transp. Res. Part A Policy Pract.* **2023**, *176*, 103831. [CrossRef]
3. Timmer, S.; Bösehans, G.; Henkel, S. Behavioural Norms or Personal Gains?—An Empirical Analysis of Commuters' Intention to Switch to Multimodal Mobility Behaviour. *Transp. Res. Part A Policy Pract.* **2023**, *170*, 103620. [CrossRef]
4. Tsouros, I.; Polydoropoulou, A.; Tsirimpa, A. Multimodal Alternatives: How Do Users Perceive the Different Combination of Modes? In Proceedings of the 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), Naples, Italy, 26–28 June 2017; IEEE: Piscataway, NJ, USA, 2017.
5. Nobis, C. Multimodality. *Transp. Res. Rec. J. Transp. Res. Board* **2007**, *2010*, 35–44. [CrossRef]
6. Buehler, R.; Hamre, A. The Multimodal Majority? Driving, Walking, Cycling, and Public Transportation Use among American Adults. *Transportation* **2014**, *42*, 1081–1101. [CrossRef]
7. Deschaintres, E.; Morency, C.; Trépanier, M. Measuring Changes in Multimodal Travel Behavior Resulting from Transport Supply Improvement. *Transp. Res. Rec. J. Transp. Res. Board* **2021**, *2675*, 533–546. [CrossRef]
8. Kent, J.L. Driving to Save Time or Saving Time to Drive? The Enduring Appeal of the Private Car. *Transp. Res. Part A Policy Pract.* **2014**, *65*, 103–115. [CrossRef]
9. Kuhnimhof, T.; Buehler, R.; Wirtz, M.; Kalinowska, D. Travel Trends among Young Adults in Germany: Increasing Multimodality and Declining Car Use for Men. *J. Transp. Geogr.* **2012**, *24*, 443–450. [CrossRef]
10. Streit, T.; Allier, C.-E.; Weiss, C.; Chlond, B.; Vortisch, P. Changes in Variability and Flexibility of Individual Travel in Germany. *Transp. Res. Rec. J. Transp. Res. Board* **2015**, *2496*, 10–19. [CrossRef]
11. Jonuschat, H.; Stephan, K.; Schelewsky, M. Understanding Multimodal and Intermodal Mobility. In *Sustainable Urban Transport*; Emerald Group Publishing Limited: Bingley, UK, 2015; pp. 149–176.
12. Molin, E.; Mokhtarian, P.; Kroesen, M. Multimodal Travel Groups and Attitudes: A Latent Class Cluster Analysis of Dutch Travelers. *Transp. Res. Part A Policy Pract.* **2016**, *83*, 14–29. [CrossRef]
13. Aarts, H.; Verplanken, B.; Van Knippenberg, A. Predicting behavior from actions in the past: Repeated decision making or a matter of habit? *J. Appl. Soc. Psychol.* **1998**, *28*, 1355–1374. [CrossRef]
14. Clifton, K.; Muhs, C.D. Capturing and Representing Multimodal Trips in Travel Surveys. *Transp. Res. Rec. J. Transp. Res. Board* **2012**, *2285*, 74–83. [CrossRef]
15. Klinger, T. Moving from Monomodality to Multimodality? Changes in Mode Choice of New Residents. *Transp. Res. Part A Policy Pract.* **2017**, *104*, 221–237. [CrossRef]
16. Diana, M.; Mokhtarian, P.L. Desire to Change One's Multimodality and Its Relationship to the Use of Different Transport Means. *Transp. Res. Part F Traffic Psychol. Behav.* **2009**, *12*, 107–119. [CrossRef]
17. Ton, D.; Zomer, L.-B.; Schneider, F.; Hoogendoorn-Lanser, S.; Duives, D.; Cats, O.; Hoogendoorn, S. Latent Classes of Daily Mobility Patterns: The Relationship with Attitudes towards Modes. *Transportation* **2019**, *47*, 1843–1866. [CrossRef]
18. Shaheen, S.; Rodier, C. *EasyConnect: Low-Speed Modes Linked to Transit Planning Project*; California PATH Research Report; Partners for Advanced Travel and Highways, University of California: Berkeley, CA, USA, 2008.
19. Maximizing Mobility in Los Angeles—First & Last Mile Strategies. University of California, Berkeley, Transportation Library; Southern California Association of Governments, 2009. Available online: http://www.scag.ca.gov/nonmotorized/pdfs/LA-Maximizing-Mobility-Final-Vol1.pdf (accessed on 20 March 2023).
20. Diana, M.; Pirra, M. A Comparative Assessment of Synthetic Indices to Measure Multimodality Behaviours. *Transp. A Transp. Sci.* **2016**, *12*, 771–793. [CrossRef]
21. Yang, M.; Zhao, J.; Wang, W.; Liu, Z.; Li, Z. Metro Commuters' Satisfaction in Multi-Type Access and Egress Transferring Groups. *Transp. Res. Part D Transp. Environ.* **2015**, *34*, 179–194. [CrossRef]
22. Krygsman, S.; Dijst, M. Multimodal Trips in the Netherlands: Microlevel Individual Attributes and Residential Context. *Transp. Res. Rec. J. Transp. Res. Board* **2001**, *1753*, 11–19. [CrossRef]
23. Sallis, J.F.; Frank, L.D.; Saelens, B.E.; Kraft, M.K. Active Transportation and Physical Activity: Opportunities for Collaboration on Transportation and Public Health Research. *Transp. Res. Part A Policy Pract.* **2004**, *38*, 249–268. [CrossRef]
24. Alessandretti, L.; Natera Orozco, L.G.; Saberi, M.; Szell, M.; Battiston, F. Multimodal Urban Mobility and Multilayer Transport Networks. *Environ. Plan. B Urban Anal. City Sci.* **2022**, *50*, 2038–2070. [CrossRef]
25. Barmpounakis, E.; Geroliminis, N. On the New Era of Urban Traffic Monitoring with Massive Drone Data: The pNEUMA Large-Scale Field Experiment. *Transp. Res. Part C Emerg. Technol.* **2020**, *111*, 50–71. [CrossRef]

26. Hatziioannidu, F.; Polydoropoulou, A. Passenger Demand And Patterns Of Tourists' Mobility In The Aegean Archipelago With Combined Use Of Big Datasets From Mobile Phones And Statistical Data From Ports And Airports. *Transp. Res. Procedia* **2017**, *25*, 2309–2329. [CrossRef]
27. Wu, L.Y.; Hasan, S.; Chung, Y.; Kang, J.E. Understanding the Heterogeneity of Human Mobility Patterns: User Characteristics and Modal Preferences. *Sustainability* **2021**, *13*, 13921. [CrossRef]
28. Pani, A.; Sahu, P.; Mishra, S. Gender disparities in multimodal travel Attitudes, Behavior, and satisfaction. *Transp. Res. Part D Transp. Environ.* **2023**, *123*, 103917. [CrossRef]
29. Barker, J.; Ademolu, E.; Bowlby, S.; Musson, S. Youth transitions: Mobility and the travel intentions of 12–20 year olds, Reading, UK. *Child. Geogr.* **2019**, *17*, 442–453. [CrossRef]
30. Lu, Y.; Prato, C.G.; Sipe, N.; Kimpton, A.; Corcoran, J. The role of household modality style in first and last mile travel mode choice. *Transp. Res. Part A Policy Pract.* **2022**, *158*, 95–109. [CrossRef]
31. Ton, D.; Shelat, S.; Nijënstein, S.; Rijsman, L.; van Oort, N.; Hoogendoorn, S. Understanding the Role of Cycling to Urban Transit Stations through a Simultaneous Access Mode and Station Choice Model. *Transp. Res. Rec. J. Transp. Res. Board* **2020**, *2674*, 823–835. [CrossRef]
32. Djurhuus, S.; Hansen, H.; Aadahl, M.; Glümer, C. The Association between Access to Public Transportation and Self-Reported Active Commuting. *Int. J. Environ. Res. Public Health* **2014**, *11*, 12632–12651. [CrossRef]
33. Chan, K.; Farber, S. Factors Underlying the Connections between Active Transportation and Public Transit at Commuter Rail in the Greater Toronto and Hamilton Area. *Transportation* **2019**, *47*, 2157–2178. [CrossRef]
34. Schneider, F.; Ton, D.; Zomer, L.-B.; Daamen, W.; Duives, D.; Hoogendoorn-Lanser, S.; Hoogendoorn, S. Trip Chain Complexity: A Comparison among Latent Classes of Daily Mobility Patterns. *Transportation* **2020**, *48*, 953–975. [CrossRef]
35. Kroesen, M. Modeling the Behavioral Determinants of Travel Behavior: An Application of Latent Transition Analysis. *Transp. Res. Part A Policy Pract.* **2014**, *65*, 56–67. [CrossRef]
36. Kroesen, M.; Cranenburgh, S.V. Revealing Transition Patterns between Mono- and Multimodal Travel Patterns over Time: A Mover-Stayer Model. *Eur. J. Transp. Infrastruct. Res.* **2016**, *16*, 754–771. [CrossRef]
37. An, Z.; Heinen, E.; Watling, D. The level and determinants of multimodal travel behavior: Does trip purpose make a difference? *Int. J. Sustain. Transp.* **2023**, *17*, 103–117. [CrossRef]
38. Buehler, R.; Hamre, A. An examination of recent trends in multimodal travel behavior among American motorists. *Int. J. Sustain. Transp.* **2016**, *10*, 354–364. [CrossRef]
39. Mao, Z.; Ettema, D.; Dijst, M. Commuting trip satisfaction in Beijing: Exploring the influence of multimodal behavior and modal flexibility. *Transp. Res. A Policy Pract.* **2016**, *94*, 592–603. [CrossRef]
40. Jia, N.; Li, L.; Ling, S.; Ma, S.; Yao, W. Influence of Attitudinal and Low-Carbon Factors on Behavioral Intention of Commuting Mode Choice—A Cross-City Study in China. *Transp. Res. Part A Policy Pract.* **2018**, *111*, 108–118. [CrossRef]
41. Wells, E.M.; Small, M.; Spurlock, C.A.; Wong-Parodi, G. Factors associated with emerging multimodal transportation behavior in the San Francisco Bay Area. *Environ. Res. Infrastruct. Sustain.* **2021**, *1*, 031004. [CrossRef]
42. De Witte, A.; Hollevoet, J.; Dobruszkes, F.; Hubert, M.; Macharis, C. Linking Modal Choice to Motility: A Comprehensive Review. *Transp. Res. Part A Policy Pract.* **2013**, *49*, 329–341. [CrossRef]
43. Magdolen, M.; von Behren, S.; Burger, L.; Chlond, B. Mobility styles and car ownership—Potentials for a sustainable urban transport. *Sustainability* **2021**, *13*, 2968. [CrossRef]
44. Vermunt, J.K.; Magidson, J. Latent class cluster analysis. In *Applied Latent Class Analysis*; Cambridge University Press: Cambridge, UK, 2002; pp. 89–106.
45. De Haas, M.; Wijgergangs, K.; Hoogendoorn-Lanser, S. Identifying diferent types of observed immobility within longitudinal travel surveys. In Proceedings of the International Conference on Transport Survey Methods, Estérel, QC, Canada, 24–29 September 2017.
46. Krueger, R.; Vij, A. Normative Beliefs and Modality Styles: A Latent Class and Latent Variable Model of Travel Behaviour. *SSRN Electron. J.* **2016**, *45*, 789–825. [CrossRef]
47. Yu, D.H.; Dong, S.H.; Yao, S. Improvement of K-Means Algorithm and Its Application in Air Passenger Grouping. *Comput. Intell. Neurosci.* **2022**, *2022*, 3958423. [CrossRef] [PubMed]
48. Lou, N. Analysis of the Intelligent Tourism Route Planning Scheme Based on the Cluster Analysis Algorithm. *Comput. Intell. Neurosci.* **2022**, *2022*, 3310676. [CrossRef]
49. Wei, H.; Lei, G.; Qin, L.; Tian, L. Cluster Analysis of Trip Purpose Based on Residents' Travel Characteristic. In Proceedings of the 2022 IEEE 7th International Conference on Intelligent Transportation Engineering, ICITE, Beijing, China, 11–13 November 2022; pp. 574–579. [CrossRef]
50. Ding, J.; Liu, H.; Yang, L.T.; Yao, T.; Zuo, W. Multiuser Multivariate Multiorder Markov-Based Multimodal User Mobility Pattern Prediction. *IEEE Internet Things J.* **2020**, *7*, 4519–4531. [CrossRef]
51. Pan, Y.; Darzi, A.; Yang, M.; Sun, Q.; Kabiri, A.; Zhao, G.; Xiong, C.; Zhang, L. National-Level Multimodal Origin–Destination Estimation Based on Passively Collected Location Data and Machine Learning Methods. *Transp. Res. Rec. J. Transp. Res. Board* **2023**. [CrossRef]
52. Marchal, P.; Madre, J.-L.; Yuan, S. Postprocessing Procedures for Person-Based Global Positioning System Data Collected in the French National Travel Survey 2007–2008. *Transp. Res. Rec. J. Transp. Res. Board* **2011**, *2246*, 47–54. [CrossRef]

53. Hood, J.; Sall, E.; Charlton, B. A GPS-Based Bicycle Route Choice Model for San Francisco, California. *Transp. Lett.* **2011**, *3*, 63–75. [CrossRef]
54. Chen, X.; Xu, X.; Yang, C. Trip Mode Inference from Mobile Phone Signaling Data Using Logarithm Gaussian Mixture Model. *J. Transp. Land Use* **2020**, *13*, 429–445. [CrossRef]
55. Liu, H. Research on Analysis Method and Model of Urban Rail Travel Behavior Based on Multi-Source Data. Ph.D. Thesis, Chongqing Jiaotong University, Chongqing, China, 2022.
56. Wu, Z.; Xie, J.; Wang, Y.; Nie, Y. (Marco) Map Matching Based on Multi-Layer Road Index. *Transp. Res. Part C Emerg. Technol.* **2020**, *118*, 102651. [CrossRef]
57. Ding, F.; Li, X.; Lyu, Y.; Wang, Y.; Jiang, L.; Ji, H.; Tong, E.; Zhang, D. C-UTBDS: Urban Traffic Travel Characteristic Mining Based on Cellular-Network Big Data. *China J. Highw. Transp.* **2022**, *36*, 165–182. [CrossRef]
58. Song, S. Research on Regional Travel Destination Choice Model Based on Cellular Signaling Data. Master's Thesis, Xi'an Jiaotong University, Xi'an, China, 2022.
59. Li, Y. User Travel Behavior Research Based on Mobile Phone Signaling Data. Master's Thesis, Chongqing University of Posts and Telecommunications, Chongqing, China, 2017.
60. Wang, Y. Research on Resident Travel Semantic Method Integrating Multi-Source Big Data and Spatio-Temporal Weight Model. Master's Thesis, Jilin Jianzhu University, Jilin, China, 2021.
61. Sulikova, S.; Brand, C. Investigating What Makes People Walk or Cycle Using a Socio-Ecological Approach in Seven European Cities. *Transp. Res. Part F Traffic Psychol. Behav.* **2021**, *83*, 351–381. [CrossRef]
62. Chen, S. Modeling of the All Day Multi Travel Mode Choice Considering the Heterogeneity of Choice Set. Master's Thesis, Southeast University, Nanjing, China, 2022.
63. An, Z.; Heinen, E.; Watling, D. Multimodal Travel Behaviour, Attitudes, and Cognitive Dissonance. *Transp. Res. Part F Traffic Psychol. Behav.* **2022**, *91*, 260–273. [CrossRef]
64. Wang, S.; Fu, S. Path Design and Planning and Investment and Construction Mode of Multimodal Transport Network Based on Big Data Analysis. *Discret. Dyn. Nat. Soc.* **2022**, *2022*, 9185372. [CrossRef]
65. Franco, P.; Johnston, R.; McCormick, E. Demand Responsive Transport: Generation of Activity Patterns from Mobile Phone Network Data to Support the Operation of New Mobility Services. *Transp. Res. Part A Policy Pract.* **2020**, *131*, 244–266. [CrossRef]
66. Bhandari, K.; Advani, M.; Parida, P.; Gangopadhyay, S. Consideration of Access and Egress Trips in Carbon Footprint Estimation of Public Transport Trips: Case Study of Delhi. *J. Clean. Prod.* **2014**, *85*, 234–240. [CrossRef]
67. Heinen, E.; Mattioli, G. Multimodality and CO2 Emissions: A Relationship Moderated by Distance. *Transp. Res. Part D Transp. Environ.* **2019**, *75*, 179–196. [CrossRef]