Urban Climate 53 (2024) 101794

Contents lists available at ScienceDirect

Urban Climate

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

# Congestion and Pollutant Emission Analysis of Urban Road Networks Based on Floating Vehicle Data

Wen-Long Shang [a,*], Xuewang Song [b], Yishui Chen [b], Xin Yang [c], Liyun Liang [d], Muhammet Deveci [e,f,g], Mengqiu Cao [h], Qiannian Xiang [b], Qing Yu [i]

[a] *Beijing Key Laboratory of Traffic Engineering, College of Metropolitan Transportation, Beijing University of Technology, Imperial College London, London, United Kingdom*
[b] *College of Metropolitan Transportation, Beijing University of Technology, Beijing, China*
[c] *State Key Laboratory of Advanced Rail Autonomous Operation, Beijing Jiaotong University, China*
[d] *Department of Earth System Science, Ministry of Education Key Laboratory for Earth System Modeling, Institute for Global Change Studies, Tsinghua University, Beijing, China*
[e] *Department of Industrial Engineering, Turkish Naval Academy, National Defence University, 34940 Tuzla, Istanbul, Turkey*
[f] *The Royal School of Mines, Imperial College London, SW7 2AZ London, United Kingdom*
[g] *Department of Electrical and Computer Engineering, Lebanese American University, Byblos, Lebanon*
[h] *School of Architecture and Cities, University of Westminster, London NW1 5LS, United Kingdom*
[i] *Research Institute of Trustworthy Autonomous Systems, Southern University of Science and Technology Shenzhen, China*

ARTICLE INFO

*Keywords:*
Climate change
Floating vehicle data
Map-matching
Emission factor
Urban congestion

ABSTRACT

Global warming caused by greenhouse gas (GHG) is receiving increasingly attention from all over the world, and urban transportation is a significant source of greenhouse gas and pollutant emission. However, the research on traffic state of urban road networks (URNs) based on sparse floating vehicle data (FVD) is insufficient. Therefore, we mainly utilize big data techniques to explore the congestion and pollutant emission of URN with FVD. Firstly, the location of vehicles is identified and matched with the URN. We then grid the FVD and city maps to more accurately identify areas of congestion and emission in later section. Following this, we use the congestion index and K-means clustering algorithm to evaluate the traffic state over time, pollutant emission is calculated based on emission calculation standards and carbon emission is estimated by using the fuel consumption-speed model. The results indicate that congestion and emission are very severe during peak hours (e.g., 8:00 a.m.), particularly in some transportation hub areas, such as high-speed rail stations. During off-peak hours (e.g., 11:00 p.m.), congestion and emission are relatively lower. The negative correlation between congestion index and emission is also revealed. This study provides some practical approaches to more accurately estimate the overall urban traffic state by using sparse traffic data, and may offer support to urban traffic managers in managing traffic congestion and pollutant emissions.

* Corresponding author.
*E-mail addresses:* shangwl_imperial@bjut.edu.cn (W.-L. Shang), songxw@emails.bjut.edu.cn (X. Song), Chenys615@emails.bjut.edu.cn (Y. Chen), 11111047@bjtu.edu.cn (X. Yang), Chelsealliang@outlook.com (L. Liang), muhammetdeveci@gmail.com (M. Deveci), m.cao@westminster.ac.uk (M. Cao), qiannian@emails.bjut.edu.cn (Q. Xiang), yuq@sustech.edu.cn (Q. Yu).

https://doi.org/10.1016/j.uclim.2023.101794
Received 9 August 2023; Received in revised form 12 November 2023; Accepted 19 December 2023
Available online 5 January 2024

## 1. Introduction

Global warming is mainly caused by greenhouse gas, which leads to an increase in temperature(Houghton, 2005). Urban areas are responsible for a significant part of the total global emissions of greenhouse gases (Gurney et al., 2022; Xin, 2023), and it becomes important to promote sustainable development to achieve carbon neutrality (Tozer et al., 2022; Hincks et al., 2023; Gokasar and Karaman, 2023). Recognizing the increasingly severe problem of climate change, many countries around the world have committed to promoting the establishment of a low-carbon economy through initiatives such as the *Paris Agreement [WWW Document*] (2016). Gasoline-powered vehicles are a major contributor to greenhouse gas emission yet as economic growth has accelerated the urbanization process in China and raised people's living standards the use of such vehicles has increased significantly. This has in turn increased the level of air pollution.

As of the end of June 2023, the motor vehicle ownership in China has reached 426 million, representing a significant increase of 70.4% compared to the number recorded in 2013 (250 million vehicles). Of the total motor vehicle ownership, cars account for 328 million, reflecting a remarkable growth of 139.4% compared to 2013 (137 million vehicles) (Fanchao, 2023). This substantial surge in motor vehicle ownership reflects a rapid expansion in China's automobile market over the past decade. However, this tremendous growth in car ownership has raised concerns about environmental sustainability and public health. The surge in car numbers has led to a corresponding increase in traffic congestion, air pollution, and greenhouse gas emission, exacerbating the challenges of climate change and urban air quality. According to the China Mobile Source Environmental Management Annual Report, cars have become a significant source of air pollution in China (Ministry of Ecology and Environment of the People's Republic of China, 2022). Among all the pollutants emitted by motor vehicles, cars account for over 90% of the emission of carbon monoxide (CO), hydrocarbons (HC), nitrogen oxides ($NO_x$), and particulate matter (PM) (Ministry of Ecology and Environment of the People's Republic of China, 2022). At the same time, carbon dioxide ($CO_2$) emission from fuel vehicles cannot be ignored. The significant contribution of these pollutants from car emission highlights their substantial impact on air pollution. As such, addressing and mitigating the emission from cars are crucial steps in tackling the air quality challenges we face.

Reducing vehicle-related pollution requires accurate calculation of vehicle pollution emission, identification of congestion areas and efficient route planning (Shindell et al., 2011; Çolak et al., 2016; Grote et al., 2016). The rapid development of information technology in the transportation sector has transformed the field from an era of data scarcity to one of big and rich data. Thus, the research on traffic congestion and pollutant emission has been able to start from micro data at the level of a single vehicle to develop a traffic congestion index evaluation model that is integrated with the actual traffic conditions of the city based on massive traffic data. Visualization technology has become an important tool to present the status of urban traffic operations in real time, allowing the development of models to evaluate traffic congestion and hence better inform sustainable urban development. Dynamic traffic data, such as data detected by detectors (e.g., video, sensor coils) and Floating Vehicle Data (FVD), are widely used for studying congestion and pollutant emissions in urban areas (Rahmani and Koutsopoulos, 2013; Rahmani et al., 2015). Various models, such as the COPERT model, the MOVES model, etc., have also been adopted to conduct related analyze s(Ekström et al., 2004; Perugu, 2019).

Based on this background, this study investigates urban congestion and pollutant emission in Shenzhen, processing floating vehicle data through the use of big data techniques. The analysis of pollutant emission is carried out by determining the pollutant emission factors. To begin with, this paper provides a review of relevant literature, followed by a summary of the current status of research on urban traffic congestion and pollutant emission. This serves as a foundation for the ideas and methods presented in this paper. Subsequently, floating vehicle data obtained from Shenzhen city is processed and analyzed in terms of its characteristics. The city and floating vehicle data are then gridded and matched to analyze the congestion index using a K-means clustering method, thereby generating results for comparison. Additionally, emission factors for each vehicle pollutant at different speeds are computed to analyze pollution emission in Shenzhen city at any given time. Carbon dioxide emission is calculated by the fuel consumption-speed model. Lastly, we analyze the characteristics of congestion during different periods in Shenzhen, explore congested areas and road sections, and assess the effect of urban congestion on pollutant emission. Furthermore, changes in congestion areas are examined to offer guidance on daily travel and energy-saving and emission-reduction measures. Overall, this study contributes to the existing literature by providing an in-depth analysis of urban congestion and pollutant emission in Shenzhen city, utilizing big data techniques. The findings of this study provide practical implications and suggestions for policymakers and urban planners to better manage traffic congestion and reduce air pollution levels in urban cities.

The remainder of this paper is organized as follows: Section 2 provides a literature review related to our research. Section 3 describes the data sources and methodology employed in this study. Section 4 presents a case study. Finally, Section 5 summarizes our research findings and discusses the limitations.

## 2. Literature review

FVD represents a valuable dataset employed for transportation research and management purposes. It involves the continuous acquisition of real-time vehicle information, including positions, velocities, travel directions, and other pertinent data, facilitated by GPS or other sensor technologies. This data is subsequently transmitted to a central collection platform. Given the dynamic nature of vehicle positions, this dataset is referred to as “floating” vehicle data. FVD holds immense potential for diverse transportation applications, encompassing traffic flow analysis, travel behaviour investigations, and driver behaviour evaluations, among others (Shang et al., 2023; Kerner et al., 2005; Rahmani et al., 2015; Nigro et al., 2022).

This section presents a literature review of studies that have used different data (e.g., FVD), and models to assess urban congestion and pollutant emission.

### 2.1. Studies to assess urban traffic congestion

Urban traffic congestion analysis is an important research area in the field of transportation. Effective research on urban traffic congestion holds great value in promoting transportation planning and management for both commuters and governments. This section aims to summarize and analyze research, models, and methods related to urban traffic congestion across four key areas: urban traffic management, congestion prediction, data analysis, and the social and economic impacts of congestion. Through the analysis of these diverse models and methods, this review seeks to identify effective approaches for studying traffic congestion in urban environments.

Urban traffic management cannot be separated from congestion management, and researchers have conducted extensive studies on traffic management to improve the efficiency of the transportation system significantly (Geroliminis and Daganzo, 2008). With the widespread application of big data technology and the development of location-based services (LBS), various vehicle data, such as GPS data, are being widely collected. This trend has led to the advancement of intelligent transportation systems that utilize advanced model techniques, data sources, and optimization algorithms to improve accuracy and efficiency in traffic management (Janecek et al., 2015; Dakic and Menendez, 2018). Overall, the increasing availability of Floating Vehicle Data (FVD) and Automatic Vehicle Location (AVL) data, coupled with deep learning techniques and optimization algorithms, has facilitated extensive research on traffic congestion (Sun et al., 2021; Zhao et al., 2021). For example, Jiang et al. (2021) proposed a multitemporal traffic flow model that takes into account various factors such as weather conditions, holidays, and rush hours. Their method successfully predicted the traffic flow and air pollution level in urban areas and effectively improved the traffic management capability. Some other scholars have used deep convolutional neural networks (DCNN) to analyze sparse data. Rempe et al. (2022) proposed this method, which improves the accuracy of traffic speed estimation and outperforms traditional linear and polynomial regression models. Following this, Chen and Zhang (2022) constructed a traffic flow prediction model based on the Deep Belief Network (DBN) algorithm, which effectively suppresses the spread of congestion in smart cities. Similarly, Zhao et al. (2022) developed a centrally scheduled parking system for connected vehicles and autonomous vehicles (CAVs) using Macro Fundamental Diagrams (MFDs) and Mixed Integer Programming to optimize the parking resource utilization.

Traffic congestion management is a difficult problem in urban traffic management, so predicting the formation and dissipation of traffic congestion becomes very important. To address these questions, researchers have proposed a variety of methods, including establishing a nonlinear regression model to predict the level of traffic congestion under carbon emission constraints (Yang et al., 2019), using elasticity indicators to study the resilience of actual traffic congestion and identify congested clusters (Zhang et al., 2019), predicting congestion based on the travel time of public transportation (Huang et al., 2019), and exploring the multiple steady states in urban transportation networks (Zeng et al., 2020). In addition, researchers have investigated the correlation between frequent congestion points and speed fluctuations and state transition times (Wang et al., 2020) and examined the acceptability of congestion pricing schemes in urban areas (Shatanawi et al., 2020). Following this, Hammami (2020) investigated the impact of urban freight transportation on traffic congestion, while Olusanya et al. (2020) proposed an IT-based solution to improve traffic flow in congested cities. After that, Tang et al. (2023) combined deep neural networks with a subset selection method for predicting spatio-temporal data of urban traffic to improve the prediction accuracy.

Data analysis is a key part of the study of traffic congestion and management, and researchers have used different data and techniques for estimation and analysis. Some researchers have proposed various methods for estimating vehicle speed and travel time based on sparse or missing floating vehicle data. To estimate vehicle travel times, Rahmani et al. (2015) proposed a consistent path inference method for sparse floating vehicle data. After that, Zhang et al. (2017) developed a data-driven method to predict travel times on urban expressways and verified the effectiveness and stability of the method under different traffic conditions. Subsequently, Rempe et al. (2017)and Dakic and Menendez (2018) investigated effective methods for estimating vehicle speeds in the spatio-temporal domain. In addition, Yan et al. (2020) analyzed the congestion characteristics of different time periods and road classifications using real-time traffic data from the Gaode LBS platform. Qian et al. (2020) investigated the impact of multinational corporations on urban traffic congestion by using large-scale trajectory data from transportation network companies. While Xu and Huang (2020) applied spectral clustering to cab GPS trajectories to detect traffic congestion. Meanwhile, Qin et al. (2021) introduced a traffic flow grid model to study the impact of cab GPS trajectories on traffic congestion. Othman et al. (2022) introduced a new method for traffic flow estimation based on floating vehicle data and road topography. In addition, Zeng et al. (2022) designed an effective framework for exploring the spatio-temporal characteristics of traffic congestion based on large-scale cab trajectory data in Shenzhen, China. They effectively analyzed the spatio-temporal characteristics of traffic congestion using a complex network approach, which lays the foundation for developing traffic management and control strategies.

Traffic congestion can also have social and economic impacts. For example, some researchers have explored the impact of smart city construction on traffic congestion and the quality of public transportation (Guo et al., 2020). Beojone and Geroliminis (2021) examined the impact of bike-sharing on urban mobility and found that the presence of bike-sharing increased traffic congestion and total travel time. In addition, Wei et al. (2022) explored the socio-economic costs of traffic congestion in China's urbanization process using big data analysis, revealing differences in the spatial and temporal patterns of urban traffic congestion in different periods. Meanwhile, Bendib (2020) investigated the impact of spatial agglomeration on social equity and urban congestion and proposed an effective strategy to study urban congestion using GIS. In terms of sustainable transportation, Tan et al. (2023) explored the design of a public transportation service network and provided methods for the sustainable development of urban transportation. Li et al. (2023) and Moslem et al. (2023) made efforts in the area of regional sustainable transportation assessment and the quality of the supply of public transportation. On the other hand, Gokasar and Karaman (2023) explored the relationship between personnel services and public transportation. They used the Geographic Information System (GIS) in their study and achieved more effective results in terms of improving peak hour congestion, among others.

This section provides an overview of research on congestion in the field of urban transportation from four perspectives, including the models and methods used. In urban transportation, it is important to identify and solve traffic congestion problems. However, there is not enough research on areas such as congestion studies and data analysis. Although some methods for identifying and predicting traffic congestion have been developed, further research is needed to improve the accuracy and reliability. Therefore, this paper builds upon previous research by using big data methods to analyze and model floating vehicle data, combining the congestion index with the K-means clustering algorithm to study the spatio-temporal characteristics of traffic congestion. This provides a better understanding of the characteristics and trends of traffic congestion, allowing for the development of effective traffic management and control strategies.

### 2.2. Studies to estimate pollutant emission of urban transportation

In recent years, the rapid growth of vehicles in the transportation sector has posed significant challenges to the environment and human health. Emission of pollutants such as carbon monoxide (CO), nitrogen oxides (NOx), and greenhouse gases have garnered widespread attention. To address this issue, scholars in the field of transportation engineering have increasingly employed various models and methods to delve into the emission of pollutants caused by transportation. This section provides an overview of these studies.

Researchers have utilized a variety of models and methods to gain a better understanding of and quantify emission resulting from transportation. Among these, the MOBILE model has been widely used to calculate emission factors and generate emission inventories based on factors such as vehicle speed, mileage, and climate (Fu et al., 2000; Mukherjee and Viswanathan, 2001; Vallamsundar and Lin, 2011). Additionally, the Calculation of Road Emission from Transport (COPERT) model has been frequently applied to study automobile tailpipe emission. For instance, Zhu et al. (2011) employed the COPERT IV model in conjunction with GPS data to determine pollution emission factors and exhaust characteristics of road vehicles in the city of Guangzhou. Researchers have found that the COPERT model offers more accuracy in predicting emission from motorized vehicles in China as compared to the MOBILE model, likely due to China's emission standards being more in line with European standards, while the MOBILE model's parameters are derived from data collected in the United States (Cheng et al., 2011). Xue et al. (2013) then modelled the relationship between pollution emission and speed based on traffic flow on the road.

In addition to these models, researchers have considered the characteristics of various vehicle types, such as the mechanical model proposed by Ehsani et al. (2016). This model takes into account different vehicle types and the impact of toll systems on $CO_2$ emission. Furthermore, the combination of floating vehicle data with the Motor Vehicle Emission Simulator (MOVES) model has analyzed the effects of various traffic conditions on pollutant emission from light-duty vehicles (Chen et al., 2016). However, the complexity of model parameters and uncertainty in default data have led to an increase in the uncertainty of the MOVES model. Some researchers have also localized parameters through the application of the MOVES model (Hao et al., 2017; Li et al., 2021). Du et al. (2017), on the other hand, used a neural network model to predict the fuel consumption patterns of different vehicles under various influencing factors based on floating vehicle data. Moreover, the International Vehicle Emission (IVE) model has been employed to establish road emission inventories (Zhou et al., 2019). Scholars have used different models for various studies to comprehend the spatiotemporal changes in urban public transportation pollution emission following the COVID-19 pandemic, analyze the impact of information and communication technology measures on emission, and investigate the environmental benefits of shared bicycles (Sui et al., 2020; Shang et al., 2021).

Researchers have also explored the effects of new technologies and policies on carbon emission reduction. They have compared fuel consumption and emission of taxi services and ride-sharing service DiDi using GPS and order data, finding that DiDi performs better in terms of energy efficiency and emission reduction (Sui et al., 2019). Studies have also compared the impact of different vehicle technologies on reducing carbon emission from taxis, suggesting that fuel cell electric vehicles (FCEVs) using renewable energy sources have greater advantages in terms of carbon emission reduction cost (Mingolla and Lu, 2021; Bai et al., 2022). Additionally, Ercan et al. (2022) investigated carbon emission from autonomous electric vehicles using polynomial logit models and system dynamics model, and the results showed that autonomous electric vehicles can significantly reduce greenhouse gas emission in the transportation sector. Yao et al. (2022b) explored carbon emission from electric vehicles through the vehicle-to-grid (V2G) model, revealing that due to battery disposal and the predominant use of coal for electricity generation, the advantages of two-way V2G over one-way V2G are limited. While the application of models and methods plays a pivotal role in studying emission caused by transportation, researchers have pointed out some challenges, including the availability of parameters and uncertainty in model default data. Future directions in research should focus on improving and localizing existing emission models to adapt to the environmental and transportation conditions of specific regions (Liu et al., 2022). This will help enhance the predictive accuracy and reliability of models.

In conclusion, the emission of pollutant in the transportation sector poses significant challenges to the environment and human health. Through continuous improvement and the application of various models and methods, researchers can gain a better understanding of and quantify the pollution emission caused by transportation. They can explore the effects of new technologies and policies on carbon emission reduction, offering critical support for the sustainable development of the transportation sector and carbon neutrality goals (Shang et al., 2023; Shang and Lv, 2023). However, model uncertainty and the complexity of parameters remain obstacles to overcome. Therefore, this study, based on the geographical features of Shenzhen, China, utilized methods compliant with Chinese pollutant emission calculation standards, along with a fuel consumption model to calculate $CO_2$ emission. We analyzed the spatio-temporal characteristics of pollutant emission and $CO_2$ emission, and examined their relationship with traffic congestion. Our research can provide a reference for governments to take feasible measures to reduce pollutant emission and carbon emission.

## 3. Methodology

In this section, we propose a framework for determining the congestion level of urban road networks and estimating vehicle emission in such networks based on taxi GPS data. Specifically, we first pre-processing the floating vehicle data and grid the study area (Section 3.1). We then analyze the congestion level of urban roads using a congestion index and K-means clustering algorithm (Section 3.2). Following this, we analyze taxi pollutant emission using domestic vehicle emission calculation standards (Section 3.3). Finally, carbon dioxide emission is calculated using the fuel consumption-speed model (Section 3.4).

The pre-processing of floating vehicle data and gridding of the study area are necessary to extract meaningful insights from the raw data. These processes aim to reduce noise and improve the accuracy of subsequent analysis. Furthermore, the use of a congestion index and the K-means clustering algorithm enables us to effectively identify and analyze congested roads within the urban network. This approach facilitates the identification of high-priority areas for congestion mitigation measures and helps to optimize traffic flow management. Finally, we analyze the pollutant emission from taxis using current vehicle pollutant emission standards, thus providing insight into the environmental impact of the urban road network. This information can inform policy decisions regarding sustainable transportation and urban planning. Overall, our framework presents a comprehensive and nuanced approach to understanding and addressing the challenges inherent in managing urban road networks. The detailed flowchart is shown in Fig. 1.

### 3.1. Data processing

The data utilized for this study is obtained from the travel records of typical taxis operating within Shenzhen city on regular workdays (Zhang et al., 2015). The equipment installed in the taxis collects data at intervals ranging from 1 to 60 s per vehicle, over 24 h from 00:00 to 23:59. To protect the privacy of the passengers, the data hide information related to the specific date and the real license plate number of the vehicle. The dataset includes key attributes, such as vehicle identification (ID), timestamp, geographical coordinates (longitude and latitude), and passenger occupancy status. The specific data is presented in Table 1. In the data table, 'ID' represents a unique identifier assigned to each taxi, 'time' indicates the timestamp of data collection spanning from 00:00 to 23:59, 'lon' corresponds to the longitude coordinate of each taxi during data collection, 'lat' pertains to the latitude coordinate of each taxi at the time of data acquisition, 'passenger' signifies the presence or absence of passengers in the taxi during data collection, and 'speed' denotes the speed of the vehicle at the time of data recording. Importantly, the binary value "1" indicates passenger presence, while "0" signifies passenger absence. We sort the collected data by ID and time to get the data of each vehicle in chronological order and use this data as the basis for data cleaning.

The floating vehicle data have potential errors, and cleaning up these erroneous data is a step that must be taken before the data can be analyzed. Therefore, we remove duplicate and abnormal data, such as data with sudden changes in passenger status. We define the data with a sudden change in passenger status as abnormal data. There are two situations in which a sudden change in passenger status

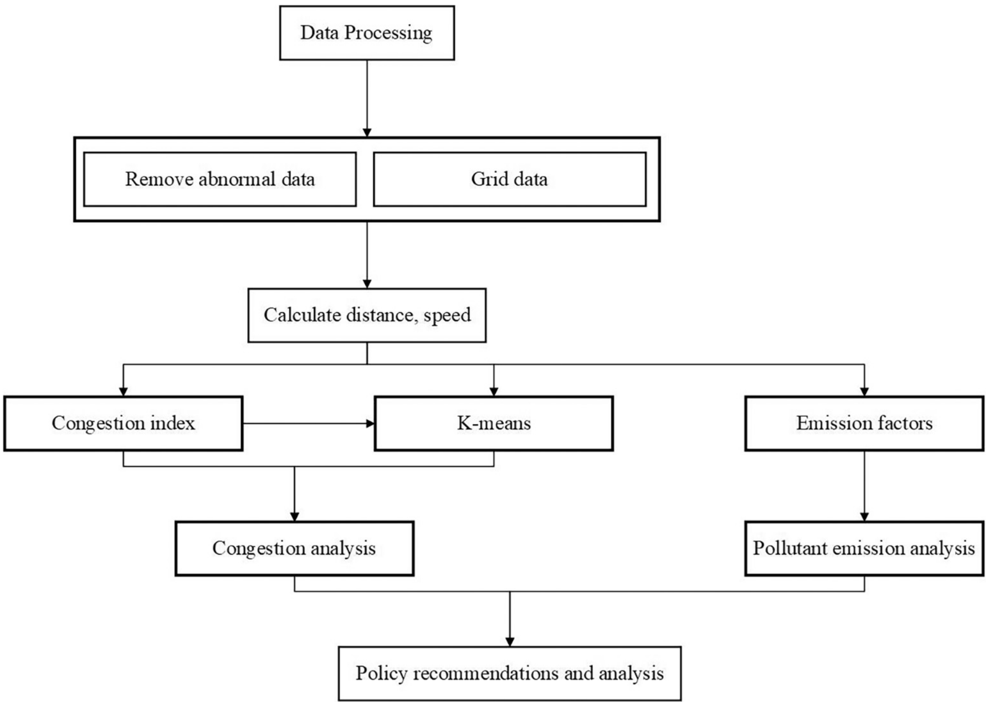



**Fig. 1.** The methodology flowchart.

**Table 1**
Floating vehicle data of Shenzhen City.

| | ID | time | lon | lat | passenger | speed |
|---|---|---|---|---|---|---|
| 0 | 22,223 | 0:00:00 | 114.145714 | 22.555317 | 1 | 80 |
| 1 | 22,223 | 0:00:06 | 114.14695 | 22.555468 | 1 | 80 |
| 2 | 22,223 | 0:00:14 | 114.148537 | 22.555866 | 1 | 80 |
| 3 | 22,223 | 0:00:16 | 114.148949 | 22.555933 | 1 | 79 |
| 4 | 22,223 | 0:00:22 | 114.1502 | 22.556116 | 1 | 84 |
| 5 | 22,223 | 0:00:30 | 114.151817 | 22.556601 | 1 | 81 |
| 6 | 22,223 | 0:00:32 | 114.152168 | 22.556816 | 1 | 81 |
| 7 | 22,223 | 0:00:45 | 114.154366 | 22.558649 | 1 | 81 |
| 8 | 22,223 | 0:01:36 | 114.16243 | 22.562517 | 1 | 50 |
| 9 | 22,223 | 0:01:44 | 114.163551 | 22.562468 | 1 | 36 |
| 10 | 22,223 | 0:01:59 | 114.16378 | 22.562349 | 1 | 0 |
| 11 | 22,223 | 0:02:16 | 114.163803 | 22.562349 | 0 | 0 |

occurs: (1) when there is a sudden transition from unoccupied to occupied in a continuous record. This situation indicates that a passenger is getting on or off the vehicle while the vehicle is unoccupied. This subset of anomalous data will be deleted. (2) When there is a sudden transition from occupied to unoccupied in a continuous record. This situation indicates that passengers are getting off and out of the vehicle quickly (Yu and Li, 2022). Therefore, the data in these cases was also removed. Since both the identification of duplicate and anomalous data rely on vehicle IDs and timestamps, we first aggregate data from the same vehicle and then classify it based on timestamps. Data cleaning operations are then performed as described above.

After data cleaning, we conducted a statistical analysis of the data and found that 95% of the data sampling intervals were below 60 s. Therefore, to reduce the error caused by excessively long data sampling intervals, we remove the data points with intervals exceeding 60 s. Subsequently, we construct the travel trajectories of each vehicle, each trajectory consisting of an origin point (O) and a destination point (D). Given the relatively short intervals between adjacent data points for each vehicle, we regard the speed variation of the vehicle as uniform acceleration motion. Assuming the speed of the vehicle at point O is $v_1$, the speed at point D is $v_2$, and the travel time between the two points is $T$ we then use Eq. (1) to compute the distance $L$ between O and D.

$$L = (v_1/7.2 + v_2/7.2) \times T \tag{1}$$

Following this, we use big data processing methods to conduct grid matching and visualization of the floating vehicle data (Yu and Li, 2022; Yu and Yuan, 2022). Firstly, grids within the geographical scope of the study area are generated, and Fig. 2 shows the generated grid map of Shenzhen City. Then the floating vehicle data is matched to each grid to analyze the spatio-temporal characteristics of traffic congestion and pollutant emission.

### 3.2. Urban congestion analysis methods

#### 3.2.1. Traffic congestion index

Once the floating vehicle data for the whole day is gridded, each grid will contain all the characteristics of the floating vehicles that passed through it during the day, including license plate number, time, and speed. Based on this data, the data for the entire day are grouped according to the respective grid, and the maximum driving speed for each grid can be determined.

In this study, the free speed of each grid is defined as the maximum travel speed and is used to calculate the congestion index ($K$) (Li et al., 2020), as shown in Eq. (2):

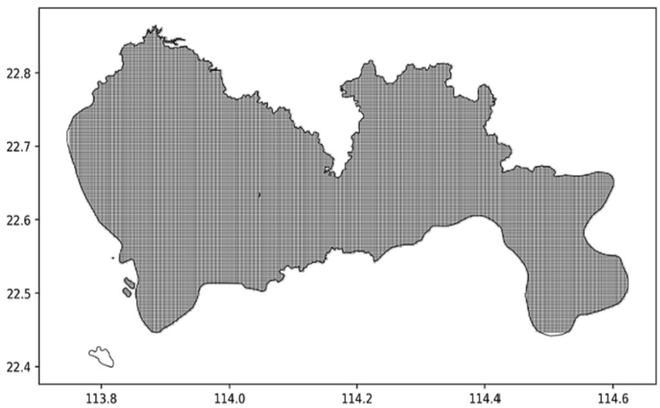



**Fig. 2.** Gridded map of Shenzhen.

$$K = v/V \tag{2}$$

The congestion index $K$ is a crucial indicator for evaluating the traffic conditions of a road network because it reflects the traffic flow through the roadway. Specifically, a smaller $K$ value indicates heavier traffic congestion, while a larger $K$ value suggests smoother traffic conditions. The road network can be divided into colour-coded regions according to $K$ values, ranging from red to green, to help road users quickly understand the current traffic situation of an area. By observing the distribution characteristics of $K$ values at different periods, it is possible to gain further insights into the dynamic changes in urban traffic conditions and develop specific strategies such as adjusting traffic signal timings, re-routing vehicles, or implementing other measures to optimize the use of road space and reduce travel times, thus enhancing the operational efficiency and service quality of the road network.

#### 3.2.2. K-means clustering

The k-means algorithm is an unsupervised machine learning algorithm used to partition a dataset into k different groups or clusters based on members with similar features (Likas et al., 2003). The K-means clustering algorithm has a simple idea and converges faster compared to other algorithms. Moreover, the clustering effect is better, and the only parameter that needs to be adjusted is the number of clusters, K. The algorithm aims to partition n observations into k clusters, in which each observation belongs to the cluster with the nearest mean. In our study, urban traffic congestion is explored, and traffic conditions are categorized into four different categories. Therefore, we select the k-means clustering algorithm that best suited the specific demands of our study. The algorithm works as shown in Table 2.

In this paper, we use the speed and $K$ congestion index of each vehicle as features to categorize the data. We then perform a k-means clustering analysis during each period to determine the traffic condition in the city. In this study, we apply the k-means algorithm to cluster the vehicle speed data into different congestion levels such as smooth flow, slow flow, congestion, and severe congestion, resulting in four clusters. We compare and analyze the obtained results with those of the congestion index. The flow of the K-means clustering algorithm is shown in Fig. 3.

### 3.3. Calculate pollutant emission based on standards

According to the current vehicle pollutant emission standards(Ministry of Environment Protection of the People's Republic of China, 2019), the emission factors of various pollutants at different speeds were calculated by combining the characteristics of vehicle models, the geographical environment of Shenzhen, climate and other influencing factors. The results are shown in Table 3.

After obtaining the pollution emission factors for different speeds, the distance traveled by the floating vehicle, the density of gasoline and the fuel consumption per km were combined to calculate the emission of different pollutants according to Eq. (3) (Ministry of Environment Protection of the People's Republic of China, 2019):

$$E_i = EF_i \times L \times 10^{-3} \tag{3}$$

Where, $E_i$ is the emission of $i$ pollutant; $EF_i$ is the emission factor of $i$ pollutant; $L$ is the distance traveled by the vehicle in meters.

The emission calculations for each type of pollutant are performed sequentially and then visualized on a map of Shenzhen. The resulting data is analyed to determine the impact of urban congestion on pollutant emission.

### 3.4. Calculate $CO_2$ emission based on the fuel consumption-speed model

The carbon emission calculation method used in this study is the fuel consumption-speed model of Peng (2014) and Yao et al. (2022a). The fuel type of the vehicle is gasoline, and the carbon emission of the vehicle between a period of OD is assumed to be C, which is calculated by Eqs. (4) and (5):

$$C = F \times P \times L \times 10^{-5} \tag{4}$$

$$F = a \times v^2 + b \times v + c \tag{5}$$

Where, $C$ is the emission of $CO_2$; $P$ is the vehicle carbon emission factor, which is 2322.21 g/L in this paper; $F$ is the fuel consumption per 100 km; $L$ is the distance between ODs; and $a$, $b$, and $c$ are the regression parameters, which take the values of 0.0052, $-0.9734$, and 54.3605, respectively.

**Table 2**
K-means algorithm process.

| Step | Description |
|---|---|
| 1 | Decide on the number of clusters ($k$) |
| 2 | Randomly select k points from the data set as initial centroids |
| 3 | Assign each data point to the closest centroid using Euclidean distance as the distance metric |
| 4 | Recompute the centroid of each cluster by taking the mean of all the points assigned to it |
| 5 | Repeat steps 3 and 4 until the centroids no longer move, or a maximum number of iterations is reached |
| 6 | Obtain k clusters of data points that are most similar to each other based on their distances to the assigned centroids |

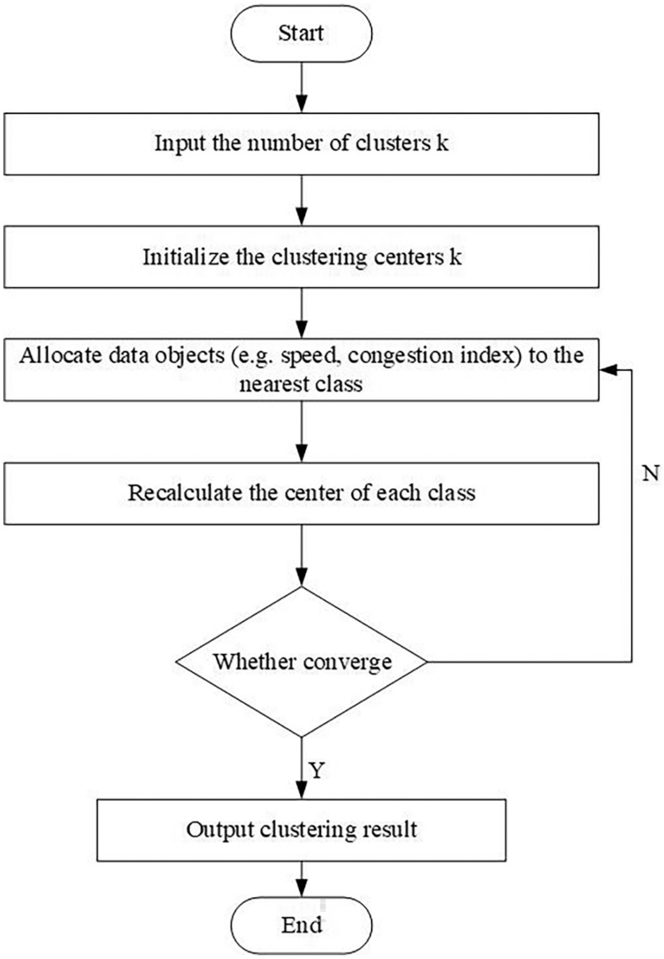



**Fig. 3.** K-means clustering algorithm flowchart.

**Table 3**
Emission factors for different speeds.

| Pollutant ($g/km$) | Speed range ($km/h$) | | | |
|---|---|---|---|---|
| | <20 | 20–30 | 30–40 | >40 |
| CO | 5.71 | 4.26 | 2.67 | 1.32 |
| HC | 0.65 | 0.48 | 0.30 | 0.12 |
| $NO_x$ | 0.13 | 0.11 | 0.083 | 0.08 |
| PM | 0.05 | 0.0375 | 0.0234 | 0.0096 |

## 4. Case study

We conduct a comprehensive analysis using Shenzhen city's taxi data as a case study. Firstly, we examine the distribution of taxi orders by dividing the geographical area of Shenzhen into grids of 2000 m × 2000 m in size. Subsequently, we perform a statistical analysis to discern the overarching order distribution patterns. The visualization of this analysis is presented in Fig. 4.

We begin by extracting taxi trajectories during passenger transportation based on a passenger status of 1 in the data. Subsequently, we derive the origin-destination (OD) pairs of the orders, outlining the starting and ending points of each order. The OD pairs provide valuable insights into the economic vitality of the respective regions. Through the OD distribution of passenger boarding and alighting, we can find that passengers are concentrated in the Bao'an, Futian, Longhua, and Luohu areas of Shenzhen (Fig. 4). It is also not difficult to find that there are more passengers in the places where airports and high-speed railway stations are located, and the flow of people in these places is also greater.

Guided by the empirical distribution of the order dataset, we establish a foundational framework to underpin our forthcoming analyzes concerning the intricate issues of traffic congestion and pollutant emission. This study focuses on the city of Shenzhen and

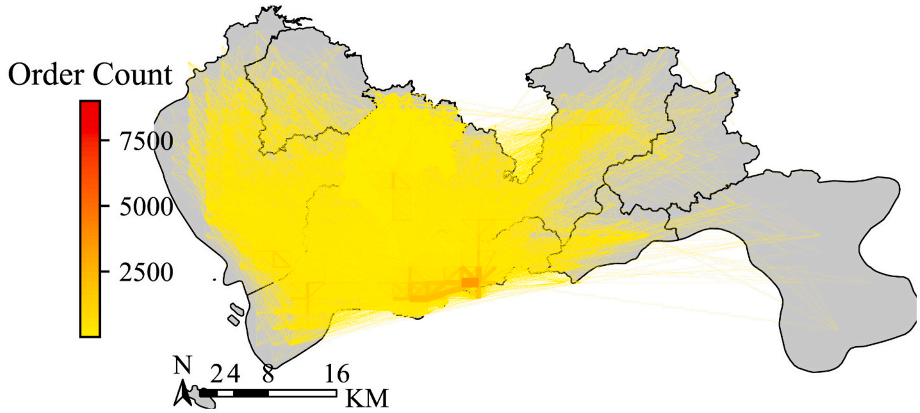
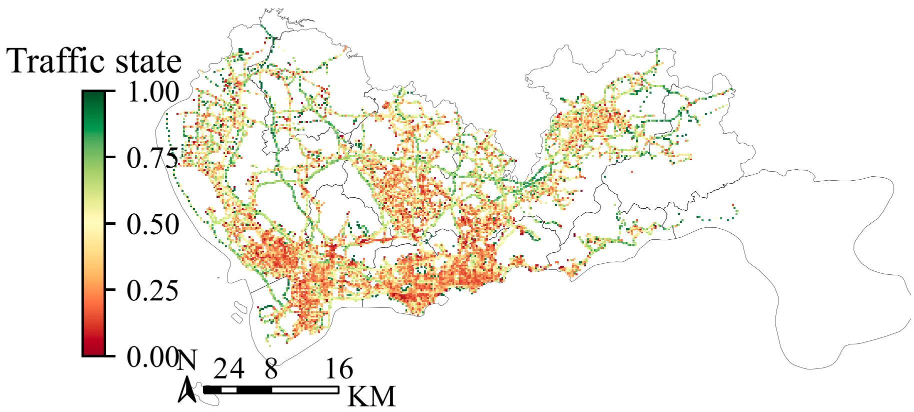
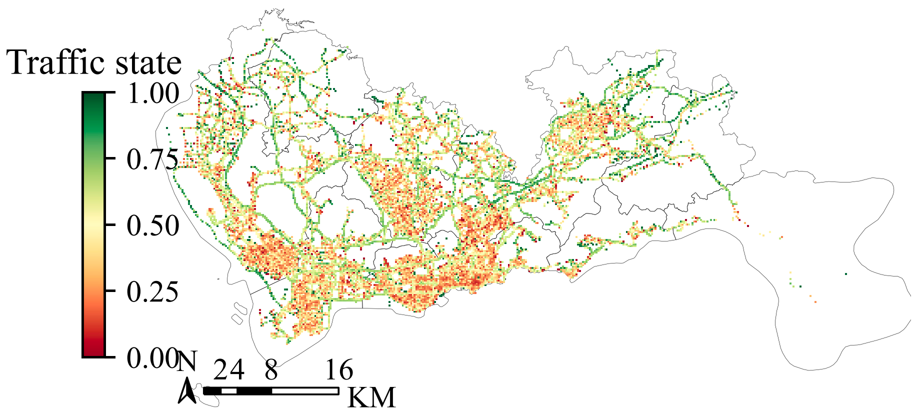



**Fig. 4.** Order distribution.




(a) 8:00-8:30




(b) 23:00-23:30

**Fig. 5.** Traffic state map based on congestion index.

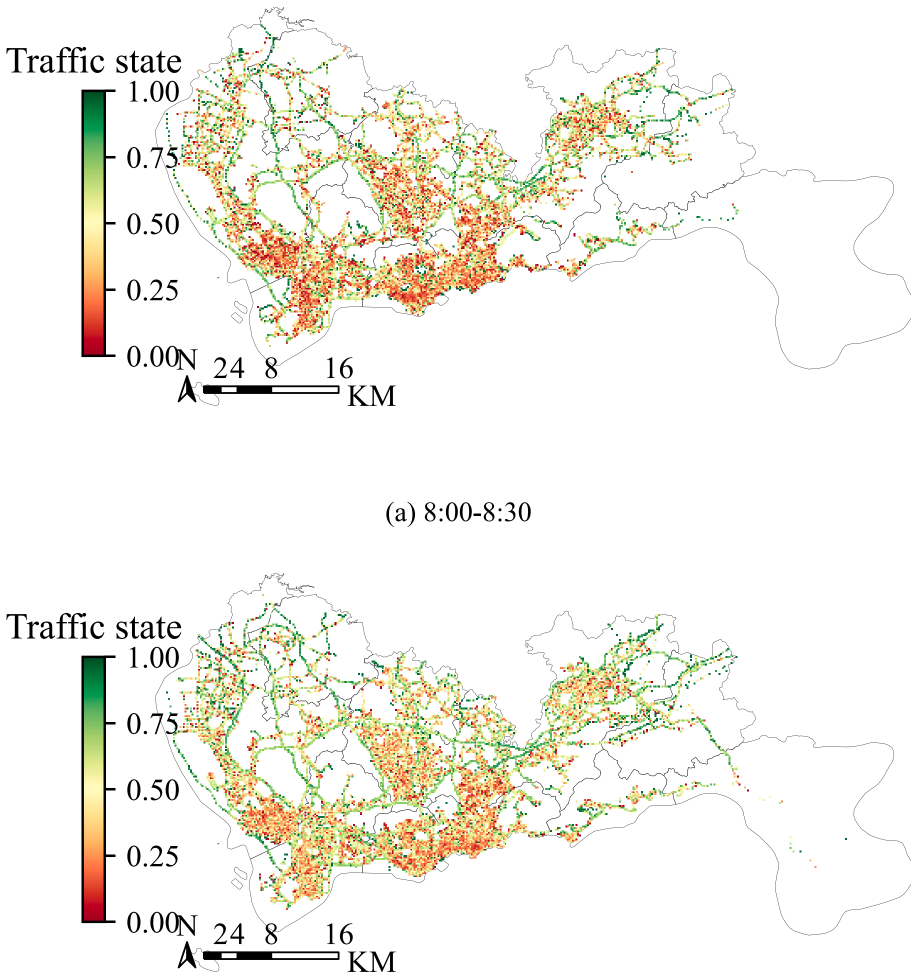
reasonably selects two distinct time intervals, namely the peak period (8:00–8:30) and the off-peak period (23:00–23:30), for an in-depth investigation of traffic congestion dynamics and pollutant emission. We have not only analyzed the relationship between traffic congestion and pollutant emission, but have also extended our research to reveal the complex interactions and causal relationships between these multi-dimensional urban phenomena.

### *4.1. Congestion index analysis*

The colour scheme utilized in this paper is based on the congestion index $K$, which ranges from smooth to congested.

Fig. 5 presents the traffic operation status during (a) the morning peak and (b) at 11:00 p.m. It is evident that more red areas are present in Fig. 5 (a), particularly during peak travel periods, with Futian, Nanshan, and Luohu experiencing severe congestion due to the high mobility of residents and various entertainment options in these areas. Although some sections of the Beijing-Hong Kong-Macao Expressway and Shenzhen Northern Ring Road remain open, certain roads near airports and stations exhibit slow traffic.

Relatively speaking, some highways in Shenzhen are smoother than other roads in the city, as shown in Fig. 5 (b). Late-night conditions are notably better than those during the daytime since there are significantly fewer vehicles on the road than in the evening and morning rush hours. Despite being crowded, roads around high-speed railway stations display greater congestion compared to non-concentrated areas, where the roads are much smoother. Fewer individuals travelling leads to a reduced number of trips and vehicles on the road, resulting in a decreased level of congestion. It can therefore be inferred that late-night traffic is considerably less congested in Shenzhen than during peak hours.




(a) 8:00-8:30

(b) 23:00-23:30

**Fig. 6.** Traffic state map based on K-means algorithm.

### 4.2. K-means clustering analysis

In this study, we use a K-mean clustering algorithm to assess the traffic congestion of floating vehicles in Shenzhen city. By categorizing the vehicles into four groups based on their speed and congestion index, we can determine the average congestion index for each grid. The resulting road traffic map of Shenzhen city, as depicted in Fig. 6, highlighted areas such as Bao'an International Airport, Nanshan, Futian, and Luohu as experiencing higher levels of congestion in comparison to smoother regions. Our analysis of the K-means clustering algorithm and the congestion index outcomes demonstrate that the approach used in this study is reasonable and accurately reflects real traffic conditions. The clustering results are consistent with those generated solely from the congestion index, confirming the effectiveness of the K-mean clustering algorithm in analyzing floating vehicle congestion.

### 4.3. Pollutant emission analysis based on the current standards

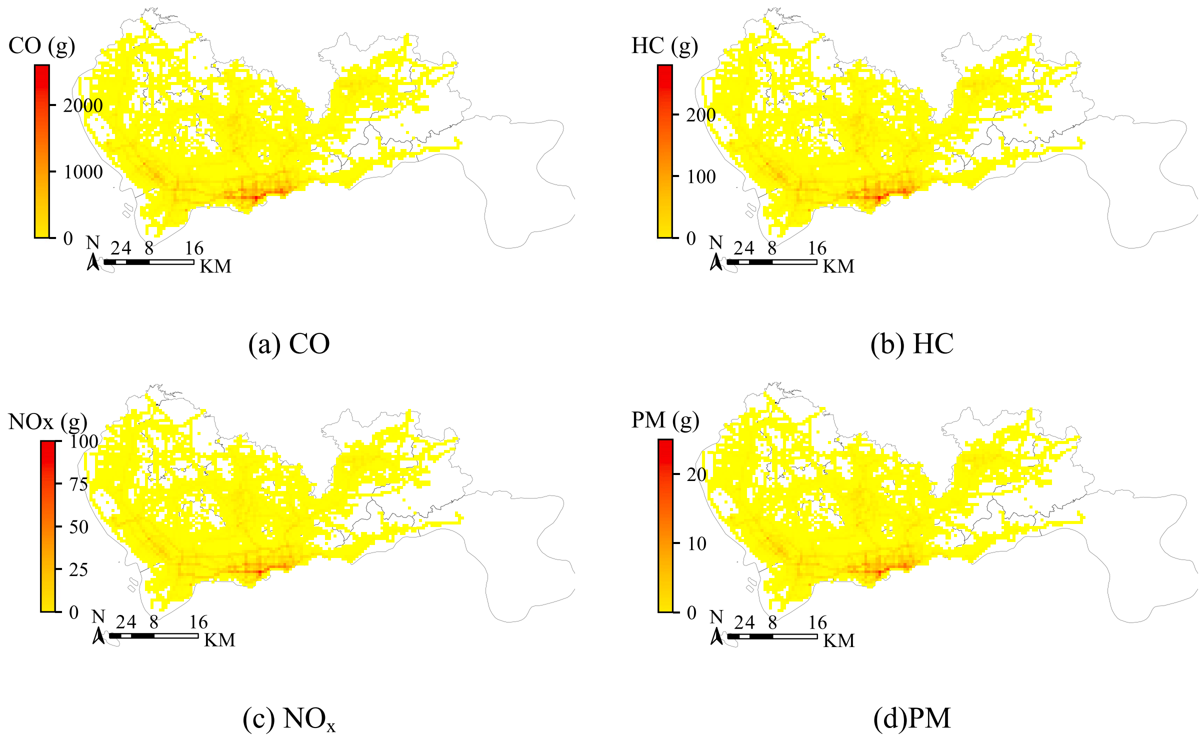
This section presents an overview of the emission of various pollutants, which is calculated based on the current vehicle emission standards. Fig. 7 and Fig. 8 show the emission of pollutants on the roads in Shenzhen during the morning peak and at night.

The results in Fig. 7 show that there is more pollutant emission in the more congested areas than in the smooth areas due to the poor road operating conditions and the lower speed values of floating vehicles travelling on the city roads at this time. For example, the roads in the areas near the Bao'an International Airport, the city government, and the railroad station show more red areas at this time. Roads in other areas, such as Longgang, have better traffic conditions so there are fewer pollutants in these areas. Based on the above analysis, it can be seen that road congestion is linked to greater pollutant emission, probably as a result of inadequate combustion of fuel in the cab and interruptions in the vehicle speed when travelling.

Pollutant emission are much lower on roads late at night compared to the morning peak, as shown in Fig. 8. At night, most of the roads in the area are in good condition, so not much pollutant is emitted. It can also be seen from the figure that pollutant emission on roads is higher only in areas with high activity, such as around the Bao'an International Airport in Bao'an District and the high-speed rail station Shenzhen Station in Luohu District. On the contrary, some highways and overpasses on the roads have low pollutant emission. Therefore, it can be inferred that the construction of large-scale transportation infrastructure and the subsequent increase in vehicle speeds can considerably reduce pollution caused by transportation.

### 4.4. $CO_2$ emission analysis based on the fuel consumption-speed model

Fig. 9 shows emission levels during the morning peak hour and at 11 p.m. As shown in Figures (a) and (b), $CO_2$ emission is higher in the Futian and Luohu districts, which are mainly concentrated near Futian and Shenzhen stations. These areas are characterized by a




(a) CO (b) HC

(c) $NO_x$ (d)PM

**Fig. 7.** Emission at 8:00 to 8:30.

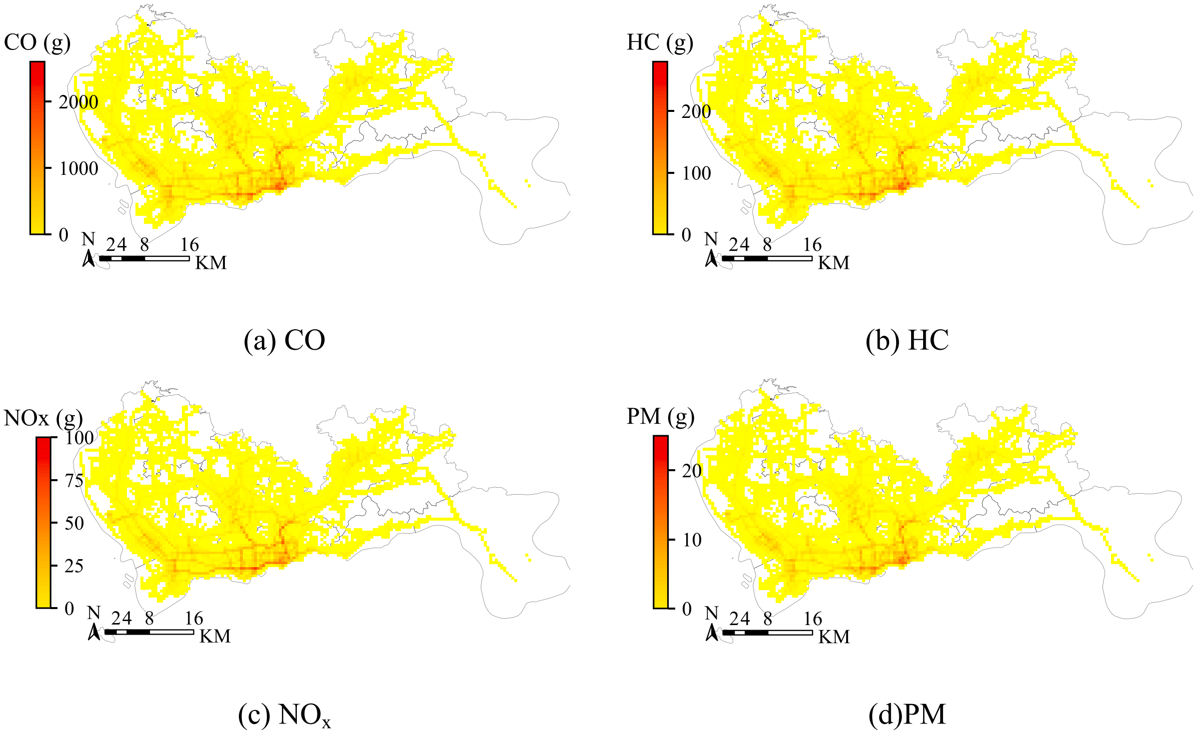



**Fig. 8.** Emission at 23:00 to 23:30.

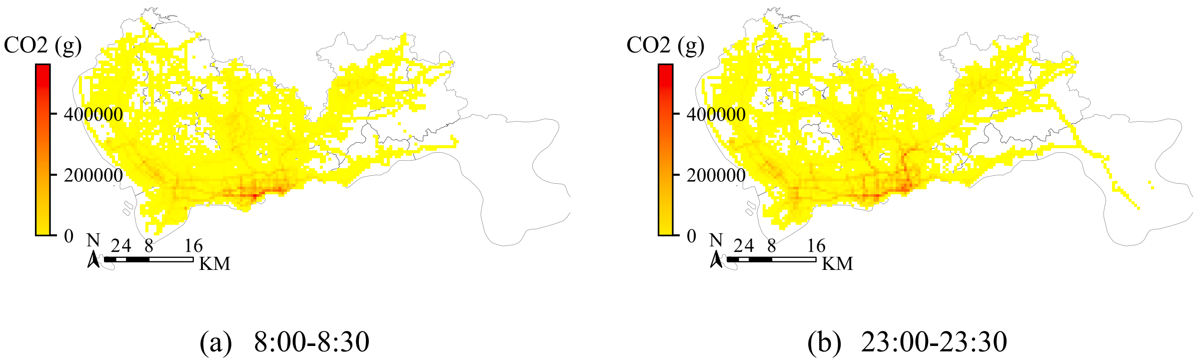



**Fig. 9.** CO2 emission.

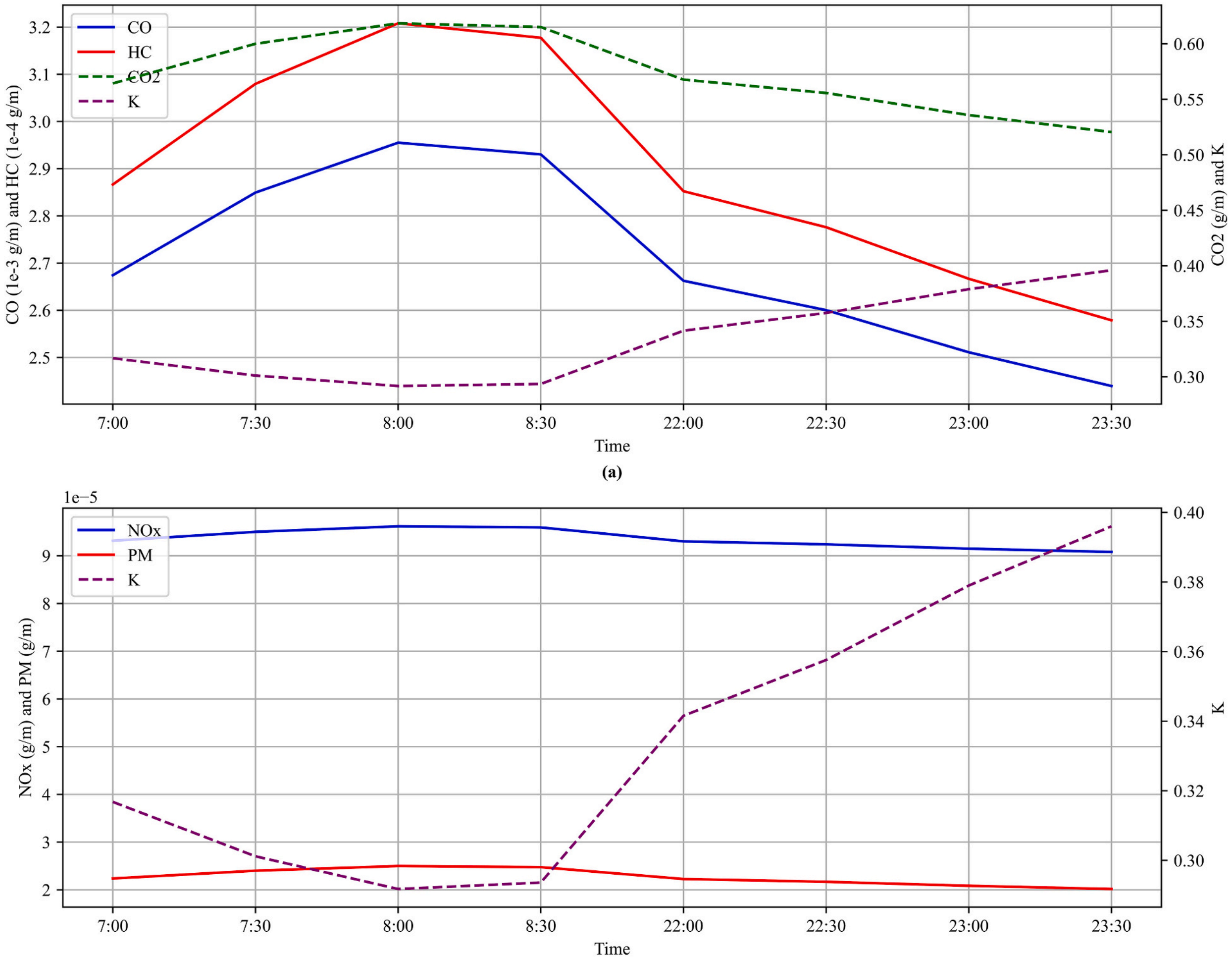
high density of office buildings, shopping malls and commercial centres, which naturally attracts more passengers, leading to a more concentrated distribution of cab orders in these areas.

Comparing the two time periods, it is clear that $CO_2$ emission during the morning peak hour is greater than that at 11 p.m. This difference highlights the relationship between traffic congestion and the increase in $CO_2$ emission. Due to the increased demand for taxi services, traffic congestion increases during the morning peak hour, leading to an increase in $CO_2$ emission.

### *4.5. The relationship between congestion and emission*

In this paper, we calculate the average pollutant emission and the average congestion index of floating vehicles for different periods in 30-min time units. We draw a comparative analysis of the relationship between various pollutant emission and congestion index. Fig. 10 and Fig. 11 describe the relationship between CO, HC, $NO_x$, PM and $CO_2$ emission and the average congestion index.

The horizontal axis in Fig. 10 represents time, and the vertical axis represents the average emission per unit distance for CO, HC, $NO_x$, PM, and $CO_2$ from floating vehicles, respectively. Meanwhile, the congestion index is drawn with them to observe congestion and emission at different times. In order to clearly demonstrate the relationship between congestion index and various pollutant emissions, we present all values on the same scale in one graph via normalization. As can be seen from Fig. 11, the congestion index shows a




(a)

(b)

**Fig. 10.** Relationship between pollutant emission per metre and congestion index.

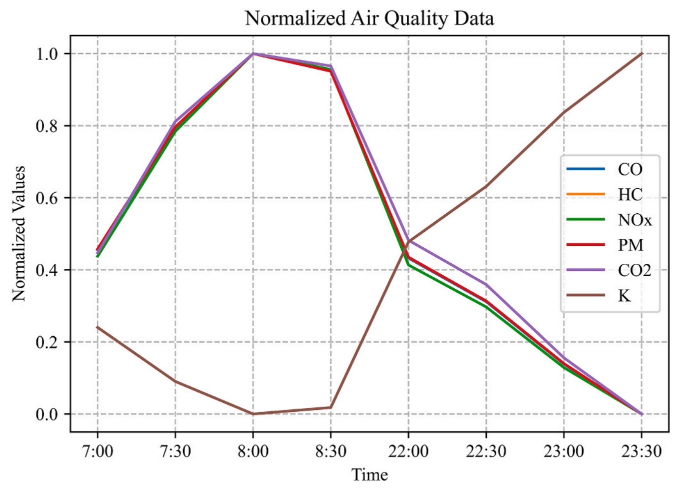



**Fig. 11.** Relationship between normalized pollutant emission and congestion index.

negative correlation with the emissions of various pollutants. Although there may be slight differences in some normalized values of these pollutants, the trend of these curves are consistent. From the figure, we can see that the higher the congestion index, the lower the emission of various pollutants, while the lower the congestion index, the higher the pollutant emission. This inverse relationship between pollutant emission and congestion index indicates that when urban road traffic is severely congested, the congestion index of motor vehicles decreases, which leads to an increase in pollutant emission. Conversely, pollutant emission decreases during hours when traffic is less congested or smooth.

Reducing traffic congestion can be an effective way to reduce emission in urban areas, as shown by the inverse relationship between emission and congestion index. These findings have important implications for rational transportation planning to improve traffic congestion and reduce pollutant emission. For example, urban traffic managers may give priority and rapid response to the areas with high levels of traffic congestion and pollutant emission identified by FVD. Following this, they can use measures such as congestion pricing, encouraging public transportation, and implementing traffic management policies to improve congestion and thus reduce pollutant emission in these areas.

## 5. Conclusions

This study focuses on analyzing traffic congestion and pollutant emission problems based on floating vehicle (taxi) data in Shenzhen. We use big data methods and clustering algorithms to analyze the spatio-temporal characteristics of congestion, $CO_2$ emission and pollutant emission in Shenzhen. Our findings indicate that taxi orders are mainly concentrated in transportation hub areas, such as Shenzhen Bao'an International Airport and Shenzhen Railway Station. From a temporal perspective, road traffic congestion in Shenzhen is more severe during the morning and evening peak hours, and pollutant emission (e.g., $CO_2$, $NO_x$) are thus also higher during these times. From a spatial perspective, traffic congestion in densely populated areas such as Bao'an, Futian, and Luohu is very severe, and the pollutant emission are higher than those in other areas of the city.

Our research provides a basis for sustainable transportation management in cities and offers references for the transportation sector to achieve carbon neutrality. Most countries promote low-carbon transportation, and reasonable traffic management and urban planning can achieve greater environmental benefits. In addition, traffic management in these areas may achieve better effects via focusing on heavily congested and high-emission transportation hubs. In terms of policy implications, this study emphasizes the importance of taking measures to reduce traffic congestion and pollutant emissions in urban areas. Several strategies can be considered, including encouraging the use of new energy vehicles, promoting shared mobility, and optimizing traffic management in congestion-prone areas. Based on this study, urban traffic managers may achieve sustainable transportation, mitigate the environmental impacts of congestion, and promote the achievement of carbon neutrality in transportation.

Inevitably, this study has several limitations. Firstly, it only utilizes floating vehicle data for taxi, which limits the accuracy of the congestion and pollution. We will expand our data collection to include a wider range of vehicle types, such as new energy vehicles, and analyze their impacts on emissions throughout their entire life-cycle. This will enable us to gain a more comprehensive understanding of the overall emission situation. In addition, the dataset used in this study is not large and thus may not fully capture the impacts of external factors, such as seasons and holidays, on traffic congestion and pollution emission. In future research, we plan to expand the dataset to incorporate data from different seasons and holidays, enabling us to gain a more comprehensive understanding of the variations in traffic patterns and emissions throughout the year. Finally, we do not consider the cost of using taxis, but the cost can influence people's willingness to take taxis or not, which in turn affects the number of orders. In the future, therefore, we may attempt to utilize other data to conduct an in-depth study on the spatio-temporal characteristics of carbon emission for urban transportation. These will provide significant theoretical and practical guidance for city managers to formulate and improve transportation planning, increase urban transportation efficiency, and reduce pollution emission.

## CRediT authorship contribution statement

**Wen-Long Shang:** Conceptualization, Methodology, Data curation, Funding acquisition, Writing – review & editing, Supervision. **Xuewang Song:** Conceptualization, Methodology, Visualization, Investigation, Software, Writing – original draft, Formal analysis. **Yishui Chen:** Conceptualization, Methodology, Software, Formal analysis. **Xin Yang:** Conceptualization, Methodology, Resources. **Liyun Liang:** Investigation, Formal analysis. **Muhammet Deveci:** Conceptualization, Methodology, Resources. **Mengqiu Cao:** Conceptualization, Methodology, Supervision. **Qiannian Xiang:** Investigation, Software. **Qing Yu:** Conceptualization, Methodology, Supervision.

## 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

Data will be made available on request.

## Acknowledgments

This research is supported in part by Beijing Natural Science Foundation (NO. 9232003 and NO. L211027) and in part by R&D Program of Beijing Municipal Education Commission (NO. KM202210005001).

## References

Bai, S., Bi, X., Han, C., et al., 2022. Evaluating R&D efficiency of China's listed lithium battery enterprises. Front. Eng. Manag. 9, 473–485. https://doi.org/10.1007/s42524-022-0213-5.

Bendib, A., 2020. The effects of spatial clustering of public facilities on social equity and urban congestion in the city of Batna (Algeria). GeoJournal 87 (2), 861–874. https://doi.org/10.1007/s10708-020-10289-y.

Beojone, C.V., Geroliminis, N., 2021. On the inefficiency of ride-sourcing services towards urban congestion. Transp. Res. Part C 124. https://doi.org/10.1016/j.trc.2020.102890.

Chen, G., Zhang, J.W., 2022. Applying artificial intelligence and deep belief network to predict traffic congestion evacuation performance in smart cities. Appl. Soft Comput. 121, 108692 https://doi.org/10.1016/j.asoc.2022.108692.

Chen, Y., Liu, Y., Lin, X., Huang, J., 2016. Evaluating the effects of traffic states on light-duty vehicle emissions in urban area based on floating Car data. Res. Environ. Sci. 29 (04), 494–502. https://doi.org/10.13198/j.issn.1001-6929.2016.04.04.

Cheng, Y., Yu, L., Wang, H.-T., Hao, Y.-Z., Song, G.-H., 2011. Comparative study of MOBILE and COPERT emission models based on PEMS. J. Transp. Syst. Eng. Inf. Technol. 11 (03), 176–181. https://doi.org/10.16097/j.cnki.1009-6744.2011.03.027.

Çolak, S., Lima, A., González, M.C., 2016. Understanding congested travel in urban areas. Nat. Commun. 7 (1), 10793. https://doi.org/10.1038/ncomms10793.

Dakic, I., Menendez, M., 2018. On the use of Lagrangian observations from public transport and probe vehicles to estimate car space-mean speeds in bi-modal urban networks. Transp. Res. Part C 91, 317–334. https://doi.org/10.1016/j.trc.2018.04.004.

Du, Y., Wu, J., Yang, S., Zhou, L., 2017. Predicting vehicle fuel consumption patterns using floating vehicle data. J. Environ. Sci. (China) 59, 24–29. https://doi.org/10.1016/j.jes.2017.03.008.

Ehsani, M., Ahmadi, A., Fadai, D., 2016. Modeling of vehicle fuel consumption and carbon dioxide emission in road transport. Renew. Sust. Energ. Rev. 53, 1638–1648. https://doi.org/10.1016/j.rser.2015.08.062.

Ekström, M., Sjödin, Å., Andreasson, K., 2004. Evaluation of the COPERT III emission model with on-road optical remote sensing measurements. Atmos. Environ. 38 (38), 6631–6641. https://doi.org/10.1016/j.atmosenv.2004.07.019.

Ercan, T., Onat, N.C., Keya, N., Tatari, O., Eluru, N., Kucukvar, M., 2022. Autonomous electric vehicles can reduce carbon emissions and air pollution in cities. Transp. Res. Part D 112, 103472. https://doi.org/10.1016/j.trd.2022.103472.

Fanchao, D., 2023. The Motor Vehicle Ownership in China Has Reached 426 Million. Legal Daily. YNET.

Fu, L., Hao, J., He, D., He, K., 2000. The emission characteristics of pollutants from motor vehicles in Beijing. Environ. Sci. 03, 68–70. https://doi.org/10.13227/j.hjkx.2000.03.016.

Geroliminis, N., Daganzo, C.F., 2008. Existence of urban-scale macroscopic fundamental diagrams: some experimental findings. Transp. Res. Part B 42 (9), 759–770. https://doi.org/10.1016/j.trb.2008.02.002.

Gokasar, I., Karaman, O., 2023. Integration of personnel services with public transportation modes: a case study of Bogazici university. J. Soft Comput. Decis. Anal. 1 (1), 1–17. https://doi.org/10.31181/jscda1120231.

Grote, M., Williams, I., Preston, J., Kemp, S., 2016. Including congestion effects in urban road traffic CO2 emissions modelling: do local government authorities have the right options? Transp. Res. Part D: Transp. Environ. 43, 95–106. https://doi.org/10.1016/j.trd.2015.12.010.

Guo, Y., Tang, Z., Guo, J., 2020. Could a Smart City ameliorate urban traffic congestion? A quasi-natural experiment based on a Smart City pilot program in China. Sustainability 12 (6), 2291. https://doi.org/10.3390/su12062291.

Gurney, K.R., et al., 2022. Greenhouse gas emissions from global cities under SSP/RCP scenarios, 1990 to 2100. Glob. Environ. Change 73, 102478. https://doi.org/10.1016/j.gloenvcha.2022.102478.

Hammami, F., 2020. The impact of optimizing delivery areas on urban traffic congestion. Res. Transp. Bus. Manag. 37, 100569 https://doi.org/10.1016/j.rtbm.2020.100569.

Hao, Y., Deng, S., Qiu, Z., Li, Q., Gao, C., Xu, Y., 2017. Vehicle emission inventory for Xi'an based on MOVES model. Environ. Pollut. Prevent. 39 (03) https://doi.org/10.15985/j.cnki.1001-3865.2017.03.001, 227-231+235.

Hincks, S., Carter, J., Connelly, A., 2023. A new typology of climate change risk for European cities and regions: principles and applications. Glob. Environ. Chang. 83, 102767 https://doi.org/10.1016/j.gloenvcha.2023.102767.

Houghton, J., 2005. Global warming. Rep. Prog. Phys. 68 (6), 1343–1403. https://doi.org/10.1088/0034-4885/68/6/R02.

Huang, Z., Xia, J., Li, F., Li, Z., Li, Q., 2019. A peak traffic congestion prediction method based on bus driving time. Entropy (Basel) 21 (7). https://doi.org/10.3390/e21070709.

Janecek, A., Valerio, D., Hummel, K.A., Ricciato, F., Hlavacs, H., 2015. The cellular network as a sensor: from Mobile phone data to real-time road traffic monitoring. IEEE Trans. Intell. Transp. Syst. 16 (5), 2551–2572. https://doi.org/10.1109/Tits.2015.2413215.

Jiang, Y., Song, G.H., Zhang, Z.Y., Zhai, Z.Q., Yu, L., 2021. Estimation of hourly traffic flows from floating Car data for vehicle emission estimation. J. Adv. Transp. 2021, 6628335. https://doi.org/10.1155/2021/6628335.

Kerner, B.S., et al., 2005. Traffic state detection with floating car data in road networks. In: Proceedings. 2005 IEEE Intelligent Transportation Systems, 2005. https://doi.org/10.1109/ITSC.2005.1520133.

Li, S., Sun, H., Chen, Y., 2020. Research on urban road congestion bottleneck identification based on floating vehicle data. In: China Intelligent Transportation Association. Proceedings of the 15th Annual Conference on Intelligent Transportation in China, vol. 2. https://doi.org/10.26914/c.cnkihy.2020.028416.

Li, M.L., Yan, M., He, H.W., Peng, J.K., 2021. Data-driven predictive energy management and emission optimization for hybrid electric buses considering speed and passengers prediction. J. Clean. Prod. 304 https://doi.org/10.1016/j.jclepro.2021.127139.

Li, Z.X., Liu, A.J., Shang, W.L., Li, J.X., Lu, H., Zhang, H.R., 2023. Sustainability assessment of regional transportation: an innovative fuzzy group decision-making model. IEEE Trans. Intell. Transp. Syst. https://doi.org/10.1109/Tits.2023.3275141.

Likas, A., Vlassis, N., Verbeek, J.J., 2003. The global k-means clustering algorithm. Pattern Recogn. 36 (2), 451–461.

Liu, Q., Li, H., Shang, W.-l., Wang, K., 2022. Spatio-temporal distribution of Chinese cities' air quality and the impact of high-speed rail. Renewable and Sustainable Energy Reviews 170, 112970. https://doi.org/10.1016/j.rser.2022.112970.

Mingolla, S., Lu, Z.M., 2021. Carbon emission and cost analysis of vehicle technologies for urban taxis. Transp. Res. Part D 99, 102994. https://doi.org/10.1016/j.trd.2021.102994.

Ministry of Ecology and Environment of the People''s Republic of China, 2022. China Mobile Source Environmental Management Annual Report(2022) [WWW Document]. https://www.mee.gov.cn/hjzl/sthjzk/ydyhjgl/202212/W020221207387013521948.pdf (accessed 11.11.23).

Ministry of Environment Protection of the People''s Republic of China, 2019. Technical Guide for the Preparation of Road Motor Vehicle Emission Inventory (for Trial Implementation) (Draft for Comments) [WWW Document]. https://www.mee.gov.cn/gkml/hbb/bgth/201407/W020140708387895271474.pdf (accessed 11.11.23).

Moslem, S., Solieman, H., Oubahman, L., Duleba, S., Senapati, T., Pilla, F., 2023. Assessing public transport supply quality: a comparative analysis of analytical network process and analytical hierarchy process. J. Soft Comput. Decis. Anal. 1 (1), 124–138. https://doi.org/10.31181/jscda11202311.

Mukherjee, P., Viswanathan, S., 2001. Carbon monoxide modeling from transportation sources. Chemosphere 45 (6–7), 1071–1083. https://doi.org/10.1016/S0045-6535(01)00128-X.
Nigro, M., et al., 2022. Exploiting floating car data to derive the shifting potential to electric micromobility. Transp. Res. Part A 157, 78–93. https://doi.org/10.1016/j.tra.2022.01.008.
Olusanya, G.S., Eze, M.O., Ebiesuwa, O., Okunbor, C., 2020. Smart transportation system for solving urban traffic congestion. Rev. Comput. Eng. Stud. 7 (3), 55–59. https://doi.org/10.18280/rces.070302.
Othman, B., De Nunzio, G., Laraki, M., Sabiron, G., IEEE, 2022. A novel approach to traffic flow estimation based on floating car data and road topography: experimental validation in Lyon, France. In: 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC). https://doi.org/10.1109/ITSC55140.2022.9922602.
Paris Agreement [WWW Document]. https://www.un.org/zh/documents/treaty/FCCC-CP-2015-L.9-Rev.1, 2016 (accessed).
Peng, B., 2014. Research about the Model of Vehicle's Fuel Consumption on the Highway. Master. Harbin Institute of Technology. https://kns.cnki.net/kcms2/article/abstract?v=xzY5Ip_Thcnkm_tl90bl-Zf0c-vfTlxuX9_GSVEJxuOStnkwKkpeN9jWL4f-b40_4M_kvglIHRMJSahdRHFSihwJaWTo4FFHYiMIDrhaZJEMBodblwGXCTgWvDKwNekY6DFQNK0KjwI=&uniplatform=NZKPT&language=CHS.
Perugu, H., 2019. Emission modelling of light-duty vehicles in India using the revamped VSP-based MOVES model: the case study of Hyderabad. Transp. Res. Part D 68, 150–163. https://doi.org/10.1016/j.trd.2018.01.031.
Qian, X.W., Lei, T., Xue, J.W., Lei, Z.X., Ukkusuri, S.V., 2020. Impact of transportation network companies on urban congestion: evidence from large-scale trajectory data. Sustain. Cities Soc. 55 https://doi.org/10.1016/j.scs.2020.102053.
Qin, J.Y., Mei, G., Xiao, L., 2021. Building the traffic flow network with taxi GPS trajectories and its application to identify urban congestion areas for traffic planning. Sustainability 13 (1). https://doi.org/10.3390/su13010266.
Rahmani, M., Koutsopoulos, H.N., 2013. Path inference from sparse floating car data for urban networks. Transp. Res. Part C 30, 41–54. https://doi.org/10.1016/j.trc.2013.02.002.
Rahmani, M., Jenelius, E., Koutsopoulos, H.N., 2015. Non-parametric estimation of route travel time distributions from low-frequency floating car data. Transp. Res. Part C 58, 343–362. https://doi.org/10.1016/j.trc.2015.01.015.
Rempe, F., Franeck, P., Fastenrath, U., Bogenberger, K., 2017. A phase-based smoothing method for accurate traffic speed estimation with floating car data. Transp. Res. Part C 85, 644–663. https://doi.org/10.1016/j.trc.2017.10.015.
Rempe, F., Franeck, P., Bogenberger, K., 2022. On the estimation of traffic speeds with Deep Convolutional Neural Networks given probe data. Transp. Res. Part C 134. https://doi.org/10.1016/j.trc.2021.103448. ARTN 103448.
Shang, W.L., Lv, Z.H., 2023. Low carbon technology for carbon neutrality in sustainable cities: a survey. Sustain. Cities Soc. 92, 104489. ARTN 104489. https://doi.org/10.1016/j.scs.2023.104489.
Shang, W.L., Chen, J., Bi, H., Sui, Y., Chen, Y., Yu, H., 2021. Impacts of COVID-19 pandemic on user behaviors and environmental benefits of bike sharing: a big-data analysis. Appl. Energy 285, 116429. https://doi.org/10.1016/j.apenergy.2020.116429.
Shang, W.L., et al., 2023. Spatio-temporal analysis of carbon footprints for urban public transport systems based on smart card data. Appl. Energy 352, 121859. https://doi.org/10.1016/j.apenergy.2023.121859.
Shang, W.-L., Zhang, M., Wu, G., Yang, L., Fang, S., Ochieng, W., 2023. Estimation of traffic energy consumption based on macro-micro modelling with sparse data from Connected and Automated Vehicles. Applied Energy 351, 121916. https://doi.org/10.1016/j.apenergy.2023.121916.
Shatanawi, M., Abdelkhalek, F., Mészáros, F., 2020. Urban congestion charging acceptability: an international comparative study. Sustainability 12 (12). https://doi.org/10.3390/su12125044.
Shindell, D., et al., 2011. Climate, health, agricultural and economic impacts of tighter vehicle-emission standards. Nat. Clim. Chang. 1 (1), 59–66. https://doi.org/10.1038/Nclimate1066.
Sui, Y., et al., 2019. GPS data in urban online ride-hailing: a comparative analysis on fuel consumption and emissions. J. Clean. Prod. 227, 495–505. https://doi.org/10.1016/j.jclepro.2019.04.159.
Sui, Y., et al., 2020. Mining urban sustainable performance: spatio-temporal emission potential changes of urban transit buses in post-COVID-19 future. Appl. Energy 280, 115966. https://doi.org/10.1016/j.apenergy.2020.115966.
Sun, T.T., Huang, Z.F., Zhu, H.D., Huang, Y.H., Zheng, P.J., 2021. Congestion pattern prediction for a busy traffic zone based on the hidden Markov model. IEEE Access 9, 2390–2400. https://doi.org/10.1109/Access.2020.3047394.
Tan, Z., Shao, S., Zhang, X., Shang, W.L., 2023. Sustainable urban mobility: flexible bus service network design in the post-pandemic era. Sustain. Cities Soc. 97, 104702 https://doi.org/10.1016/j.scs.2023.104702.
Tang, W.M., Yiu, K.F.C., Chan, K.Y., Zhang, K., 2023. Conjoining congestion speed-cycle patterns and deep learning neural network for short-term traffic speed forecasting. Appl. Soft Comput. 138, 110154 https://doi.org/10.1016/j.asoc.2023.110154.
Tozer, L., Bulkeley, H., van der Jagt, A., Toxopeus, H., Xie, L.J., Runhaar, H., 2022. Catalyzing sustainability pathways: navigating urban nature based solutions in Europe. Glob. Environ. Change 74, 102521. https://doi.org/10.1016/j.gloenvcha.2022.102521.
Vallamsundar, S., Lin, J., 2011. MOVES versus MOBILE comparison of greenhouse gas and criterion pollutant emissions. Transp. Res. Rec. 2233, 27–35. https://doi.org/10.3141/2233-04.
Wang, Y., Bian, Y., Zhang, X., 2020. Determination and analysis of the characteristics of theRecurrent congestion points of Beijing. J. Transp. Eng. 20 (06), 57–61+68. https://doi.org/10.13986/j.cnki.jote.2020.06.010.
Wei, X., Ren, Y., Shen, L., Shu, T., 2022. Exploring the spatiotemporal pattern of traffic congestion performance of large cities in China: a real-time data based investigation. Environ. Impact Assess. Rev. 95, 106808 https://doi.org/10.1016/j.eiar.2022.106808.
Xin, Y.X., 2023. The Management Model of Bike Sharing System. Journal of Soft Computing and Decision Analytics 1 (1), 209–218. https://doi.org/10.31181/jscda1120239.
Xu, W., Huang, Y., 2020. Mining urban congestion evolution characteristics based on taxi GPS trajectories. Am. J. Traffic Transp. Eng. 5 (1), 1–7. https://doi.org/10.11648/j.ajtte.20200501.11.
Xue, H., Jiang, S., Liang, B., 2013. A study on the model of traffic flow and vehicle exhaust emission. Math. Probl. Eng. 2013 https://doi.org/10.1155/2013/736285.
Yan, C.D., Wei, X.B., Liu, X., Liu, Z.G., Guo, J.X., Li, Z.W., 2020. A new method for real-time evaluation of urban traffic congestion: a case study in Xi'an, China. Geocarto Int. 35 (10), 1033–1048. https://doi.org/10.1080/10106049.2018.1552325.
Yang, S.X., Ji, Y., Zhang, D., Fu, J., 2019. Equilibrium between road traffic congestion and low-carbon economy: a case study from Beijing, China. Sustainability 11 (1). https://doi.org/10.3390/su11010219.
Yao, T., Chen, L., Luo, Y., Zhou, Y., Yang, L., 2022a. Research on the measurement of transportation carbon emission of typical road sections in Nanjing City. J. Green Sci. Technol. 24 (12) https://doi.org/10.16663/j.cnki.lskj.2022.12.059, 12-17+21.
Yao, X., Fan, Y., Zhao, F., Ma, S.-C., 2022b. Economic and climate benefits of vehicle-to-grid for low-carbon transitions of power systems: a case study of China's 2030 renewable energy target. J. Clean. Prod. 330, 129833 https://doi.org/10.1016/j.jclepro.2021.129833.
Yu, Q., Li, W., 2022. Transportation Spatio-Temporal Big Data Analysis, Mining and Visualization (Python Version). Tsinghua University Press.
Yu, Q., Yuan, J., 2022. TransBigData: A Python package for transportation spatio-temporal big data processing, analysis and visualization. J. Open Source Softw. 7 (71), 4021. https://doi.org/10.21105/joss.04021.
Zeng, G., et al., 2020. Multiple metastable network states in urban traffic. Proc. Natl. Acad. Sci. U. S. A. 117 (30), 17528–17534. https://doi.org/10.1073/pnas.1907493117.
Zeng, A., et al., 2022. Battery technology and recycling alone will not save the electric mobility transition from future cobalt shortages. Nat. Commun. 13 (1), 1341. https://doi.org/10.1038/s41467-022-29022-z.
Zhang, D., Zhao, J., Zhang, F., He, T., 2015. UrbanCPS. In: Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems. https://doi.org/10.1145/2735960.2735985.

Zhang, Z.H., Wang, Y.P., Chen, P., He, Z.B., Yu, G.Z., 2017. Probe data-driven travel time forecasting for urban expressways by matching similar spatiotemporal traffic patterns. Transp. Res. Part C 85, 476–493. https://doi.org/10.1016/j.trc.2017.10.010.

Zhang, L., Zeng, G., Li, D., Huang, H.J., Stanley, H.E., Havlin, S., 2019. Scale-free resilience of real traffic jams. Proc. Natl. Acad. Sci. U. S. A. 116 (18), 8673–8678. https://doi.org/10.1073/pnas.1814982116.

Zhao, C., Liao, F.X., Li, X.H., Du, Y.C., 2021. Macroscopic modeling and dynamic control of on-street cruising-for-parking of autonomous vehicles in a multi-region urban road network. Transp. Res. Part C 128. https://doi.org/10.1016/j.trc.2021.103176.

Zhao, C., Cao, J., Zhang, X.Y., Du, Y.C., 2022. From search-for-parking to dispatch-for-parking in an era of connected and automated vehicles: a macroscopic approach. J. Transp. Eng. Part A 148 (2). https://doi.org/10.1061/Jtepbs.0000640.

Zhou, Z.H., et al., 2019. Emission characteristics and high-resolution spatial and temporal distribution of pollutants from motor vehicles in Chengdu, China. Atmos. Pollut. Res. 10 (3), 749–758. https://doi.org/10.1016/j.apr.2018.12.002.

Zhu, Q., Liu, Y., Zeng, W., Xu, W.-J., Huang, M., 2011. Research on distribution characteristics of motor vehicle exhaust emissions based on GPS floating car method. Res. Environ. Sci. 24 (10), 1097–1103. https://doi.org/10.13198/j.res.2011.10.19.zhuqr.014.