Understanding space-time patterns of vehicular emission flows in urban areas using geospatial technique

Zihan Kan[a,b], Man Sing Wong[a,c,*], Rui Zhu[d]

[a] *Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, 181 Chatham Road South, Kowloon, Hong Kong*
[b] *State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan, Hubei 430079, China*
[c] *Research Institute for Sustainable Urban Development, The Hong Kong Polytechnic University, Hong Kong*
[d] *Senseable City Laboratory, Singapore-MIT Alliance for Research and Technology, Singapore*

ARTICLE INFO

ABSTRACT

Traffic-related emissions are well-known factors in urban environment which may have adverse implication on human health. Estimating vehicular emissions in urban areas provides an understanding of the air pollution caused by traffic. However, existing microscopic approaches cannot simulate the traffic flows and emissions for an entire city and most of the macroscopic approaches are usually highly complex and require priori knowledge about vehicles' route options. This study, therefore, proposes a straightforward and robust approach to simulate vehicular flows and estimated transport emissions at a city scale via a deterministic approach and by applying the Cell Transmission Model (CTM) to simplify the modeling of vehicles' route selections. Under a space-time integrated framework, we firstly simulate a time-dependent distribution of urban vehicular flows and then estimate pollutant emissions of Carbon Monoxide (CO), Nitrogen Oxide (NOx) and Violate Organic Compounds (VOC) for traffic flows on weekday and weekend. Finally, the spatiotemporal patterns of traffic flows as well as traffic emissions were visualized and illustrated under a space-time integrated framework. With accuracies of around 67.4% to 70%, the results demonstrated the feasibility of the proposed approach for estimating city-scale traffic flows and emissions from road transport.

1. Introduction

Road transport sector is one of the key contributors for air pollutant emission (EPA, 2014). According to the Hong Kong Environmental Protection Department (2016), 54% of total Carbon Monoxide (CO) emissions were released by the road transport sector in Hong Kong. With the adverse impact of air pollution on human health, the analysis of the traffic emissions provides insights about the spatio-temporal patterns and underlying process of transportation-related emissions. However, existing microscopic approaches analyze vehicular emissions using sampled trajectories of vehicles, which have not sufficiently considered the volumes of traffic flows of a city. While most of the macroscopic approaches can simulate traffic flows at a city scale, they are usually highly complex and require priori knowledge about vehicles' route options. This study, therefore proposes a straightforward and robust approach to simulate vehicular flows and estimate transport emissions at a city scale, through incorporating a deterministic approach and using the Cell Transmission Model (CTM).

Traditionally, information of air pollution is acquired and collected from discrete monitoring stations or collected through large-scale fuel-used survey data (Cai & Xie, 2007), which have several limitations in air pollution monitoring in a city. Firstly, the air pollution monitoring stations are spatially-restricted in which there is a lack of holistic view of atmospheric conditions throughout a city from station-data (Nyhan et al., 2016). Secondly, large-scale fuel-used data are often collected at a scale of a city or even a country, leading to a rough estimation of total emissions that vehicles might release. In addition, since the survey data has relatively low temporal resolution, the spatio-dynamic of emissions distribution is difficult to predict from survey data. In the past decades, environmental agencies in the world have developed numerous emission/fuel consumption estimation models for different types of vehicles and fuel, including the COPERT model developed by European Environment Agency (EEA), the MOBILE model and MOVES model of U.S. Environmental Protection Agency (EPA), EMFAC model released by California Air Resources Board, and CMEM and IVE models developed by the University of California at Riverside (Abo-Qudais & Abuqdais,

2005; Barth et al., 2000; CARB, 2006; EPA, 2009; Ntziachristos et al., 2000; Rakha, Ahn, & Trani, 2003; Sharma & Khare, 2001). The development of these emission models enables the estimation of emissions at various levels, i.e., street level, city level and country level. In order to quantify the emissions released by different types of vehicles operating on road networks with various traffic conditions, parameters regarding vehicle technology and movement in these models have been acquired from fixed sensors including video cameras (Yang, Boriboonsomsin, & Barth, 2011) and loop detectors (Chang et al., 2013), as well as large-scale statistics (Burón, López, Aparicio, Martín, & García, 2004) from past studies.

In addition to data collected from fixed sensors or archive statistics, vehicular trajectories have been widely applied in emission estimation in the past few years especially after the enhancement of smart phones and GPS data. Previous studies have explored both nation-wide and city-wide emission inventories based on GPS trajectory data and the emission models (Kan et al., 2018; Kan, Tang, Kwan, & Zhang, 2018; Luo et al., 2017; Nyhan et al., 2016; Shang, Zheng, Tong, et al., 2014; Sun, Hao, Ban, & Yang, 2015; Yang et al., 2011; Zhao, Kwan, & Qin, 2017). Among them, taxi GPS data are especially popular in estimating vehicular emissions thanks to their relatively high accessibility to researchers. However, there are limitations when estimating traffic emissions from taxi GPS trajectories. First, although taxis play an important role in urban public transportation, there is still a small sample of the entire urban vehicle fleet. For instance, taxis only account for 2.5% of the registered vehicles in Hong Kong (Annual traffic census, 2015). With such a small share of traffic flows, the spatial and temporal coverages of taxi GPS trajectories are limited and the dynamics of traffic flow in a whole city is thus hard to be revealed. Second, the behaviors of taxi drivers are often various in spatio-temporal domain. GPS trajectories record behaviors of taxi drivers that would be highly affected by the demands of taxi drivers themselves instead of traffic flow (Zhao, Liu, Kwan, & Shi, 2018), such as taking rests, having meals, refueling, waiting for customers, and picking up/dropping off customers. These behaviors would interfere the detection of the movement of traffic flow in a city. Recently, GPS data acquired by smartphones have also been used to estimate vehicular fuel consumption (Astarita, Guido, Mongelli, & Giofre, 2015; Gately, Hutyra, Peterson, & Wing, 2017). A most recent study used GPS trajectories of car-hailing service from Didi Chuxing Technology Co., a Chinese ride-sharing company, to estimate vehicular emissions (Sun, Zhang, & Shen, 2018). Different from taxis GPS trajectory data, both taxis and private cars could register in the Didi platform and provide ride services to the public. Thus, the trajectories from Didi contain a more mixed fleet than trajectories solely from taxis.

Since the trajectories of all vehicles in a city are difficult to be acquired, most existing studies only estimate a portion of vehicular emissions from the sampled trajectory data. However, understanding traffic emissions of a city requires an accurate number of moving vehicles on each road link during a time period. In the past decades, significant effort has been devoted to simulating traffic flows so as to estimate traffic emissions. Existing traffic flow simulation approaches mainly include microscopic traffic modeling and macroscopic modeling. Microscopic models simulate traffic flow at the level of individuals (Chen & Wu, 2011, Zamith et al., 2015), which is used to examine how individual movement patterns impact traffic flows (Sentoff, Aultman-Hall, & Holmén, 2015). For microscopic models, several software packages have been developed, such as TRANSIMS (Zietsman & Rilett, 2001), INTEGRATION (Rakha & Ahn, 2004) and VISSIM (PTV Planung Transport Verkehr, 2005). The microscopic simulation models and emission models have been integrated in existing studies to estimate traffic emissions. For instance, Amirjamshidi, Mostafa, Misra, and Roorda (2013) simulated accelerations and decelerations of individual vehicle in a driving cycle and estimated the emissions under different moving patterns. Fontes, Pereira, Fernandes, Bandeira, and Coelho (2015) used various traffic micro-simulation tools for assessing the impacts of road traffic on the environment. Abou-

Senna, Radwan, Westerlund, and Cooper (2013) cooperated both microscopic simulation model (VISSIM) and a microscopic emission model (MOVES) for estimating emissions in a section of an interstate highway. Similar approaches have been applied to estimate emissions of parts of road network such as interchanges (Xie, Chowdhury, Bhavsar, & Zhou, 2012), intersections (Jie, Van Zuylen, Chen, Viti, & Wilmink, 2013) and roundabouts (Quaassdorff et al., 2016).

In contrast to microscopic approaches, macroscopic models estimate traffic flows using aggregated parameters such as flow density and average speed (Delis, Nikolos, & Papageorgiou, 2015; Spiliopoulou, Kontorinaki, Papageorgiou, & Kopelias, 2014; Zhu, Wong, Guilbert, & Chan, 2017). Using data collected from traffic sensors such as loop detectors and traffic counting stations, macroscopic models assign traffic flows to road networks. In contrast to microscopic models which estimate individual travel behaviors, macroscopic models simulate traffic flow patterns through assigning traffic flows to road networks. Numerous Dynamic Traffic Assignment (DTA) models have been developed based on formulating principles of vehicles' travel options. The DTA models assign traffic flows to road networks by assuming that vehicles' route choices are either stochastic (Parry & Hazelton, 2012) or deterministic (Siripirote, Sumalee, Ho, & Lam, 2015), which has been demonstrated that it can capture more realistic characteristics of traffic flow (Wang et al., 2018). While prior knowledge of vehicles' route choice is not necessary in stochastic approaches, the searching capability of the stochastic approaches is relatively weak and time consuming. Deterministic approaches simulate vehicles' route choices based on assumptions of vehicle's demands and preferences, which requires the knowledge of designated paths for vehicles. However, in the case of assigning all the traffic flows to road network in a city, it is necessary to work with hundreds of thousands of vehicles. For instance, there were 728,263 vehicles licensed in Hong Kong at the end of 2015 and the number has been increasing for the past years. As a result, the designated paths for all vehicles can hardly be estimated at the same time. Some studies have simplified the problem with specifications of traffic conditions, destinations and vehicular routes (Friesz, Bernstein, Suo, & Tobin, 2001; Javani, Babazadeh, & Ceder, 2018; Zheng & Chiu, 2011), which however, cannot adequately reflect vehicle behaviors on road networks (Xia & Shao, 2005).

In order to tackle this problem, this study incorporates a deterministic approach and assumes that vehicles take the shortest paths to reach their destinations in a road network and adopted the Cell Transmission Model (CTM) to simplify the modeling of vehicles' route selections based on the Lighthill-Whitham-Richards (LWR) method (Lighthill & Whitham, 1955; Richards, 1956). The CTM was first proposed by Daganzo (1994), which is the discrete analogue of the hydrodynamic flow-density differential equations. The author further extended the model by introducing three-legged junctions which can represent more complex networks (Daganzo, 1995). The CTM can capture the traffic revolution by solving LWR and provide relatively realistic details about the queue formation, propagation and dissipation of congestions through kinematic waves (Nie & Zhang, 2008). It is one of the most comprehensive models among the existing macroscopic traffic models (Chuo et al., 2016), which has been widely used in modeling and resolving dynamic network problems (Lo & Szeto, 2002; Xie & Duthie, 2015; Ziliaskopoulos, 2000). However, since the CTM needs to solve equilibrium and optimization equations of high computation complexities in order to simulate the propagation of traffic flows, most of the existing studies have applied it on a relatively small area, such as an intersection (Canudas-de-Wit and Ferrara, 2018), a road (Chuo et al., 2016), or a small part of road networks (Islam, Vu, Panda, Hoang, & Ngoduy, 2017; Levin & Boyles, 2016). Based on the CTM, this study proposes a straightforward and robust approach for simulating the time-dependent distribution of traffic flows and estimating vehicular emissions using traffic counting data under the framework of CTM. In the proposed framework, the study area is first divided into grid cells. Comparing to the traditional CTM in which the

cell length is considered equal to the distance traveled by traffic flows, the proposed approach maintains the actual lengths and topological structures of the original road networks. In the proposed approach, the road networks in each cell are generalized into the visiting probabilities from possible origins to possible destinations through the shortest path calculation between the boundaries of the cells. Based on the visiting probabilities of each cell, four types of generated cell-based vehicular flows are simulated, i.e., station-to-station flows, station-to-boundary flows, boundary-to-boundary flows and boundary-to-station flows. In addition, since the traffic emissions are summarized by cells in this approach, the numbers and the types of vehicles instead of their accurate locations in each cell are required in the traffic flow simulation, which could significantly reduce the computational cost. Therefore, this approach can be implemented in a large area such as the entire road networks in a city. Since the routes of vehicles between origins and destinations are designated in the proposed framework, the emissions of each traffic flow are then estimated based on the volume and length of each traffic flow. In a case study, we simulate the time-dependent traffic flows and estimate pollutant emissions of Carbon Monoxide (CO), Nitrogen Oxide (NOx) and Violate Organic Compounds (VOC) in Hong Kong. The spatiotemporal patterns of traffic flows as well as emissions are visualized and illustrated in the developed space-time integrated framework. Results of simulation of traffic flows and estimation of emissions are validated using ground truth data both from traffic counting stations and statistics from the Environmental Protection Department of Hong Kong, with accuracies of around 78.6% and 70%.

This article is organized as follows. Section 2 introduces the data and methodology for simulating traffic flows and estimating traffic emissions. A case study and validation are presented in Section 3. Some related issues and limitations of this work are discussed in Section 4. Section 5 concludes this study and discusses the future work.

2. Methodology

This section introduces the proposed framework for simulating traffic flows and estimating vehicular emissions at the scale of a city. First, datasets and study area are introduced. The study area is divided by cells with resolution of 800 m × 800 m, which is the basic unit for simulating traffic flows and estimating vehicular emissions. Then we build a cell model to obtain the cell features including route lengths and visiting probabilities of counting stations and boundaries of each cell. The cell features obtained from the cell model determine the moving direction of traffic flows, which are the rationale for simulating traffic flow movement. Lastly, a flow model was proposed to simulate the time-dependent movement of traffic flow including origin, destination, volume as well as emissions of each traffic flow.

2.1. Data and study area

This study uses the entire city of Hong Kong as a test-bed. Traffic counting data is obtained from 169 counting stations. In addition to the traffic counting data, road network data is used to estimate road transport-related emissions. The road network of Hong Kong, the distribution of counting stations and dividing cells for the study area are illustrated in Fig. 1.

2.1.1. Traffic counting data

The traffic counting data is obtained from the annual traffic census (ATC) provided by the Transport Department of Hong Kong (TDHK, 2015). The original traffic counting data represent the average traffic flows for the entire year of 2015, and is recorded as the total number of vehicles passing through each of the 169 counting stations in 24 h and hourly percentage of the number of vehicles. The traffic flow data is divided into three main groups, namely the average traffic flow on Weekday (Monday to Friday), Saturday and Sunday. Based on the

annual traffic census data, hourly number of vehicles passing through each counting station on Weekday, Saturday and Sunday is obtained The traffic counting data is a reliable data source because it is the most frequently updated and qualitatively accurate among different data sources (Xie & Duthie, 2015). Recorded as hourly traffic counts in each counting station, the traffic flow data can also reflect people's socio-economic activities. In this study, the traffic counts are updated at each hour, which can provide us the dynamic traffic flow information to simulate the time-dependent traffic demands.

2.1.2. Road network in Hong Kong

Road network used in this study is obtained from OpenStreetMap, which contains 26,337 road polylines with seven types, i.e., motorway, trunk, major, secondary, territory, residential and services. The types of roads are used to determine the weights and probabilities of roads visited by vehicles. The road networks in the study area are first discretized by the cells at the boundaries of each cell, which create the sub-road networks in each cell and the nodes on the boundaries of each cell. With each cell as the basic unit for simulating traffic flows and associated emissions, the topologies of the original road networks are divided into two levels of topologies in the discretized road networks, i.e., the topologies within each cell and the topologies between the cells. For the topologies of the road networks within each cell, the sub-road networks still retain their own topologies and flow restrictions including connectivity between roads, crossing, grades of roads, turn restrictions. These features are considered when calculating the shortest paths between nodes on the boundaries and between counting stations. For the topologies between cells, the connectivity between cells is enabled by the nodes on the boundaries since the original road networks have been generalized into cells which are the basic unit in traffic flow simulation.

2.2. Cell model

The cell model is the basis for simulating traffic flow movement, which derives several cell-features to determine the moving direction of traffic flows. Fig. 2 shows the elements in a cell model. With (h, v) denoting the horizontal and vertical coordinates, a cell model is defined as:

$$\text{Cell}(h, v) = \{S, B, N, SP, V, P\}$$

In the cell model,

(1) S represents a list of counting stations in the cell, which can be described as a list: $\{[ID_1, (x_1, y_1)], …, [ID_{Ns}, (x_{Ns}, y_{Ns})]\}$ with ID and locations (x, y) of each counting station. S is optional in a cell model since not all the cells are located with counting stations.

(2) B denotes four boundaries of a cell, i.e., B_i, $i \in$\{North, South, West, East\}. Accordingly, since road network can intersect with each boundary of a cell at different nodes, N denotes the intersecting node sets for each boundary of a cell, i.e., N_i, $i \in$\{North, South, West, East\}. For instance, the road network in Fig. 2 intersect with the cell at seven nodes, i.e., P_1—P_7. The node sets on the four boundaries in the cell are $N_{North} = \{P_7\}$, $N_{South} = \{P_4\}$, $N_{West} = \{P_1, P_2, P_3\}$ and $N_{East} = \{P_6, P_5\}$, respectively.

(3) SP is the average length of the shortest paths between possible origins and destinations in a cell. The origin and destination can either be counting stations or boundaries of a cell. Hence, the SP from a counting station to a boundary, from a boundary to a boundary, from a boundary to a counting station and from a counting station to a counting station are included in SP. The SP in a cell is calculated based on the assumption that all vehicles tend to reach their destinations in the shortest path, which are calculated as:
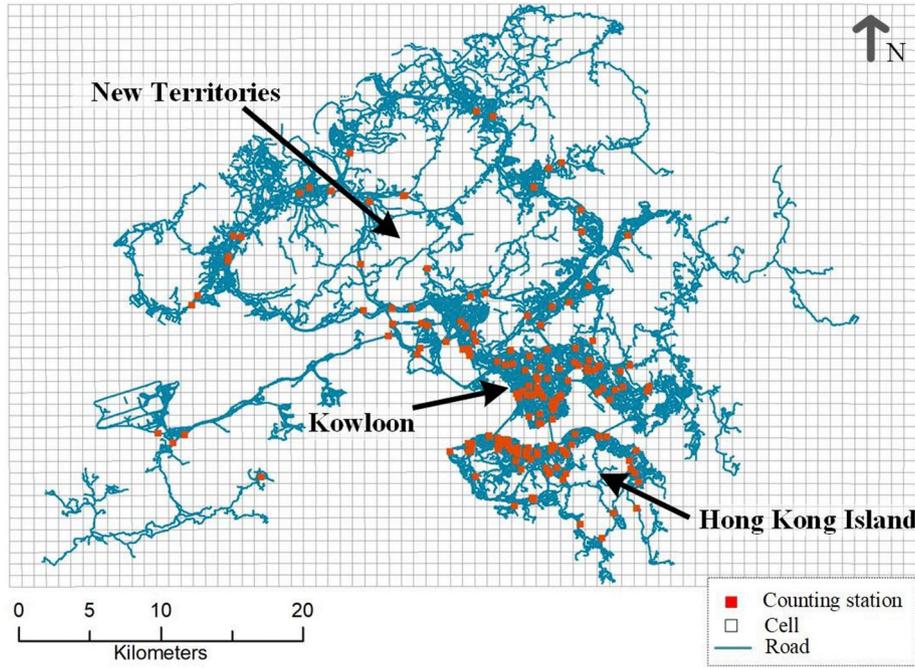
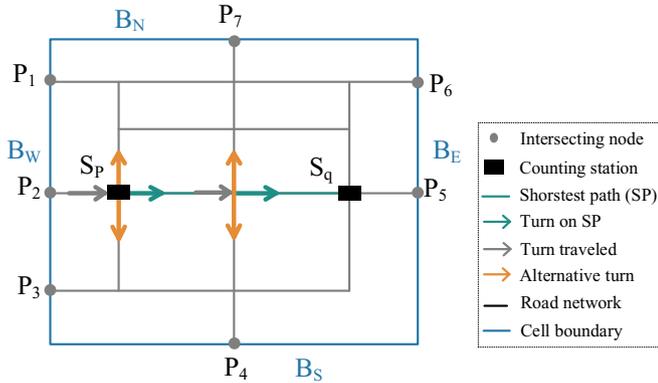**Fig. 1.** Road networks and distribution of counting stations in the study area.

Fig. 2. Elements of a cell model.

$$SP_{Sp \to Bi} = \frac{\sum SP(Sp \to Ni)}{\#Ni}, i \in \{North, South, West, East\} \qquad (1)$$

$$SP_{Bj \to Bi} = \frac{\sum \sum SP(Ni \to Nj)}{(\#Ni) \times (\#Nj)}, i, j \in \{North, South, West, East\} \qquad (2)$$

$$SP_{Bi \to Sp} = \frac{\sum SP(Ni \to Sp)}{\#Ni}, i \in \{North, South, West, East\} \qquad (3)$$

$$SP_{Sq \to Sp} = \frac{\sum SP(Sq \to Sp)}{\#Sq}, q = 1, 2, ...,(Ns - 1) \qquad (4)$$

where $SP(Sp \to N_i)$, $SP(N_i \to Sp)$ are the length of the shortest path from a counting station Sp to a node N_i and from a node N_i to a counting station Sp. Since there may be more than one node in each boundary, $\Sigma SP(Sp \to N_i)$ is the sum of lengths of the shortest paths from a station Sp to each node (N_i) of a boundary. $\#Ni$ is the total number of nodes in a boundary. Similarly, $\Sigma\Sigma SP(N_i \to N_j)$ is the sum of the length of shortest path between each node in boundary B_i and each node in boundary B_j.

(4) V consists of the maximum speed *Vmax* and average speed *Vave* in each cell, which are obtained as average values of the maximum speed and average speed of all roads in the cell.

(5) P represents the visiting probabilities from possible origins to possible destinations in a cell, including $P_{S \to Bi}$, $P_{Bi \to Bj}$, $P_{Bi \to S}$ and $P_{Sp \to Sq}$. The visiting probabilities determine the moving directions and volumes of vehicular flows in a cell. The probability that a boundary to be visited ($P_{S \to Bi}$, $P_{Bi \to Bj}$) is determined by the connectivity and accessibility of the boundary, which are related to the number of intersecting nodes. As it is shown in Eq. (5), the visiting probability either from a counting station or a boundary to a boundary is proportional to the number of intersecting nodes on the destination boundary.

$$P_{S \to Bi} = P_{Bj \to Bi} = \frac{\#Nodes\ i}{\sum \#Nodes\ j}, i, j \in \{N, S, W, E\} \qquad (5)$$

For the calculation of $P_{Bi \to S}$ and $P_{S\text{-}S}$, in contrast, it is obtained through accumulating turning probabilities at road intersections in the shortest path. In addition, since roads with high level tend to have higher probabilities to be visited than roads with lower level, each road is then assigned with a weight corresponding to its type (i.e., motorway, trunk, major, secondary, territory, residential and services). The turning probability P_i at a road intersection is thus calculated as the weights of the road that a vehicle turns to divided by the sum of the weights of alternative turns (U-turn is not considered in this model). Therefore, the probability of a single turn in the shortest path and visiting probability to S_q can be calculated as Eqs. (6) and (7) shows.

$$P_j = \frac{W_{SP} \times (\#Turn_{SP})}{\sum (W_{alternative} \times \#Turn_{alternative}) - W_{traveled} * \times Turn_{traveled}} \qquad (6)$$

$$P_{Sq \to Sq} = P_{Bi \to Sq} = \prod_{j=1}^{N} P_j \qquad (7)$$

Take the visiting probabilities in Fig. 2 for an instance, the $P_{S\text{-}S}$, $P_{B\text{-}S}$, $P_{S\text{-}B}$, $P_{B\text{-}B}$ in the cell is shown in Table 1, based on Eqs. (5)–(7).

2.3. Flow model

Based on the cell model, a flow model was proposed under the space-time integrated framework to simulate the movement and dispersion of time-dependent vehicular flows. In this study, a cell is the basic context unit for simulating traffic flows. The simulation of flows is

Table 1
Vising probabilities of S-S, B-S, S-B, B-B of the cell in Fig. 2.

O	D	Visiting probability
Sp	Sq	1/9
B_N	Sq	1/9
Sp	B_N	1/7
B_N	B_S	1/7
B_S	B_W	3/7
B_W	B_E	2/7

Fig. 3. Simulation of flow dispersion across space.

Fig. 4. Probability of space-time path of a flow.

flow. Fig. 4 shows the dispersion process of flows in a cell which originate from a counting station S_P and end at four boundaries of the cell. The destination and volume of each flow is determined by the visiting probabilities $P_{S \to Bi}$. Let Vs be the recorded vehicle numbers at the counting station S_P, the volume of each flow is calculated as Eq. (8) shows, in which $P_{S \to Bi}$ is the visiting probability from S to B_i in Eq. (5).

$$Volume_{S \to Bi} = \text{Vs} \times P_{S \to Bi}, i \in \{\text{North, South, West, East}\} \qquad (8)$$

However, the volume of each flow needs to be calibrated due to two reasons. First, since more than one counting station can locate in the same cell, the number of vehicles in a cell with two or more counting stations is over-recorded. Second, since some vehicles may park at parking lots after driving for a period of time (which cannot be detected by counting stations), the total traffic volumes recorded by counting stations are not consistent over time. Therefore, we need to determine the number of parking vehicles during each hour in order to simulate the traffic flows more accurately. To eliminate the over-recorded problem, the number of vehicles from each counting station S_P is adjusted to $V_{Sp}*(1\text{-}P_{SP \to SQ})$ for the cells with two or more counting stations. To solve the problem of inconsistency of traffic volume, we assume that the temporal distribution of vehicles parking at parking lots is opposite to the total number of vehicles moving on road network. First, the total number of vehicles moving on road network is considered as the total number of vehicles passing through a counting station during a day, which is denoted as N_S. Then we define hourly vehicular flow percent which represent the percentage of vehicles in a cell, i.e., P_S, obtained by dividing the volume of hourly vehicular flow by Ns. After that, the maximum, minimum and average values of Ps throughout a day can be derived, i.e., P_{max}, P_{min} and P_{avg}, based on which the parking percent of vehicular flow can be modeled as $P_{Park} = P_{max} + P_{min}\text{-}P_{avg}$. As a result, the volume of flow can be revised as $V_{SP}*(1\text{-} P_{Park})$.

(5) Emission. We estimate emissions for pollutants CO, NOx and VOC for each flow based on the attributes of volume, speed and length. To estimate the emission of a vehicle fleet, categories of the vehicle population is required. Since the emissions of vehicles are mainly determined. Since the traffic counts data recorded by the traffic counting stations cannot be distinguished between different types of vehicles, this study assumes that the composition of vehicles in the simulated traffic flows follows the distribution of vehicle types provided by the Hong Kong annual traffic census (TDHK, 2015), as shown in Table 2. Since the emissions of vehicles are mainly determined by vehicle types, fuel types and travel speeds, the emissions can be estimated based on the simulated traffic flows and the composition of the vehicles.

This study adopts COPERT V emission model for estimating emissions for the traffic flow. COPERT is a road transport emission inventory model financed by European Environment Agency (EEA). It categorizes vehicles into over 450 types according to vehicle parameters such as
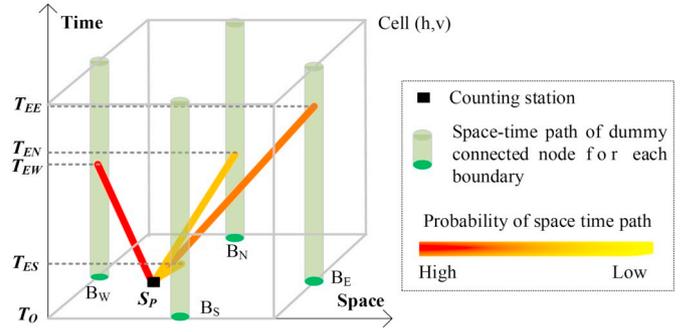
conducted periodically in each hour because the numbers of vehicles are recorded by counting stations in each hour. At the beginning of each simulation period, counting stations are considered as the origins of traffic flows, and the original volumes of flows are the numbers of vehicles recorded in each counting station. Dispersion of flows in space and time dimensions is simulated as space-time paths of flows based on visiting probabilities of each cell. A flow arriving at a boundary of the current cell will generate a new flow starting from the shared boundary of the adjacent cell.

Fig. 3 demonstrates the dispersion process of flows originated from a counting station S_P. In the figure, t_0—t_n is a time period for flow simulation. At t_0, the vehicular flow originated from S_P dispersed into four flows with different volumes. The four flows move to four boundaries of the cell and then pass the boundaries to reach their adjacent cells. Note that the time spent to move from Sp to their destinations are different because the lengths between Sp and each of the four boundaries are different. Thus, each flow needs to be tracked to obtain exact time of arrival at its boundary. When a flow has reached the boundary, this flow ends and becomes the original flow in the next cell, which is further dispersed based on the visiting probabilities in the next cell. In this way, the dispersion process is simulated and flows in each cell are recorded during each time period.

In summary, the flow model can be described as:

Flow = {ID, Cell(h, v), Origin$(S/B, T_O)$, Destination$(S/B, T_E)$, Speed

, Length, Volume, Emission}

In the model,

(1) ID is the ID number of a flow;
(2) Cell (h, v) is the cell where the flow currently locates.
(3) Origin $(S/B, T_O)$, Destination $(S/B, T_E)$ are tuples containing both location component (counting station S or boundary B of a cell) and temporal component (start time T_O and end time T_E). Speed is the average moving speed of the flow. Length is the distance traveled by a flow from its origin to its destination in the cell, which is calculated as SP in Eqs. (1)–(4).
(4) The Volume of a flow denotes the number of vehicles in the current

Table 2
Types and composition of vehicles registered in Hong Kong.

Vehicle type	Fuel type	Proportion
Motor cycle	Petrol	6.6%
Private car	Petrol	71.4%
	Diesel	0.7%
Taxi	LPG	2.5%
Single deck bus	Diesel	1.1%
Double deck bus	Diesel	0.8%
Light bus	Diesel	0.5%
	LPG	0.5%
Light goods vehicle	Petrol	0.1%
	Diesel	9.7%
Medium goods vehicle	Diesel	5.0%
Heavy goods vehicles	Diesel	0.9%

technical specifications, fuel and emission standards, and obtains the best-fit emission parameters for each type of vehicles through numerous bench tests and mathematical modeling. The COPERT model is adopted in this study because Hong Kong has been adopting the European emission standard since 1995, and the emission standards of the vehicles in Hong Kong in the year of 2015 mainly range from Euro II to Euro V. Since the volume of vehicular emissions is basically determined by vehicle's type, fuel, and moving condition, the emission parameters in COPERT are thus suitable for the fleet in Hong Kong. The feasibility of COPERT model in estimating emissions in Hong Kong has also been demonstrated in some related studies(Cen, Lo, & Li, 2016; Wang, Fu, Zhou, Du, & Ge, 2010; Xia & Shao, 2005), in which the emission factors for different types of vehicles are directly applied. Therefore, the emission factors in COPERT is considered valid for the traffic emissions in Hong Kong. In COPERT, basic emission factors (g/km) for vehicles can be calculated as in Eq. (9). The parameters *a-f* for each pollutant, vehicle category, fuel and emission standard can be referred to the COPERT handbook(https://copert.emisia.com/manual/).

$$EF_P = (a * v^2 + b * v + c + d/v)/(e * v^2 + f * v + g), p \in \{CO, NOx, VOC\} \tag{9}$$

For each vehicular flow *i*, the emission is thus obtained through taking its volume and length into account, as Eq. (10) shows.

$$E^i_p = EF_P * volume\ i * length\ i, p \in \{CO, NOx, VOC\} \tag{10}$$

Finally, the emission of each cell during any period [*ts*, *te*] can be obtained based on the emission for each flow. Suppose there are *M* flows in cell *i*: f_1^i, ..., f_M^i with origin time and destination times (T_{O1}^i, T_{E1}^i), ..., (T_{OM}^i, T_{EM}^i). Emissions for these flows are: E_1^i, ..., E_M^i. Therefore, the emissions for the cell *i* during time period [*ts*, *te*] can be calculated as:

$$Emission^i_{[ts,te]} = \sum_j^M \frac{[ts, te] \cap [T_{Oj}^i, T_{Ej}^i]}{[T_{Oj}^i, T_{Ej}^i]} Emission^i_j \tag{11}$$

3. Results

The proposed method is implemented using the spatial DBMS of PostgreSQL 11 with DBeaver 5.3 as a management tool in the database development. Within the spatial DMBS, geometry index is also used to accelerate the computation. The simulated traffic flow and emissions are visualized in the GIS software of ESRI ArcScene 10.2.

3.1. Traffic flow patterns during weekday and weekend

Fig. 5 shows the space-time patterns of traffic flow volumes on weekday and weekend during six time periods, i.e., 0–1 am, 4–5 am,

8–9 am, 12–13 pm, 16–17 pm, 20–21 pm. The traffic flows were then quantified as total moving length in each cell and visualized in a space-time integrated framework, which can both reveal the vertical differences of traffic flow patterns between different periods in a day and horizontal differences between the traffic flow during the same period of different days.

It can be observed from Fig. 5 that most traffic flows concentrated on areas in southern part of Kowloon Peninsula and the northern part of Hong Kong Island. In contrast, the hotspots of traffic flows in the New Territories are smaller and more dispersed. Fig. 5 also shows obvious differences in the traffic flow patterns between weekday and weekend. On weekday, there are obvious hotspots of traffic flows during 8 am–9 am, whereas concentrated areas of traffic flows on Sunday appear during 12 pm–13 pm. The observed differences of traffic flow patterns are mainly caused by the different travel patterns during weekday and weekend. In addition, according to annual surveys conducted by the Transport Department (Transport Advisory Committee, 2014), vehicles' average speed on the Hong Kong Island remains at around 20 km/h, and on some major traffic corridors, the vehicles' speed during weekday morning peak hours approaches to 10 km/h. Therefore, the areas with high traffic flow volumes in Fig. 5 are mainly caused by excessive vehicles.

3.2. Spatiotemporal patterns of vehicular emissions in Hong Kong urban area

Figs. 6, 7 and 8 show the spatiotemporal distributions of CO, NOx and VOC emissions on weekday, Saturday and Sunday, respectively, which are sampled at the same time periods as in Fig. 5. By comparing the patterns of emissions with the patterns of traffic flows in Fig. 5, it is revealed that the volumes of emissions are highly correlated to the distribution of simulated traffic volume, which indicates that traffic volume is the dominant contributor to the patterns of traffic emissions. In the aspects of emissions volume, Figs. 6–8 show that the volumes of CO emission are far higher than NOx and VOC emissions, and the volume of VOC emission is the lowest among all the three pollutants. Elevated areas of emissions are observed in the southern part of Kowloon Peninsula and the northern part of Hong Kong Island as well as their connected tunnels, which serve as transportation hubs for Hong Kong with large amount of traffic flows. By comparing the emissions on weekday with that on weekend, it can be observed that the hot spots of emissions on weekday appear earlier than that on weekend. Elevated areas of CO emissions are identified from 8 am on weekday while CO emissions on weekend show a more even distribution during the same time. Instead, obvious high values of CO emissions on Saturday and Sunday are observed after 12 am.

Compared with CO emissions, more elevated areas of NOx emissions are identified. In Fig. 7, more cells are observed to have relatively high values of NOx emissions than that in Fig. 6, which is probably due to different physical mechanisms of two pollutants. The traffic-related CO emissions are mainly released under vehicles' incomplete fuel combustion, which usually occurs when vehicles accelerate and decelerate in a short time period. Vehicles involved in traffic congestions often present such behaviors. As a result, cells in the southern Kowloon and the northern Hong Kong Island with higher probability of traffic congestions are identified to have much higher concentration of CO than other areas. In contrast, NOx emissions are released under the overloading condition of vehicle engines. Vehicles moving in high speed and over long distances are more likely to release high amount of NOx emissions. Therefore, not only areas with high traffic volumes but also areas with smooth traffic condition have relatively high values of NOx emissions. The similar pattern was demonstrated in a study which examined vehicular CO and NOx emissions under different traffic conditions (Zhang et al., 2011). The evidence showed that congestion was associated with the highest emissions of CO, while NOx emission rates
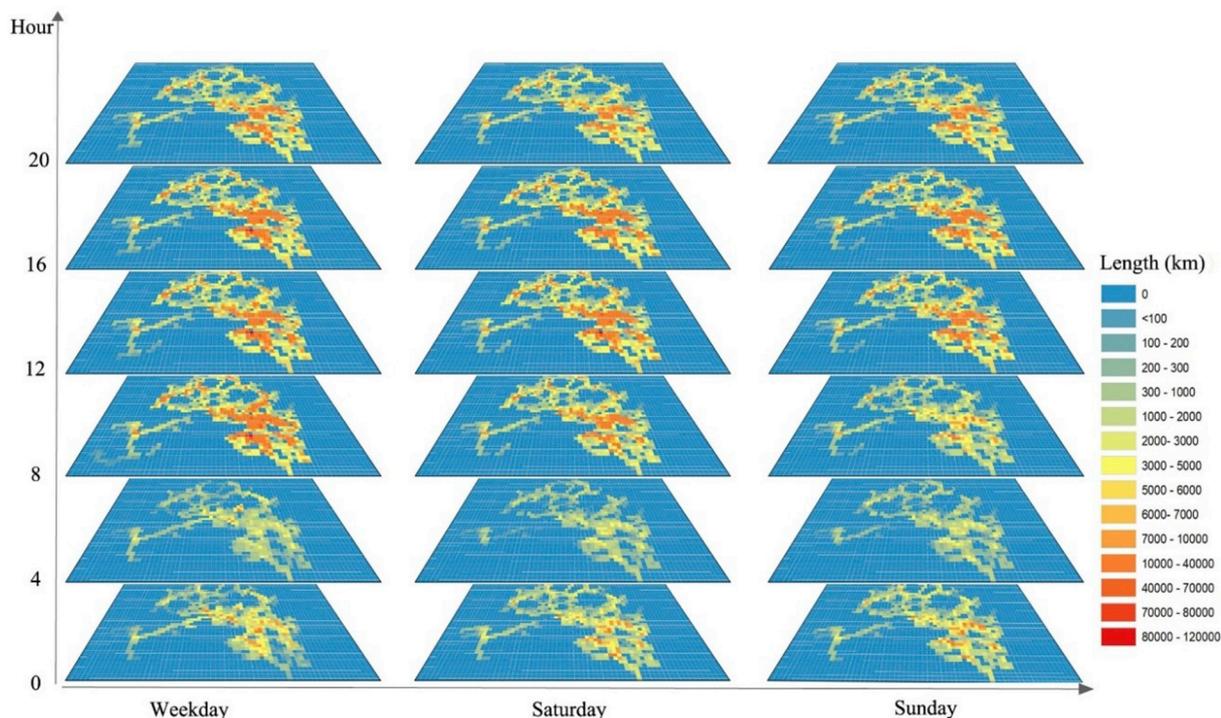
Fig. 5. Spatiotemporal patterns of traffic flows during weekday and weekend.

under the different traffic conditions were similar (Zhang et al., 2011).

The volume of violate organic compounds (VOC) emissions is the least compared with CO and NOx emissions. In Fig. 8, though VOC emissions have relatively low values, they present a more uneven distribution than that of NOx. As VOC has similar emission mechanism to CO, the elevated areas of VOC emissions are also observed in the southern Kowloon, northern Hong Kong Island as well as the cross-harbor tunnels.

3.3. Validation

3.3.1. Validation of the traffic flows simulation

In order to validate the proposed approach for simulating vehicular flows based on traffic counting stations, the simulated volumes of traffic flow in each hour on Weekday, Saturday and Sunday were compared with the count numbers of the counting stations for each hour across the study area. The accuracy of traffic simulation is shown in Fig. 9.
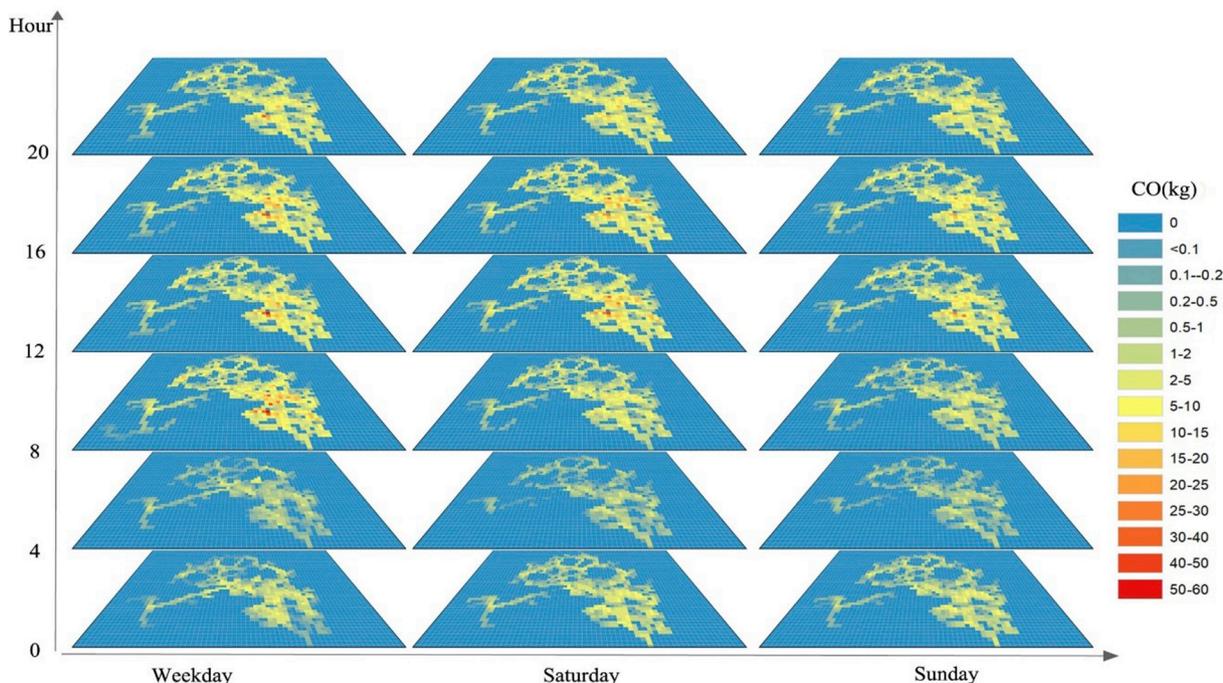

Fig. 6. Spatiotemporal patterns of CO emissions during weekday and weekend.
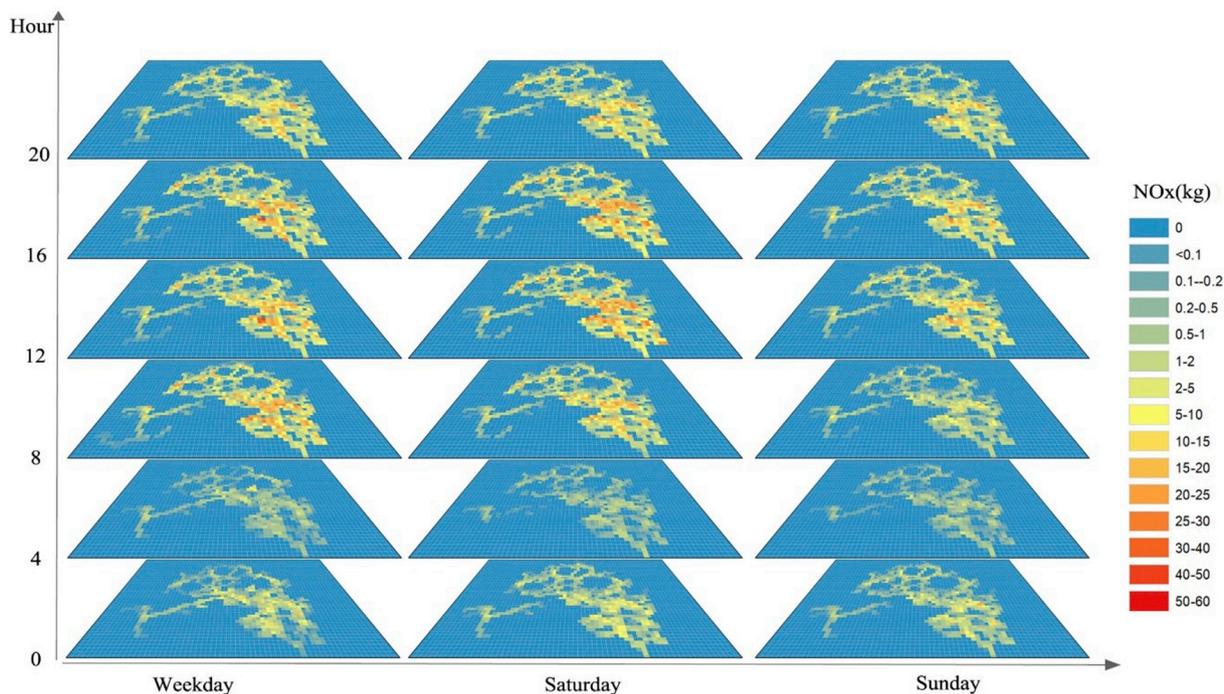
Fig. 7. Spatiotemporal patterns of NOx emissions during weekday and weekend.

Fig. 9 shows that the simulated traffic volumes after 9 am are more accurate compared with that between 0 am and 9 am. In addition, weekday has a higher simulation accuracy than Saturday and Sunday. The differences are probably due to a more predictable traffic on weekday than on weekend, especially during morning rush hours. Another factor contributing to the differences between simulated traffic volumes and the recorded traffic volumes is the lack of traffic counting stations in most cells. There are only 169 counting stations in the study area, only a small portion of cells (116 cells) in the study area are located with counting stations. The uneven distribution of traffic counting stations would increase the uncertainty of traffic flows simulation.

Nonetheless, the average accuracy of traffic simulation is 67.4%, and the average accuracy for weekday is 78.6%, which demonstrates the feasibility of the proposed approach to simulate traffic flows.

We further examine the spatial distribution of the average accuracies for the cells located with traffic counting stations on Weekday, Saturday and Sunday as shown in Fig. 10. Among the study areas, Kowloon Peninsula and the Hong Kong Island have denser population and traffic, while the population and traffic in the New Territories are sparser and smoother. Particularly, the traffic in Kowloon Peninsula is the most congested among all areas in Hong Kong. Fig. 10 shows that the cells with high accuracies cluster around Kowloon Peninsula, while
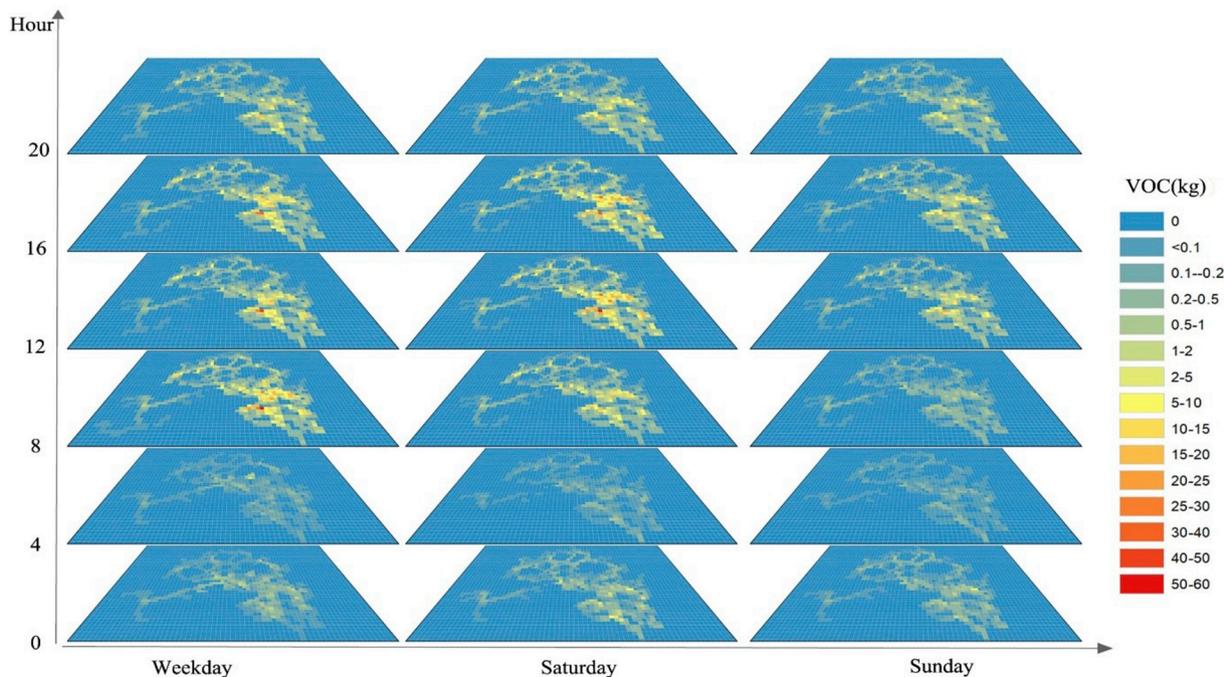

Fig. 8. Spatiotemporal patterns of VOC (Violate Organic Compounds) emissions during weekday and weekend.
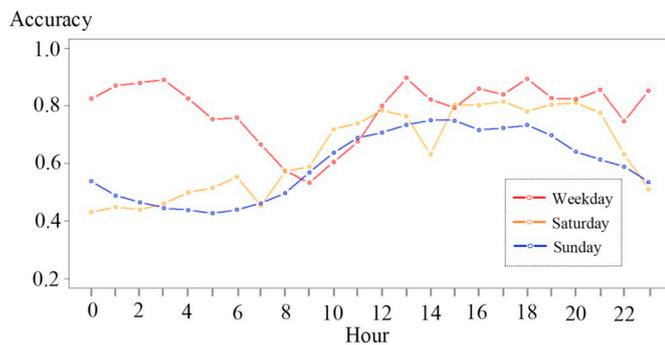
Fig. 9. Accuracy of traffic flow simulation.

there are more cells with low accuracies in Hong Kong Island. The spatial patterns of accuracies in these two areas are generally consistent on Weekday and Weekends. In the north of the study area, the New Territories also has some cells with low accuracies on Weekday and even more cells with low accuracies on Saturday and Sunday. By overlaying the distribution of accuracies with the locations of traffic counting stations, Fig. 10 further shows that the areas where counting stations cluster (such as the western Kowloon Peninsula) tend to have higher accuracies than the areas with fewer counting stations (such as the eastern New Territories). In general, Fig. 10 illustrates the spatial

variations of the accuracies across the cells with counting stations. The accuracies for traffic flow simulation tend to be higher in the cells with denser traffic flows and more counting stations.

3.3.2. Validating the emission estimation

The effectiveness of the proposed approach in emission estimation was evaluated. Measuring the exact volume of emissions requires professional equipment installed on individual vehicles, which can hardly be implemented in practice. Therefore, we validated our estimation results in a coarse granularity based on the statistics of road transport emissions in year 2015. According to the Environment Protection Department Hong Kong (HKEPD, 2016), the yearly emitting volumes of CO, NOx and VOC pollutants from road transport sector are 31,400, 18,100 and 4800 tons, which are equivalent to the daily emissions of 86,027, 49,589 and 13,150 kg, respectively. The statistics of the emission volume for CO, NOx, and VOC are considered as reference values, compared with which the daily estimations of the three pollutants and the corresponding accuracies are shown in Table 3.

In Table 3, it can be observed that the estimated volumes of the three pollutants are all lower than reference values. In addition to the errors of the estimation model, the underestimation is also caused by the traffic flow simulation. As demonstrated in Fig. 9, the accuracy of traffic simulation on weekday is higher than that on weekend, which is consistent with the fact revealed by Table 3 that the estimating accuracies on weekday are higher than that on weekend for all the
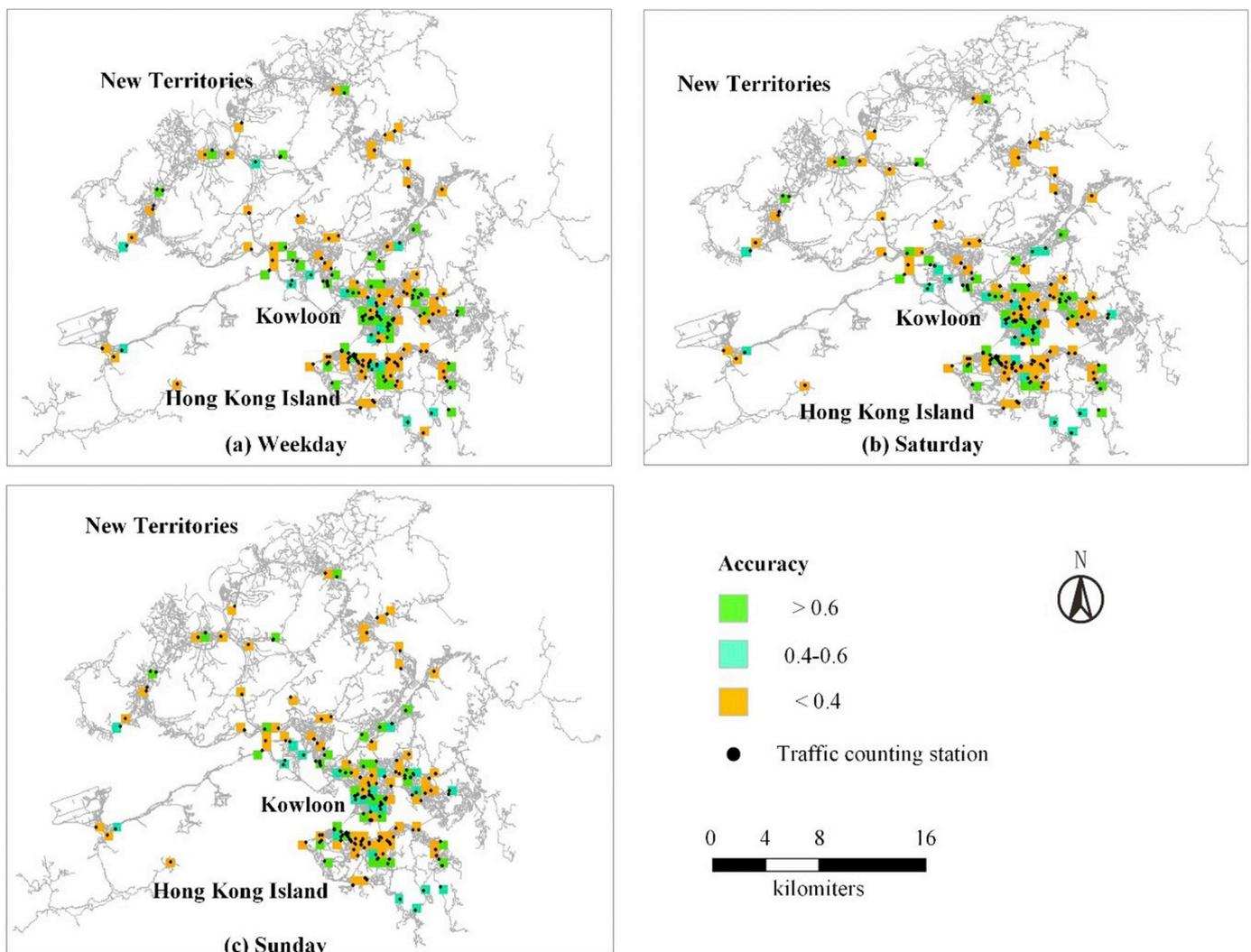

Fig. 10. The spatial variations of the accuracies of cells located with counting stations.

Table 3

Evaluation of the emission estimation results (tons).

Pollutant	CO		NOx		VOC	
Day	Estimation	Accuracy	Estimation	Accuracy	Estimation	Accuracy
Weekday	60,124.35	69.9%	47,109.14	95.0%	9401.0	71.5%
Saturday	56,038.31	65.1%	37,562.54	75.7%	8631.49	65.6%
Sunday	50,642.49	58.9%	37,378.98	75.4%	7837.53	59.6%
Ground truth	86,027		49,589		13,150	

pollutants. Table 3 also shows that the estimating accuracies for NOx and VOC are higher than that for CO. More specifically, NOx has the highest estimating accuracy, which is 95.0% for weekday, and around 75% for weekend. In summary, Table 3 shows that except for CO and VOC on Sunday (the accuracies are both below 60%), the estimated emissions are overall consistent with the ground truth with the accuracies of most estimations around 70%.

4. Discussion

This study simulated the time-dependent traffic flows and estimated pollutant emissions of Carbon Monoxide (CO), Nitrogen Oxide (NOx) and Violate Organic Compounds (VOC) in Hong Kong using space-time GIS techniques. The spatiotemporal patterns of traffic flows as well as emissions were analyzed and visualized in the developed space-time integrated framework, which provided insights about quantities and underlying mechanisms of emissions from road traffic. Results of traffic flows simulation and emissions estimation were cross-compared with reference values from traffic counting stations and census data. The effectiveness of the proposed approach in simulating traffic flows and estimating emissions is demonstrated with accuracies of 67.4% and 70%.

Instead of analyzing vehicular emissions in each road link, this study divided the study area into cells and deemed cells as basic unit for estimating traffic flows and vehicular emissions. Since the complex road network is generated as routes in each cell and further processed as visiting probabilities of different boundaries in a cell, the process of estimating total traffic emissions is notably simplified. The proposed approach utilizing cells is appropriate for studying traffic emissions in a large scale such as a city. In the case study, concentrations and distributions of CO, NOx and VOC emissions in Hong Kong have been studied, which may provide a holistic view for both transportation department and environmental department to plan to reduce traffic emissions in the coming future. For instance, inhabitant who are sensitive to a participle type of traffic pollutants can be advised to going to the areas with high concentrations of pollutant. The proposed simple and straightforward model has been found to be efficient and consistent with reference data. Moreover, the approach in this study has the potential to be further extended for evaluating traffic control strategies targeted at reducing traffic emissions for different pollutants.

There are also limitations in this study. First, it is advisable to consider the impact of urban structure and socioeconomic activities when simulating traffic flows since urban economic activity and employment density would affect residents' daily commuting behaviors as well as traffic flows. However, our study area may have some limitations on examining this impact. Hong Kong has one of the most developed and sustainable public transport system in the world. According to a report from the Transport Department of Hong Kong (TDHK, 2017), Hong Kong has the highest public transport usage rate in the world. In Hong Kong, more than 12 million trips are made through public transport services each day including railways, buses, taxis, trams and ferries, which account for over 90% of the total trips. For other major cities with renowned public transport systems, in comparison, the public transport usage rate is around 60% in Singapore, 70% in Tokyo and 30% in London and New York. Therefore, most of

socioeconomic and employment activities are conducted through public transport in Hong Kong. As shown in Table 2, taxis and buses only account for 5.4% of the total registered vehicles in Hong Kong, and the other vehicles on road network are motor cycles, private cars and goods vehicles. As a result, the traffic flows on urban road network in Hong Kong has a high level of uncertainty. In this situation, the hourly traffic counting data is feasible to simulate the traffic flow patterns. Second, the proposed approach is efficient for discovering and understanding patterns for traffic emissions at a large scale, which however, has a relatively weak support for explaining the emissions in a local area such as an intersection of roads and small lanes. Third, there are uncertainties in the processes of both traffic flows simulation and emission estimation. For the results of traffic flows simulation, the accuracy of simulation is greatly influenced by the distribution of traffic counting stations. Better results could be obtained when the spatial density of the stations is higher. In addition to the lack of traffic counting stations in some cells, the strategy for deploying counting stations is to locate more stations at areas with high traffic volume with low traffic volume. This strategy can help to collect more reliable traffic information for congested while locate fewer stations at areas, while in areas with less traffic, the traffic flows distribution become more uncertain. For the results of emissions estimation, in addition to the accumulating factor caused by uncertain traffic flow simulation, the proposed approach also inherits the limitations of COPERT, the underlying emission model. In COPERT, average speed for each cell was used in emissions estimation, which might render the approach not accurate for traffic conditions where the speed varies. Possible ways to improve the simulation of traffic flows and associated emissions include adopting people's travel patterns in traffic flows simulation and incorporating other microscopic traffic data such as vehicle GPS trajectories when estimating the traffic flow speed. This study focuses on the traffic volume simulation and assignment in the cell-based road networks based on traffic counts data. With available trajectory data or traffic flow speed data, the estimated traffic flows could be closer to the real-world traffic by incorporating the microscopic approaches in the traffic flow simulation.

The codes of traffic flow simulation and emissions estimation can be shared to the readers upon request.

5. Conclusion

This study proposed an approach for simulating time-dependent distribution of traffic flows and estimating vehicular emissions for all the vehicles in Hong Kong using traffic counting data. Under a space-time integrated framework, we simulated the time-dependent traffic flows and estimated pollutant emissions of Carbon Monoxide (CO), Nitrogen Oxide (NOx) and Violate Organic Compounds (VOC) in Hong Kong. The spatiotemporal patterns of traffic flows as well as emissions were analyzed and visualized in the proposed space-time integrated framework using space-time GIS techniques. With accuracies of 67.4% and 70%, the results demonstrated the feasibility of the proposed approach for estimating city-scale traffic flows and traffic emissions. Future work will focus on improving the existing CTM framework through considering the impact of urban structure and socioeconomic activities on traffic flow patterns.

References

Abo-Qudais, S., & Abuqdais, H. (2005). Performance evaluation of vehicles emissions prediction models. *Clean Technologies and Environmental Policy, 7*(4), 279–284.

Abou-Senna, H., Radwan, E., Westerlund, K., & Cooper, C. D. (2013). Using a traffic simulation model (VISSIM) with an emissions model (MOVES) to predict emissions from vehicles on a limited-access highway. *Journal of the Air & Waste Management Association, 63*(7), 819–831.

Amirjamshidi, G., Mostafa, T. S., Misra, A., & Roorda, M. J. (2013). Integrated model for microsimulating vehicle emissions, pollutant dispersion and population exposure. *Transportation Research Part D: Transport and Environment, 18*, 16–24.

Astarita, V., Guido, G., Mongelli, D., & Giofre, V. P. (2015). A co-operative methodology to estimate car fuel consumption by using smartphone sensors. *Transport, 30*(3), 307–311.

Barth, M., An, F., Younglove, T., Scora, G., Levine, C., Ross, M., & Wenzel, T. (2000). The development of a comprehensive modal emissions model. *NCHRP web-only document 122 Contractor's final report for NCHRP project 25-11* (pp. 307). National Cooperative Highway Research Program.

Burón, J. M., López, J. M., Aparicio, F., Martín, M.Á., & García, A. (2004). Estimation of road transportation emissions in Spain from 1988 to 1999 using COPERT III program. *Atmospheric Environment, 38*(5), 715–724.

Cai, H., & Xie, S. (2007). Estimation of vehicular emission inventories in China from 1980 to 2005. *Atmospheric Environment, 41*(39), 8963–8979.

California Air Resource Board (CARB) (2006). *EMFAC version 2.30 user guide: Calculating emission inventories for vehicles in California*.

Canudas-de-Wit, C., & Ferrara, A. (2018). A variable-length Cell Transmission Model for road traffic systems. *Transportation Research Part C: Emerging Technologies, 97*, 428–455.

Cen, X., Lo, H. K., & Li, L. (2016). A framework for estimating traffic emissions: The development of passenger Car emission unit. *Transportation Research Part D: Transport and Environment, 44*, 78–92.

Chang, X., Chen, B. Y., Li, Q., Cui, X., Tang, L., & Liu, C. (2013). Estimating real-time traffic carbon dioxide emissions based on intelligent transportation system technologies. *IEEE Transactions on Intelligent Transportation Systems, 14*(1), 469–479.

Chen, S. R., & Wu, J. (2011). Modeling stochastic live load for long-span bridge based on microscopic traffic flow simulation. *Computers & Structures, 89*(9–10), 813–824.

Chuo, H. S. E., et al. (2016). Computation of cell transmission model for congestion and recovery traffic flow. *2016 IEEE international conference on consumer electronics-Asia (ICCE-Asia)* (pp. 1–4). IEEE.

Daganzo, C. F. (1994). The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological, 28*(4), 269–287.

Daganzo, C. F. (1995). The cell transmission model, part II: Network traffic. *Transportation Research Part B: Methodological, 29*(2), 79–93.

Delis, A. I., Nikolos, I. K., & Papageorgiou, M. (2015). Macroscopic traffic flow modeling with adaptive cruise control: Development and numerical solution. *Computers & Mathematics with Applications, 70*(8), 1921–1947.

Fontes, T., Pereira, S. R., Fernandes, P., Bandeira, J. M., & Coelho, M. C. (2015). How to combine different microsimulation tools to assess the environmental impacts of road traffic? Lessons and directions. *Transportation Research Part D: Transport and Environment, 34*, 293–306.

Friesz, T. L., Bernstein, D., Suo, Z., & Tobin, R. L. (2001). Dynamic network user equilibrium with state-dependent time lags. *Networks and Spatial Economics, 1*(3–4), 319–347.

Gately, C. K., Hutyra, L. R., Peterson, S., & Wing, I. S. (2017). Urban emissions hotspots: Quantifying vehicle congestion and air pollution using mobile phone GPS data. *Environmental Pollution, 229*, 496–504.

Islam, T., Vu, H. L., Panda, M., Hoang, N., & Ngoduy, D. (2017). The accuracy of cell-based dynamic traffic assignment: Impact of signal control on system optimality. *arXiv preprint*. https://arxiv.org/abs/1708.03759 arXiv:1708.03759).

Javani, B., Babazadeh, A., & Ceder, A. (2018). Path-based capacity-restrained dynamic traffic assignment algorithm. *Transportmetrica B: Transport Dynamics*, 1–24.

Jie, L., Van Zuylen, H., Chen, Y., Viti, F., & Wilmink, I. (2013). Calibration of a microscopic simulation model for emission calculation. *Transportation Research Part C: Emerging Technologies, 31*, 172–184.

Kan, Z., Tang, L., Kwan, M. P., Ren, C., Liu, D., Pei, T., & Li, Q. (2018). Fine-grained analysis on fuel-consumption and emission from vehicles trace. *Journal of Cleaner Production, 203*, 340–352.

Kan, Z., Tang, L., Kwan, M. P., & Zhang, X. (2018). Estimating vehicle fuel consumption and emissions using GPS big data. *International Journal of Environmental Research and Public Health, 15*(4), 566.

Levin, M. W., & Boyles, S. D. (2016). A multiclass cell transmission model for shared human and autonomous vehicle roads. *Transportation Research Part C: Emerging Technologies, 62*, 103–116.

Lighthill, M. J., & Whitham, G. B. (1955). On kinematic waves II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A:*

Mathematical and Physical Sciences, 229(1178), 317–345.

Lo, H. K., & Szeto, W. Y. (2002). A cell-based variational inequality formulation of the dynamic user optimal assignment problem. *Transportation Research Part B: Methodological, 36*(5), 421–443.

Luo, X., Dong, L., Dou, Y., Zhang, N., Ren, J., Li, Y., ... Yao, S. (2017). Analysis on spatial-temporal features of taxis' emissions from big data informed travel patterns: A case of Shanghai, China. *Journal of Cleaner Production, 142*, 926–935.

Nie, Y., & Zhang, H. (2008). A variational inequality formulation for inferring dynamic origin–destination travel demands. *Transportation Research Part B: Methodological, 42*(7–8), 635–662.

Ntziachristos, L., Samaras, Z., Eggleston, S., Gorissen, N., Hassel, D., & Hickman, A. J. (2000). *COPERT III. Computer programme to calculate emissions from road transport, methodology and emission factors (version 2.1)*. Copenhagen: European Energy Agency (EEA).

Nyhan, M., Sobolevsky, S., Kang, C., Robinson, P., Corti, A., Szell, M., ... Ratti, C. (2016). Predicting vehicular emissions in high spatial resolution using pervasively measured transportation data and microscopic emissions model. *Atmospheric Environment, 140*, 352–363.

Parry, K., & Hazelton, M. L. (2012). Estimation of origin–destination matrices from link counts and sporadic routing data. *Transportation Research Part B: Methodological, 46*(1), 175–188.

PTV Planung Transport Verkehr AG (2005). *VISSIM 4.10 User Manual, Karlsruhe, Germany*.

Quaassdorff, C., Borge, R., Pérez, J., Lumbreras, J., de la Paz, D., & de Andrés, J. M. (2016). Microscale traffic simulation and emission estimation in a heavily trafficked roundabout in Madrid (Spain). *Science of the Total Environment, 566*, 416–427 (s).

Rakha, H., & Ahn, K. (2004). Integration modeling framework for estimating mobile source emissions. *Journal of Transportation Engineering, 130*(2), 183–193.

Rakha, H., Ahn, K., & Trani, A. (2003). Comparison of MOBILE5a, MOBILE6, VT-MICRO, and CMEM models for estimating hot-stabilized light-duty gasoline vehicle emissions. *Canadian Journal of Civil Engineering, 30*(6), 1010–1021.

Richards, P. I. (1956). Shock waves on the highway. *Operations Research, 4*(1), 42–51.

Sentoff, K. M., Aultman-Hall, L., & Holmén, B. A. (2015). Implications of driving style and road grade for accurate vehicle activity data and emissions estimates. *Transportation Research Part D: Transport and Environment, 35*, 175–188.

Shang, J., Zheng, Y., Tong, W., et al. (2014). Inferring gas consumption and pollution emission of vehicles throughout a city. *Proceedings of the 20th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1027–1036). ACM.

Sharma, P., & Khare, M. (2001). Modeling of vehicular exhausts—A review. *Transportation Research Part D: Transport and Environment, 6*(3), 179–198.

Siripirote, T., Sumalee, A., Ho, H. W., & Lam, W. H. (2015). Statistical approach for activity-based model calibration based on plate scanning and traffic counts data. *Transportation Research Part B: Methodological, 78*, 280–300.

Spiliopoulou, A., Kontorinaki, M., Papageorgiou, M., & Kopelias, P. (2014). Macroscopic traffic flow model validation at congested freeway off-ramp areas. *Transportation Research Part C: Emerging Technologies, 41*, 18–29.

Sun, D. J., Zhang, K., & Shen, S. (2018). Analyzing spatiotemporal traffic line source emissions based on massive didi online car-hailing service data. *Transportation Research Part D: Transport and Environment, 62*, 699–714.

Sun, Z., Hao, P., Ban, X., & Yang, D. (2015). Trajectory-based vehicle energy/emissions estimation for signalized arterials using mobile sensing data. *Transportation Research Part D-Transport and Environment, 34*.

The Hong Kong Environmental Protection Department, (2016)

Transport Advisory Committee (2014). *Report on study of road traffic congestion in Hong Kong*. Hong Kong.

Transport Department of Hong Kong (2015). The annual traffic census. https://www.td.gov.hk/en/publications_and_press_releases/publications/free_publications/the_annual_traffic_census_2015/index.html.

Transport Department of Hong Kong (TDHK) (2017). Public transport strategy study. https://www.td.gov.hk/filemanager/en/publication/ptss_final_report_eng.pdf.

U.S. Environmental Protection Agency (2009). *Motor vehicle emission simulator (MOVES) 2010: User guide. Report no. EPA-420-B-09-041, Ann Arbor, MI, December*.

U.S. Environmental Protection Agency (2014). *National Emissions Inventory Report*. (Accessed 8 April 2019).

Wang, H., Fu, L., Zhou, Y., Du, X., & Ge, W. (2010). Trends in vehicular emissions in China's mega cities from 1995 to 2005. *Environmental Pollution, 158*(2), 394–400.

Wang, Y., Szeto, W. Y., Han, K., & Friesz, T. L. (2018). Dynamic traffic assignment: A review of the methodological advances for environmentally sustainable road transportation applications. *Transportation Research Part B: Methodological, 111*, 370–394.

Xia, L., & Shao, Y. (2005). Modelling of traffic flow and air pollution emission with application to Hong Kong Island. *Environmental Modelling & Software, 20*(9), 1175–1188.

Xie, C., & Duthie, J. (2015). An excess-demand dynamic traffic assignment approach for inferring origin-destination trip matrices. *Networks and Spatial Economics, 15*(4), 947–979.

Xie, Y., Chowdhury, M., Bhavsar, P., & Zhou, Y. (2012). An integrated modeling approach for facilitating emission estimations of alternative fueled vehicles. *Transportation Research Part D: Transport and Environment, 17*(1), 15–20 (Calibration of a microscopic simulation model for emission calculation).

Yang, Q., Boriboonsomsin, K., & Barth, M. (2011). Arterial roadway energy/emissions estimation using modal-based trajectory reconstruction. *IEEE Conf. Intell. Transp. Syst. Proceedings, ITSC* (pp. 809–814). .

Zamith, M., Leal-Toledo, R. C. P., Clua, E., Toledo, E. M., & Magalhães, G. V. (2015). A new stochastic cellular automata model for traffic flow simulation with drivers' behavior prediction. *Journal of computational science, 9*, 51–56.

Zhao, P., Kwan, M. P., & Qin, K. (2017). Uncovering the spatiotemporal patterns of CO2 emissions by taxis based on Individuals' daily travel. *Journal of Transport Geography,*

62, 122–135.

Zhao, P., Liu, X., Kwan, M. P., & Shi, W. (2018). Unveiling cabdrivers' dining behavior patterns for site selection of 'taxi canteen'using taxi trajectory data. *Transportmetrica A: Transport Science,* 1–24.

Zhang, K., Batterman, S., & Dion, F. (2011). Vehicle emissions in congestion: Comparison of work zone, rush hour and free-flow conditions. *Atmospheric Environment, 45*(11), 1929–1939.

Zheng, H., & Chiu, Y. C. (2011). A network flow algorithm for the cell-based single-destination system optimal dynamic traffic assignment problem. *Transportation Science, 45*(1), 121–137.

Zhu, R., Wong, M. S., Guilbert, É., & Chan, P. W. (2017). Understanding heat patterns produced by vehicular flows in urban areas. *Scientific Reports, 7*(1), 16309.

Zietsman, J., & Rilett, L. R. (2001). Analysis of aggregation effects in vehicular emission estimation. *Transportation Research Record, 1750*(1), 56–63.

Ziliaskopoulos, A. K. (2000). A linear programming model for the single destination system optimum dynamic traffic assignment problem. *Transportation Science, 34*(1), 37–49.