Journal of Cleaner Production 142 (2017) 926–935

Contents lists available at ScienceDirect

Journal of Cleaner Production

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

Analysis on spatial-temporal features of taxis' emissions from big data informed travel patterns: a case of Shanghai, China

Xiao Luo [a, b], Liang Dong [c, d, *], Yi Dou [d, e], Ning Zhang [f, g, **], Jingzheng Ren [h], Ye Li [i], Lu Sun [d, j], Shengyong Yao [k]

[a] Tongji University, No. 1239, Siping Road, Shanghai, PR China
[b] Shanghai Tongji Urban Planning & Design Institute, No. 1111, North Zhongshan Road, Shanghai, PR China
[c] Institute of Environmental Sciences, CML, Leiden University, Einsteinweg 2, 2333 CC Leiden, The Netherlands
[d] Center for Social and Environmental Systems Research, National Institute for Environmental Studies (NIES), 16-2 Onogawa, Tsukuba-City, Ibaraki 305-8506, Japan
[e] Graduate School of Environmental Studies, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan
[f] Department of Economics, Jinan University, Guangzhou, Guangdong 510632, PR China
[g] Institute of Resource, Environment and Sustainable Development, Jinan University, Guangzhou, Guangdong 510632, PR China
[h] Centre for Engineering Operations Management, Department of Technology and Innovation, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark
[i] Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Department of Traffic Engineering, Tongji University, 4800 Cao'an Road, Shanghai 201804, PR China
[j] Department of Environment Systems, Graduate School of Frontier Sciences, The University of Tokyo 5-1-5 Kashiwanoha, Kashiwa-shi, Chiba 277-8563, Japan
[k] School of Traffic and Transportation, Shijiazhuang Tiedao University, Shijiazhuang City, PR China

ARTICLE INFO

Article history:
Received 30 November 2015
Received in revised form
23 May 2016
Accepted 25 May 2016
Available online 1 June 2016

Keywords:
GPS
Big data mining
Spatial-temporal emissions distribution
Taxi travel pattern
Shanghai

ABSTRACT

Air pollutions from transportation sector have become a serious urban environmental problem, especially in developing countries with expending urbanization. Cleaner technologies advancement and optimal regulation on the transporting behaviors and related design in infrastructures is critical to address above issue. To understand the spatial and temporal emissions pattern within transportation lays the foundation for design on better infrastructures and guidance on low-carbon transportation behaviors. The feasibility of Global Positioning System (GPS) and emerging big data analysis technique enable the in-depth analysis on this topic, while to date, applications had been rather few. With this circumstance, this paper analyzed the taxi's energy consumption and emissions and their spatial-temporal distribution in Shanghai, one of the most famous mega cities in China, applying big data analysis on GPS data of taxies. Spatial and temporal features of energy consumptions and pollutants emissions were further mapped with geographical information system (GIS). Results highlighted that, spatially, the energy consumption and emission presented a distribution of dual-core cyclic structure, in which, two hubs were identified. One was the city center, the other was Hongqiao transport hub, the activities and emission was more concentrated in the west par of Huangpu River. Temporally, the highest activity and emission moment was 9–10AM, the second peak occurred in 7–8PM, which were both the traffic rush period. The lowest activity/emission moment was 3–4AM. Causal mechanism for such distribution was further investigated, so as to improve the driving behaviors. Through the exploration of spatial and temporal emissions distribution of taxis via big dada technique, this paper provided enlightening insights to policy makers for better understanding on the travel patterns and related environmental implications in Shanghai metropolis, so as to support better planning on infrastructures system, demand side management and the promotion on low-carbon life styles.

* Corresponding author. CML, Leiden University, P.O. Box 9518, 2333 RA Leiden, The Netherlands. Tel.: +31 (071) 527 5608.
** Corresponding author. Institute of Resource, Environment and Sustainable Development, Jinan University, Guangzhou, Guangdong 510632, PR China.
E-mail addresses: dong0926@163.com (L. Dong), zn928@naver.com (N. Zhang).

http://dx.doi.org/10.1016/j.jclepro.2016.05.161
0959-6526/

1. Introduction

With the rapid economic growth and infrastructure development under the process of urbanization and motorization in developing countries, fossil fuel-based transportation demand is increasing quickly (Mitric, 2013; Paravantis and Georgakellos, 2007), which results in significantly increasing contribution to greenhouse gas (GHG) emissions and urban air pollution (Cai and Xu, 2013; Li et al., 2016). It not only significantly affects residents' health, but also decrease the overall amenity of the city (Wang et al., 2015). Particularly to developing countries like China, urban transport is identified as one of the main contributors to the deteriorating urban air pollution (Li et al., 2014; Zhang and Lahr, 2014). Hence, it is quite serious to seek a package of efficient solutions to better regulate the car usage for reducing the emissions (Hao et al., 2014; Tao et al., 2014; Yan et al., 2015).

With this circumstance, China's government spare no efforts to improve the urban transports system so as to significantly reduce the air pollutants and GHGs emissions (An et al., 2011; Cai et al., 2014; Geng et al., 2013). The main hardware efforts include the enhancement of fuels quality (An et al., 2011; Hao et al., 2006, 2011) and technology improvement on the engines (Peng et al., 2016), encouragement of electric vehicles and natural gas vehicles (Cai et al., 2014; Cai and Xu, 2013), as well as urban public transports promotion (Geng et al., 2013; Wang and Liu, 2014). The software efforts included regulation on driving (e.g. the odd and even number lines regulation in Beijing) (Wang and Liu, 2014; Wang et al., 2014), improvement on infrastructures (Asamer et al., 2016; Tu et al., 2016; Yang et al., 2015), as well as encouragement on low-carbon transports behaviors (Geng et al., 2013; Peng et al., 2015; Wang et al., 2015; Willoughby, 2013). To support the above various implications, to better understand the travel patterns and the related emissions features is critical (Tang et al., 2016), especially for the design and the improvement of infrastructures, as well as low-carbon behaviors transformation (Asamer et al., 2016; Cai and Xu, 2013; Gao and Kitirattragarn, 2008).

The emerging ICT technologies (Information and Communication Technology), development of geographical information system (GIS) and GPS technique, as well as big data mining approach (Lécué et al., 2014; Re Calegari et al., 2016; Togawa et al., 2016), lays a foundation on in-depth investigation on the travel patterns of vehicles and the analysis on related emissions features (Cai et al., 2014; Tu et al., 2016). With data recorded from GPS terminals as a large amount of real-time GPS trajectory data, big data mining technique can help to analyze the driver's behavior from the trajectories that links to the estimation and influencing mechanism of vehicle emissions (Liu et al., 2009). In this way, causal mechanism between spatial and temporal emissions features and travel patterns can be investigated, so as to support further implications on a series of infrastructures design and improvement, as well as travel behaviors optimization. However, up to date, relative studies had been rather few. With the feasibility of installation of GPS and data collection, urban taxis were the main source of researches.

The enlightening studies included but not limited to: Wang et al. (2008) measured the on-road vehicles' emissions distribution with the help of GPS data. In this study, vehicles emissions' inventory was further developed for Shanghai. Results showed that about 20% of the total emissions were emitted during the cold start period and more than 50% of the total emissions were emitted on the arterial roads. Also, light-duty vehicles were the main source of CO emissions. However, travel patterns were not in-depth analyzed. Liu et al. (2012) analyzed taxis' fuel use pattern was and the influencing factors of taxis' fuel choice by conducting a survey of taxi drivers in Nanjing City. The analytical results highlighted impacts of various personal characteristics on the fuel consumptions and driving patterns, such as how old age, long experience, and high education affected the adoption of natural gas taxis to substitute traditional fossil fuel based taxis. Their findings provided primary evidence of how personal characteristics and behaviors affected fuel consumption (and related emissions), but big data techniques were not fully applied. Cai et al. (2014), Cai and Xu (2013) investigated real-time vehicle trajectory data for more than ten thousands taxis in Beijing, applied with “big data” mining techniques, so as to characterize the travel patterns of taxis. With this as basis, they further analyzed the on life cycle GHGs emissions of projection of plug-in hybrid electric vehicles substituting fossil fuel based taxies. The characterized individual travel patterns analysis provided critical insights on understanding the driving behaviors and related design on plug infrastructures in Beijing. Tu et al. (2016) applied a spatial–temporal demand coverage approach to optimize the location design of electric vehicles charging stations. The real time GPS data from taxis fleet provided basic information for the model to analyze the spatial and temporal patterns of taxis in Shenzhen, one of the most developed cities in China. However, the spatial and temporal fuel consumption and emissions patterns were not addressed.

Existing studies verified the feasibility and advancement of GPS data and related big data mining technique on the analysis of urban travel patterns and the related environmental impacts. Based on the above, this paper applies big data mining technique to examine a massive data set containing real time trajectories of about 13,600 taxis in Shanghai for one month, in total more than four billion data, to explore the travel patterns and related spatial and temporal features of emissions. Visualization of emissions patterns is mapping with GIS. Given that the data set represents more than 20% of Shanghai's taxi fleet, our results will provide critical insights for a series of follow-up social, economic and environmental implications of urban taxis management, such as adoption of electric vehicles, optimal location of infrastructures, which are hot spots of Chinese government's promotions on mega cities. To our best knowledge, this study represents the first of a series of researches exploring the role of big data in spatial and temporal analysis on urban taxi emissions.

The remainder of the paper is organized as follows: after this introduction section, Section 2 describes the materials and methods, including the data set of GPS trajectories and the model for the real-time emissions analysis and distribution mapping; Section 3 introduces the case area; Section 4 presents the results, discussions and related implications; and finally, Section 5 draws the conclusions.

2. Materials and methods

2.1. Data

This study aims at conducting a generalized process to make use of vehicle GPS data for analyzing emission distribution from urban transportation sector. At the first step, we apply the method to a general case in Shanghai based on a full set of Taxi GPS data. Generally, taxi service is related with the actual travel demand, not only the service strength is proportional to travel demand, but also its distribution is corresponding to travel demand distribution. In addition, taxi GPS data is likely to be shared as an open data in the future, so that this paper could support a basic analysis method for its application. The taxis GPS data chosen in this study is supported by Shanghai Qiangsheng Holding Co., Ltd., which is the largest taxi company in Shanghai and occupies around 30% of all the taxis in the city. The GPS data set contains approximate 140 million records per day, which belongs to 13,675 taxis, covering more than 20% of the whole taxi fleet in Shanghai. Among the one month data series, we extracted the GPS record on April 1st of 2015 (typical workday). The mechanism of the GPS device installed in the taxis is to locate and send the real-time longitude and latitude information back to the receiver server every some seconds, while the ID of the taxi, vacant statue, time stamp of GPS measuring, driving speed, and the driving direction at that time are recorded in a table. In this way, it is able to provide critical spatial and temporal information to analyze the travel patterns. A sample record is shown in Table 1 to present the data structure.

2.2. Accounting approach for fuel consumption and emission

A taxi's trajectory consists of many short distance segments that we can estimate the distance and average speed for each segment, following the calculation of fuel consumption and emission.

Firstly, fuel consumption rate on each segment could be estimated due to its average driving speed. Then, by an integral equation, the total fuel consumption would be summed for each trajectory through which emissions could be estimated by multiplying the emission factor with fuel consumption. Finally, to present the spatial distribution, the emissions of all the trajectories would be summed and divided into a set of meshes with resolution of 100 m.

$$q_{i,j} = \left(a + cv_{i,j} + ev_{i,j}^2\right) \Big/ \left(1 + bv_{i,j} + dv_{i,j}^2\right) \quad (1)$$

$$Q = \sum_{i,j} q_{i,j} \quad (2)$$

$$E_k = Q \cdot e_k \quad (3)$$

where $q_{i,j}$ is the fuel consumption at segment i of taxi j, and $v_{i,j}$ is the average driving speed at segment i of taxi j. Following the total fuel consumption Q is the sum of $q_{i,j}$, and the emission of type k is the product of Q and emission factor of type k written as e_k. Referring to the COPERT model (Nikoleris et al., 2011), the parameters of fuel efficiency joint with emission factors due to driving speed are summarized in Table 2.

2.3. Mapping technique for spatial distribution of emissions

For summarizing and presenting the distribution of emissions, it is necessary to integrate the result with geographic information. Since GPS device records coordinate every time, it is possible to transfer the points into segments with coordinates by ArcGIS software. Thus it will return a set of trajectories with emissions data in the attribute table. Corresponding to a set of 100 m meshes, all these lines are intersected into many segments within each mesh. Finally, all the segments of emissions value are summed by meshes, which is shown in Fig. 1, the segment AD will be intersected into AB, BC and CD, so that the travel distance, fuel consumption and emissions value would be corresponded to mesh3, mesh1 and mesh2, respectively. Consequently, the distribution of emissions could be presented as mesh data.

3. Case description

As one of the two most developed mega-cities in mainland of China (Beijing and Shanghai), Shanghai is a typical dense and productive city that is suitable as the case area of this study (Dou et al., 2016). As shown in Fig. 2, it locates in the Yangtze River Delta, which is the wealthiest commercial region in China. On the one hand, Shanghai is a very dense city and has a distinct core-periphery structure. Due to the Statistical Yearbook of Shanghai in 2015, the total municipal administrative area of Shanghai is 6340 km^2, where in total with 24.26 million permanent population by 2014. Most of the residents are concentrating in the city center within the outer ring expressway, mainly including Huangpu, Jing'an, Changning, Xuhui, Putuo, Zhabei, Yangpu and north part of Pudong New District. The peak population density even reached 90,000 person/km^2 in the CBD (Central Business District) around the People's Square. On the other hand, Shanghai is also one of the most important engines of China's economy that always keeps a rapid economic growth. During the period from 2000 to 2014, Shanghai's GDP increased from 477 billion CNY to 2357 billion CNY that maintains an annual growth rate by 12%. By 2014, the tertiary industry possesses the majority of GDP by 64.8%, following the secondary industry shares 34.7%.

As a consequence of dense land use and rapid economic growth, residents' traffic demand is increasing fast which uncontrollably leads to serious traffic jam and vehicle emissions. According to the 3rd and 4th Comprehensive Travel Survey of Shanghai, the total trip number generated by urban passenger transportation has approached 50 million trips per day by 2009 that means daily one person generates 2.43 trips. At the same time, the number of individual civil motor vehicles in Shanghai increased from 0.85 million to 2.25 million during the period from 2002 to 2014. Not only the average driving speed in the city decreased, but also the air quality in the city became worse due to the vehicle emissions. According to the modal split of urban passenger transportation in Shanghai listed in Table 3, referring the Comprehensive Travel Survey of the year 1986, 1995, 2004, and 2009. Although railway system shares more and more proportion of passenger traffic, the share of small cars and taxi increases faster from almost 0%–12% and 6%, respectively, that reveals the severity of controlling the usage of private cars and taxi for emissions reduction.

Based on above, to improve the cars regulation, promote better land use planning, and optimize the design on related infrastructures is critical to transform Shanghai into low-carbon city (Dou et al., 2016; Jiang et al., 2016). The China's central and Shanghai's governments has launched a series policies implications to promote the projection of public transport, adaptation of electric

Table 1
A sample record of taxis GPS information.

Taxi ID	Vacant stature	Receive time of GPS	Measuring time of GPS	Longitude	Latitude	Speed	Direction
2	1	0:59:01	0:58:55	121.428555	31.27356	8	109
2	1	0:59:51	0:59:45	121.428548	31.273563	10	109
2	1	0:59:20	0:59:15	121.428555	31.27356	14	109
2	1	0:59:40	0:59:35	121.428555	31.27356	18	109
2	1	0:59:30	0:59:25	121.428555	31.27356	20	109
2	1	0:59:11	0:59:05	121.428555	31.27356	22	109
4	1	0:59:29	0:59:24	121.396163	31.233547	0	350
4	1	0:59:09	0:59:04	121.396163	31.233547	0	350
4	1	0:59:40	0:59:34	121.396163	31.233547	0	350

Table 2
Parameters for estimating fuel consumption and emissions.

	a	b	c	d	e
CO	71.7	35.4	11.4	−0.248	0
HC	5.57×10^{-2}	3.65×10^{-2}	-1.1×10^{-3}	-1.88×10^{-4}	1.25×10^{-5}
NOx	9.29×10^{-2}	-1.22×10^{-2}	-1.49×10^{-3}	3.97×10^{-5}	6.53×10^{-6}
Fuel consumption	217	9.6×10^{-2}	0.253	-4.12×10^{-4}	9.65×10^{-3}

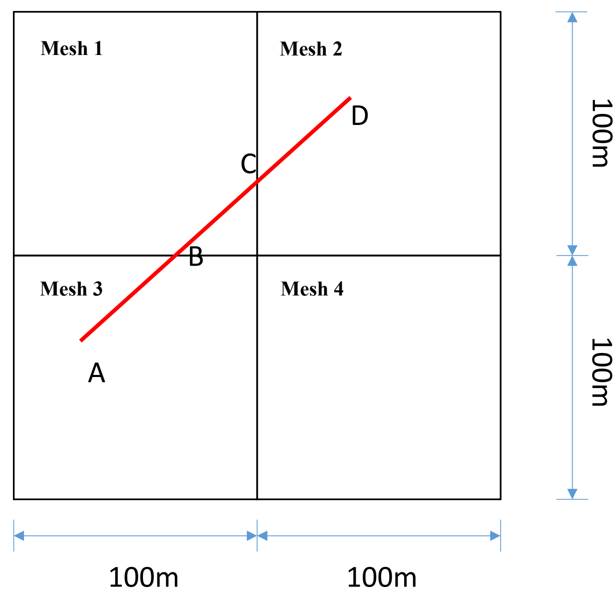

Fig. 1. Summation of segment value into meshes.

Table 3
Modal split of urban passenger transportation in Shanghai.

Year	Rail	Bus	Taxi	Small-car	Large car	Motor cycle	Bike	Walk
1986	0%	24%	0%	1%	0%	0%	32%	42%
1995	1%	18%	2%	4%	2%	2%	42%	30%
2004	3%	18%	7%	9%	3%	6%	29%	26%
2009	8%	16%	6%	12%	2%	4%	27%	24%

vehicles, construction of EVs charging stations. Acted as a key role in urban transport, improvement on taxis fleet in Shanghai will be critical. This paper will represent a first practice on the researches on taxis' travel behaviors and analysis on spatial and temporal emissions patterns in Shanghai, by exploring the role of big data and spatial mapping technique.

4. Result and discussion

4.1. Spatial and temporal characters of travel distance

As a key parameter for travel patterns, travel distance is firstly analyzed. Result is illustrated in Fig. 3. From the time series perspective, calculating from the taxi GPS data supported by Qiangsheng Company of Shanghai, the total accumulated travel distance achieves 6.4 million kilometers per day by 13,675 taxis, so that the average travel distance of one taxi is estimated around 470 km. Apparently, the travel demand varies a lot during a day. It reaches the peak value around 10AM and 8PM, which represents the rush hour in the morning period and evening period. The lowest value appears at 4AM, then immediately increases to the morning peak time. Interestingly, it decreases a lot from 4PM to

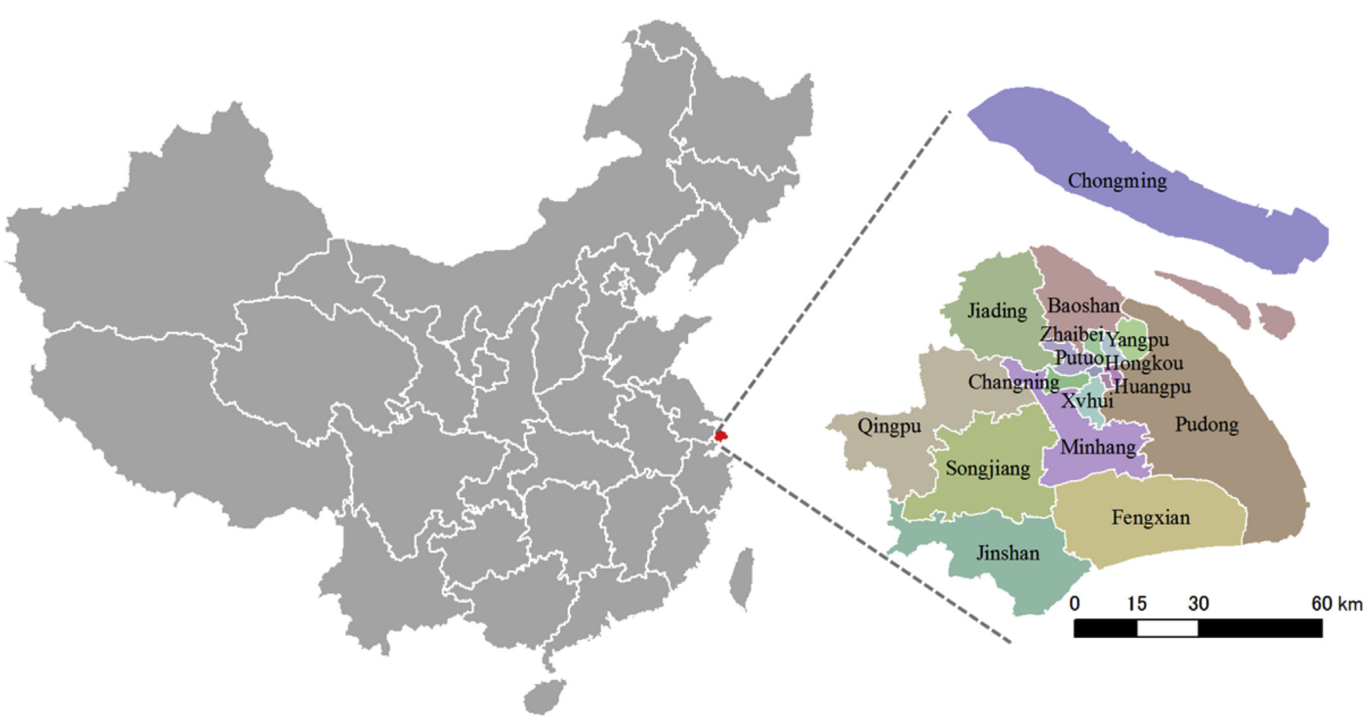

Fig. 2. Geographic information of Shanghai City.

6PM when it should be the actual peak time of traffic demand due to the Urban Comprehensive Traffic Survey of Shanghai. The main reason could be the serious congestion in city center, which decreases the average travel speed significantly. In addition it is also the timing that many of the taxi drivers shift the work to the next worker. Therefore, even though the traffic demand (trip generation) is very high, the actual travel distance differentiated in this period is not so high. After 8PM, the travel distance decreases rapidly due to the citizens' daily schedule, but immediately experiences a peak at 12PM when the service sectors end their business and citizens close their leisure activities in the night.

From spatial perspective, obvious various travel patterns are presented with different time spans. Results are illustrated in Fig. 4, in which, relative values of each spatial meshes compared to the maximum calculation are presented. In this way, the spatial distribution of travel distance is visualized. It is highlighted that, starting from the lowest value at 3AM, citizens' travel activities begin from the southeast part of the city along the metro line, and then quickly expand to the west direction of Shanghai. There appears a hot spot where travel demand is concentrating in the end of the metro line. In fact, the hot spot at west area is just the Hongqiao Airport Business Area, in which the city's high-speed railway station is also located. The other hot spot in eastern area is the People's Square, which locates within the Central Business District (CBD) of Shanghai. In the peak hour (9AM–10AM), the increasing traffic demand quickly expands from Hongqiao Airport to the CBD that seems to be a corridor connecting the west to the city center, even extends to the east part of Pudong District. Followed by the period 10AM–12AM, the hot spot moves to the Puxi District. From the time period 3PM–4PM, the concentration effect in the Hongqiao Airport area gradually decreases until the airport and railway system end the schedule, and the travel activities finally shrink into the city center.

4.2. Emission estimation and the spatiotemporal characteristics

The above travel distance can present the travel patterns to a large extent. With this as basis, we further analyze the emissions patterns. Results of temporal analysis are highlighted in Fig. 5. It presents that, the total fuel consumption of the taxis in the day (April 1st of 2015) is estimated about 456t, while the total CO, HC and NOx emissions are 81t, 0.3t and 0.9t respectively. CO emission occupies the majority. Being similar with the variation of travel distance, the variation of emissions follows the same trend in a day. Starting from the valley period of 4AM–5AM, the quantity of emissions increases to the peak value at the time period

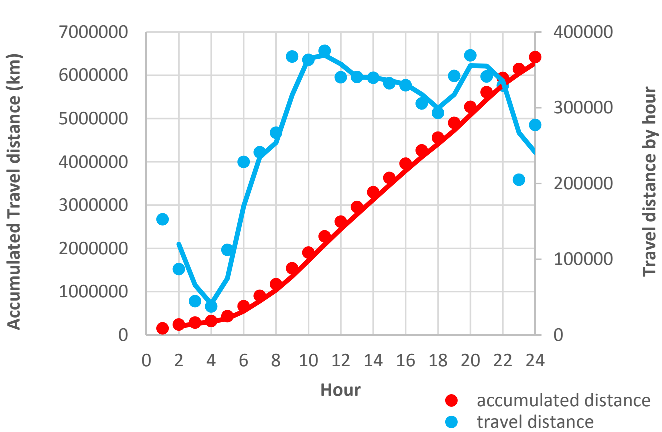

Fig. 3. Accumulated travel distance and travel distance of Taxi by hour.

9AM–12AM, and begins to decrease until 6PM. Its peak in the evening is around 8PM, then gradually decreases until mid-night.

According to the method provided in the section 3, it is possible to estimate the spatial distribution of travel distance (traffic demand), fuel consumption and emissions using taxi GPS data. The accumulated quantity of them is summarized in Fig. 6, in which, relative values of each spatial meshes compared to the maximum calculation are presented, in terms of emissions in line with the travel distances. As a whole, since the emissions are related to the fuel consumption, which is proportional to the travel distance, all these figures indicate a quite similar circular structure. Obviously, the travel activities are concentrated in the city center that most of them occur within the outer ring, which is basically coinciding with the urban structure of Shanghai. However, two cores appear within the outer ring, of which one in the east area is the planned city center around the People's Square (Point A in Fig. 6), and another one is around the Hongqiao Transport Hub (Point B in Fig. 6) that is the city center in the meaning of regional transportation. A corridor with high value links the two cores that shows the close contact between planned center and regional communicational center. Interestingly, the boundary of high value area stops crossing the Huangpu River to newly developed city center known as Pudong New District, where the travel activities are obstructed by the river. Otherwise, another relatively high value area appears to the east of the city center where another transport hub, namely Pudong International Airport (Point C in Fig. 6), is located. With the increasing traffic demand from Pudong International Airport, it may become a hot spot in the future.

It is highlighted that, since the quantity of fuel consumption and emissions is generally based on the travel activities, the distribution of fuel consumption, CO emission and NOx emission reveals similar to the distribution of travel distance. However, considering the decrement of average travel speed in the city center and Hongqiao Transport Hub area, the fuel consumption and emissions in these two areas are relatively higher but lower along the corridor between them.

Specifically, the more precise sight on time series changes of emissions brings a meaningful understanding on spatiotemporal inequality. Considering the similarity between travel distance, fuel consumption and emissions, here takes the emission of NOx as a typical pollutant, which occupies the majority and is one of the most important sources caused the smog (Fig. 7).

Although the distribution of the accumulated NOx pollution reveals a core-periphery structure in a whole day, the instantaneous distribution of NOx indicates a variable pattern due to the period. From 3:00AM, the relatively highlighted areas of NOx emission firstly appear in the city center with the north and west sub-centers. Then the emissions gradually extend to the whole urban area. At the period of 7AM–8AM, the identified hot spot firstly appears at Hongqiao Transport Hub when many travelers begin their works or flights at the airport. Later the highlight area spreads to the city center. In the day, the core-periphery structure of NOx emission represents a sharp gradient that reveals the travel activities excessively concentrate in the city. This proves that Shanghai is still monocentric, even though the governmental master plan aims to guide the city towards a polycentric structure. The worst situation of emissions happens at the peak hour of commuting when the regular congestion decreases the average travel speed so as to increase the emission factor of the pollutants.

4.3. Causal mechanism of travel patterns, fuel consumption and emissions

According to the equations for estimating fuel consumption and emissions provided in methodology section, two critical factors

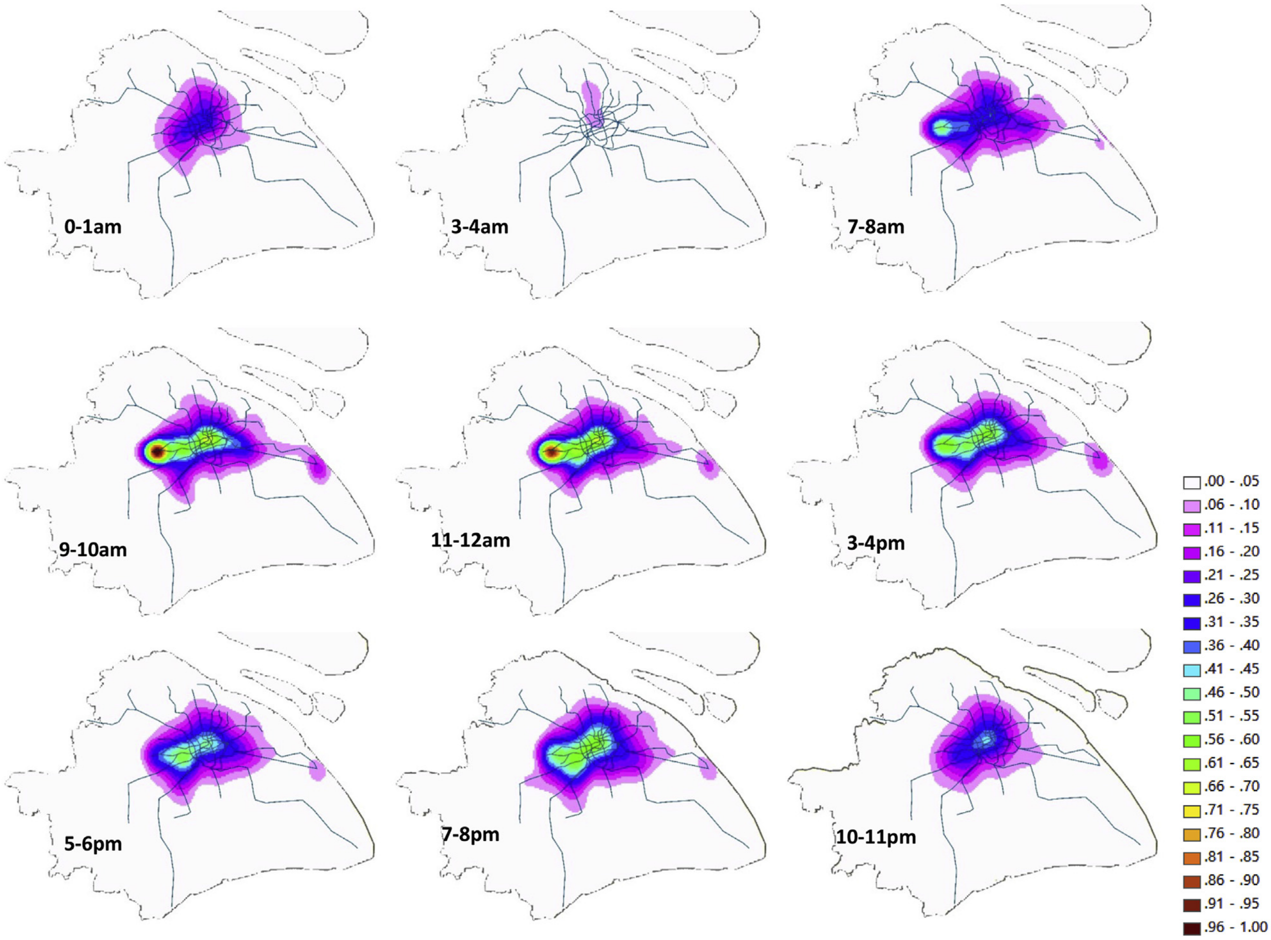

Fig. 4. Travel distance distribution in typical time periods. Note: The value of the labels present the relative values of each spatial meshes to the maximum calculation. In this way, we can easily visualize the spatial distribution of travel distances.

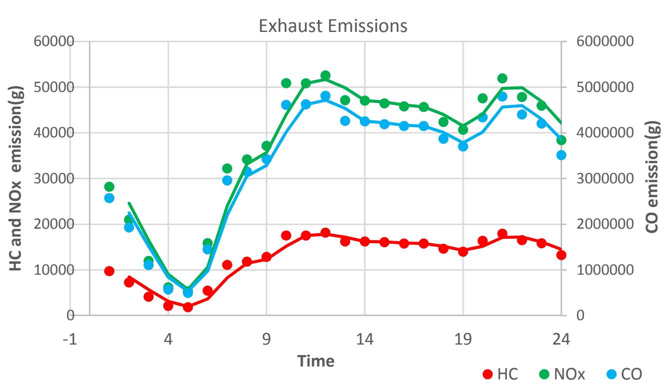

Fig. 5. Emission quantity variation d a day.

could be identified. One is the travel distance, and the other is the emission factor. The former depends on the actual travel activity, while the latter is influenced by real-time travel speed. The relationship between emission factor and travel speed has been shown in Fig. 8.

Setting the fuel consumption and emission factors of the cold start (travel speed ≈ 0) as 100%, when the value of fuel consumption and emission of HC, CO, NOx are 198 g/km, 0.05 g/km, 2.29 g/km, 0.092 g/km, respectively. Commonly, with the increment of travel speed, the fuel consumption and emission factors at first decrease a lot, then gradually increase and overpass the starting value. However, the path of each section indicates quite a difference. The travel speed around 60 km/h is identified as the most efficient speed for reducing the pollutions, while it is of significant importance to avoid low travel speed below 20 km/h. The difference of efficiency between the high and the low value could be five times.

On the other side, the accumulated travel distance in different travel speed is summarized in Fig. 9. The result show that the travel speed varies a lot, fortunately there is an impulse in the speed of 60 km/h, namely the most efficient speed for fuel efficiency, maybe that's because most of the taxi drivers want to control to the speed to the most fuel efficiency speed when the can control. In the other speed the frequency are similar to some degrees, because the road condition and the traffic jam cannot be avoided.

In fact, beside the equations provided before, there is also a critical factor which would affect the fuel efficiency and emissions, known as “load factor” (or effective driving ratio). More passengers in a taxi may increase the total fuel consumption and emissions, but the efficiency (quantity per person) would significantly increase. Due to the Fourth Comprehensive Traffic Survey Report of Shanghai, the average load factor of all taxis is only 1.66 person/trip that indicates a great waste of transport resource. Since currently the data have not recorded the number of passengers, this study simply defines an “effective driving ratio” instead of load factor as the accumulated travel distance of taxis with passengers. As shown in Fig. 10, the overall effective driving ratio of taxis is around 65% as

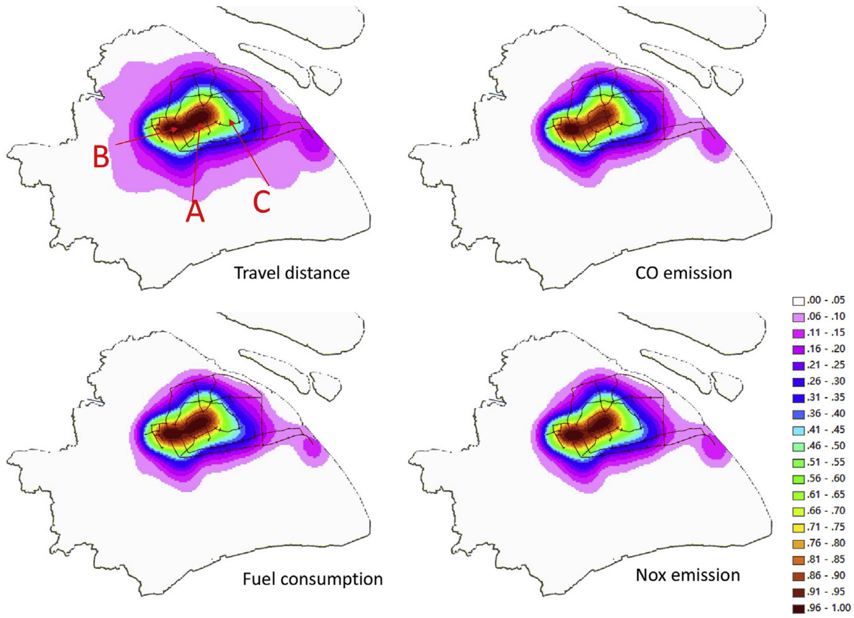

Fig. 6. The distribution of total travel distance, fuel consumption and emissions in one day. Note: The value of the labels present the relative values of each spatial meshes to the maximum calculation. In this way, we can easily visualize the spatial distribution of emissions.

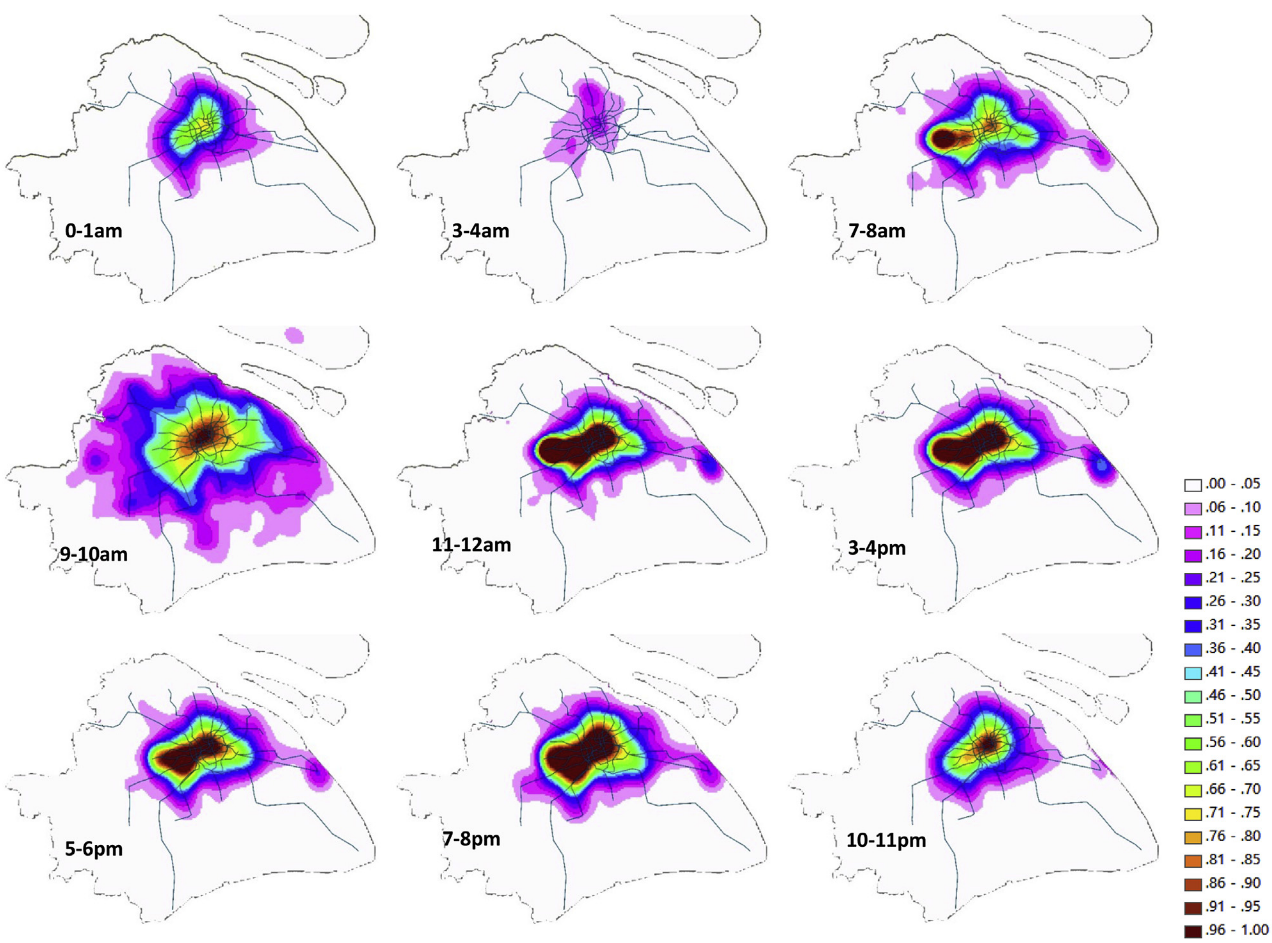

Fig. 7. NOx emission of Shanghai taxi in different typical time. Note: The value of the labels present the relative values of each spatial meshes to the maximum calculation. In this way, we can easily visualize the spatial distribution of NOx emissions.

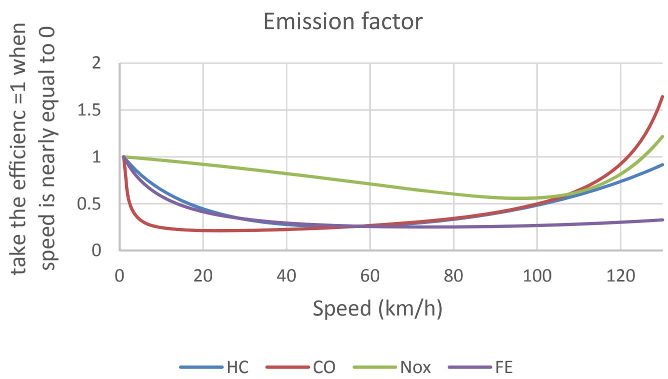

Fig. 8. Changes of the fuel consumption and emission factor by real-time travel speed.

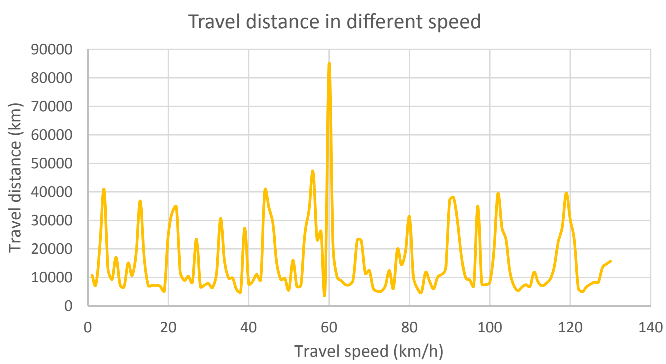

Fig. 9. Accumulated travel distance in different travel speed.

a whole. It meets the low value at 4AM then increases around 70% and keeps stable until mid-night.

However, from the drivers' perspective, the effective driving ratio of each taxi reveals a great variation. As shown in Fig. 11, the variation of the numbers of taxis in different effective driving ratio follows the normal distribution. Some drivers can very efficiently attract passengers and make use of the travel distance, by contract some of them are very inefficient. Driving behavior may be one main cause to the inefficiency so that it is necessary for these drivers to find a better way of supporting service. One feasible solution could be supported by Information Communication Technology (ICT) such as the worldwide used ‘Uber’ or Chinese local application ‘DiDi taxi’, which helps the passengers to search the nearest taxi so as to avoid unnecessary empty drive at the same time.

4.4. Discussion and implications

Based on above spatial and temporal analysis on the travel patterns, fuel consumptions as well as emissions, several critical insights are highlighted.

a. Obvious spatial and temporal patterns of travel behaviors and emissions are verified with the big data mining. One advantage of big data technique is that, based on real time data, it can further investigate the social behaviors. Our results identify that travel patterns have inconvenient impacts on related fuel consumption and emissions. Some travel patterns result from improper design of the infrastructures, such as urban functional unit layout, road condition, traffic regulation, inefficient public transport, and so on. However, some are caused by social behaviors, e.g. refusing behaviors of the taxi drivers, unhealthy competition of driving in the road, as well as group culture, which results in some traffic jam in certain area. Via big data, we can further analyze these social factors in-depth and provides critical insights on how to make ever-improvement of them.
b. Big data mining on travel patterns and spatial and temporal analysis can strongly support the optimal design on urban infrastructures. It was well known that under surging urbanization, China's urban infrastructures planning is not so rational some time. Typical problem is the concentration of administrative and economic resources, which reflect on the urban planning, is the typical concentration effects, and the separation of living functional areas and commercial functional areas, which cause longer travel distance and emissions. Our spatial analysis identified this problem.
c. As a discussion following up above viewpoint, it is vital important to promote transit oriented development (TOD) in Shanghai and other Chinese mega cities. By improved design on urban public transport, taxi reception stations, taxi booking systems and other infrastructures, it is expected that the accessibility of transport can be enhanced and as a result, the city can be compact and the travel pattern will be more environmental efficient.
d. The other important implications of our research is to support better design on infrastructures for adoption of electric vehicles and natural gas based taxis, which are the two main policies implication in Chinese large cities to “green” urban taxis. Fully understanding the travel patterns of taxis lay the foundation for optimal spatial planning on more efficient charging stations and reasonable adoption of electric or natural gas vehicles.
e. Finally, there is still research limitations and future concerns. This paper can be the fundamental analysis, that is to test the feasibility of big data mining approach, mapping technique, and

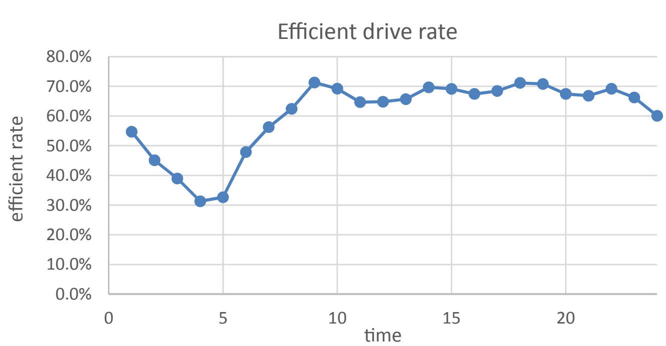

Fig. 10. Effective driving ratio.

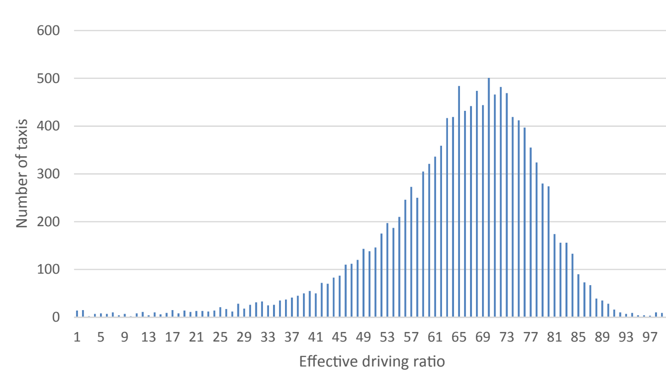

Fig. 11. Number of taxis by effective driving ratio.

to characterize the travel patterns and related emissions features. The analysis is some basic as usual analysis. In future, we need further combine the future urban planning scenarios, to in-depth discuss a series of environmental implications of taxis system, e.g. projection of EVs and natural gas vehicles, and the related design on charging stations and natural gas stations. As well, in methodologies perspective, this paper applies standard accounting approach, not fully applies the life cycle analysis. Combined with scenarios design, in the future, to integrate life cycle assessment, spatial analysis and big data mining, it is expected to support spatial and temporal explicit analysis on infrastructures design and adoption of more environmental friendly vehicles.

5. Conclusions

Applying with big data mining on GPS data set containing one month real time trajectories of 13,675 taxis in Shanghai, this paper explored the travel patterns and related spatial and temporal features of emissions. Visualization of emissions patterns was realized via GIS accordingly. Results verified obvious spatial and temporal disparities. Spatially, the energy consumption and emission presented a distribution of dual-core cyclic structure. Two hubs were identified, including the city center and the Hongqiao transport hub. In general, the activities and emission was more concentrated in the west par of Huangpu River. Temporally, the highest activity and emission moment was 9–10AM, the second peak occurred in 7–8PM, which were both the traffic rush period. The lowest activity/emission moment was 3–4AM. Causal mechanism for such distribution was further investigated in-depth, so as to improve the driving behaviors. The findings of this paper will provide critical insights for a series of future social, economic and environmental implications on urban transport management, including demand side management, decision making on the adoption of electric vehicles, optimal location of infrastructures, as well as guidance on transformation of low-carbon travel patterns.

Based on the analytical results, critical insights for policies implications were discussed in-depth, from the perspectives of identification and guidance of social behaviors; optimal design on urban planning, especially improve the constraints caused by concentration of urban administrative and economic resources; promotion of transit oriented development (TOD), which focusing on the improvement of accessibility; optimal design on transport infrastructures, especially spatial planning on more efficient charging stations and reasonable adoption of electric or natural gas vehicles; as well as future academic concerns on integrating life cycle assessment, spatial analysis and big data mining, to support spatial and temporal explicit analysis on infrastructures design and adoption of more environmental friendly vehicles.

Finally, it is highlighted that, the data set in this paper represents approximate 20% of Shanghai's taxi fleet, as a result, our results will be presentative and able to provide critical insights for a series of follow-up practical social, economic and environmental implications of urban taxis management in Chinese mega cities. The approaches were also expected to extend to future studies on other urban transport systems, e.g. urban public transport, and even the whole urban infrastructures system, with the expecting development of smart city promotion.

Acknowledgments

This research was supported by Shanghai Sailing Program (16YF1410900): Evaluation of urban built-up environment based on mobile phone signaling data; the project of “Smart Industrial Parks (SIPs) in China: towards Joint Design and Institutionalization” (No. 467-14-003). The fourth author was also funded by the National Natural Science Foundation of China (No. 41461118), National Social Science Foundation of China (15ZDA054), the last author was funded by the National Natural Science Foundation of China (NO. 51308358).

References

An, S., Hu, X., Wang, J., 2011. Urban taxis and air pollution: a case study in Harbin, China. J. Transp. Geogr. 19, 960–967.
Asamer, J., Reinthaler, M., Ruthmair, M., Straub, M., Puchinger, J., 2016. Optimizing charging station locations for urban taxi providers. Transp. Res. Part A Policy Pract. 85, 233–246.
Cai, H., Jia, X., Chiu, A.S.F., Hu, X., Xu, M., 2014. Siting public electric vehicle charging stations in Beijing using big-data informed travel patterns of the taxi fleet. Transp. Res. Part D Transp. Environ. 33, 39–46.
Cai, H., Xu, M., 2013. Greenhouse gas implications of fleet electrification based on big data-informed individual travel patterns. Environ. Sci. Technol. 47, 9035–9043.
Dou, Y., Luo, X., Dong, L., Wu, C., Liang, H., Ren, J., 2016. An empirical study on transit-oriented low-carbon urban land use planning: exploratory spatial data analysis (ESDA) on Shanghai, China. Habitat Int. 53, 379–389.
Gao, H.O., Kitirattragarn, V., 2008. Taxi owners' buying preferences of hybrid-electric vehicles and their implications for emissions in New York City. Transp. Res. Part A Policy Pract. 42, 1064–1073.
Geng, Y., Ma, Z., Xue, B., Ren, W., Liu, Z., Fujita, T., 2013. Co-benefit evaluation for urban public transportation sector – a case of Shenyang, China. J. Clean. Prod. 58, 82–91.
Hao, H., Geng, Y., Wang, H., Ouyang, M., 2014. Regional disparity of urban passenger transport associated GHG (greenhouse gas) emissions in China: a review. Energy 68, 783–793.
Hao, H., Wang, H., Ouyang, M., 2011. Fuel conservation and GHG (greenhouse gas) emissions mitigation scenarios for China's passenger vehicle fleet. Energy 36, 6520–6528.
Hao, J., Hu, J., Fu, L., 2006. Controlling vehicular emissions in Beijing during the last decade. Transp. Res. Part A Policy Pract. 40, 639–651.
Jiang, Y., Waley, P., Gonzalez, S., 2016. Shifting land-based coalitions in Shanghai's second hub. Cities 52, 30–38.
Lécué, F., Tallevi-Diotallevi, S., Hayes, J., Tucker, R., Bicer, V., Sbodio, M., Tommasi, P., 2014. Smart traffic analytics in the semantic web with STAR-CITY: scenarios, system and lessons learned in Dublin City. Web Semant. Sci. Serv. Agents World Wide Web 27–28, 26–33.
Li, W., Bai, H., Yin, J., Xu, H., 2016. Life cycle assessment of end-of-life vehicle recycling processes in China—take Corolla taxis for example. J. Clean. Prod. 117, 176–187.
Li, Z., Pan, L., Fu, F., Liu, P., Ma, L., Amorelli, A., 2014. China's regional disparities in energy consumption: an input–output analysis. Energy 78, 426–438.
Liu, B., Chen, C., Zhang, B., Bu, M., Bi, J., Yu, Y., 2012. Fuel use pattern and determinants of taxi drivers' fuel choice in Nanjing, China. J. Clean. Prod. 33, 60–66.
Liu, K., Yamamoto, T., Morikawa, T., 2009. Feasibility of using taxi dispatch system as probes for collecting traffic information. J. Intell. Transp. Syst. 13, 16–27.
Mitric, S., 2013. Urban transport lending by the World Bank: the last decade. Res. Transp. Econ. 40, 19–33.
Nikoleris, T., Gupta, G., Kistler, M., 2011. Detailed estimation of fuel consumption and emissions during aircraft taxi operations at Dallas/Fort Worth International Airport. Transp. Res. Part D Transp. Environ. 16, 302–308.
Paravantis, J.A., Georgakellos, D.A., 2007. Trends in energy consumption and carbon dioxide emissions of passenger cars and buses. Technol. Forecast. Soc. Change 74, 682–707.
Peng, B.-B., Fan, Y., Xu, J.-H., 2016. Integrated assessment of energy efficiency technologies and CO_2 abatement cost curves in China's road passenger car sector. Energy Convers. Manag. 109, 195–212.
Peng, B., Du, H., Ma, S., Fan, Y., Broadstock, D.C., 2015. Urban passenger transport energy saving and emission reduction potential: a case study for Tianjin, China. Energy Convers. Manag. 102, 4–16.
Re Calegari, G., Celino, I., Peroni, D., 2016. City data dating: emerging affinities between diverse urban datasets. Inf. Syst. 57, 223–240.
Tang, T.-Q., Xu, K.-W., Yang, S.-C., Ding, C., 2016. Impacts of SOC on car-following behavior and travel time in the heterogeneous traffic system. Phys. A Stat. Mech. Appl. 441, 221–229.
Tao, M., Chen, L., Wang, Z., Ma, P., Tao, J., Jia, S., 2014. A study of urban pollution and haze clouds over northern China during the dusty season based on satellite and surface observations. Atmos. Environ. 82, 183–192.
Togawa, T., Fujita, T., Dong, L., Ohnishi, S., Fujii, M., 2016. Integrating GIS databases and ICT applications for the design of energy circulation systems. J. Clean. Prod. 114, 224–232.
Tu, W., Li, Q., Fang, Z., Shaw, S.-l., Zhou, B., Chang, X., 2016. Optimizing the locations of electric taxi charging stations: a spatial–temporal demand coverage approach. Trans. Res. Part C Emerg. Technol. 65, 172–189.
Wang, H., Chen, C., Huang, C., Fu, L., 2008. On-road vehicle emission inventory and its uncertainty analysis for Shanghai, China. Sci. Total Environ. 398, 60–67.
Wang, J.J., Liu, Q., 2014. Understanding the parking supply mechanism in China: a

case study of Shenzhen. J. Transp. Geogr. 40, 77–88.
Wang, Q., Feng, X., Liu, Y., Wang, X., Zhang, H., 2014. Urban travel mode split optimization based on travel costs. Procedia Soc. Behav. Sci. 138, 706–714.
Wang, Z., Chen, F., Fujiyama, T., 2015. Carbon emission from urban passenger transportation in Beijing. Transp. Res. Part D Transp. Environ. 41, 217–227.
Willoughby, C., 2013. How much can public private partnership really do for urban transport in developing countries? Res. Transp. Econ. 40, 34–55.
Yan, Y., Sun, Y.B., Weiss, D., Liang, L.J., Chen, H.Y., 2015. Polluted dust derived from long-range transport as a major end member of urban aerosols and its implication of non-point pollution in northern China. Sci. Total Environ. 506–507, 538–545.
Yang, J., Chen, J., Zhou, M., Wang, Z., 2015. Major issues for biking revival in urban China. Habitat Int. 47, 176–182.
Zhang, H., Lahr, M.L., 2014. China's energy consumption change from 1987 to 2007: a multi-regional structural decomposition analysis. Energy Policy 67, 682–693.