

Research Article

Estimation of Hourly Traffic Flows from Floating Car Data for Vehicle Emission Estimation

Yun Jiang ¹, Guohua Song ¹, Zeyu Zhang ¹, Zhiqiang Zhai ¹ and Lei Yu ^{1,2,3}

¹Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing 100044, China

²Department of Transportation Studies, Texas Southern University, Houston, TX 77004, USA

³School of Intelligent Transportation, Xuchang University, Xuchang 461000, Henan, China

Correspondence should be addressed to Guohua Song; ghsong@bjtu.edu.cn

Received 16 October 2020; Revised 10 February 2021; Accepted 13 March 2021; Published 25 March 2021

Academic Editor: Dujuan Yang

Copyright © 2021 Yun Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to model air quality in heavy pollution days, a dynamic emission monitoring system is implemented in the Beijing road network, which requires the input of hourly traffic flows. Floating car data (FCD) is increasingly employed for flow estimation based on the fundamental diagrams to supplement data provided by stationary detectors. However, existing studies often used a typical fundamental diagram without considering the hysteresis phenomena and the uncertainty of traffic flow estimation. This study aims to develop a multiperiod fundamental diagram for the traffic flow estimation from FCD considering the hysteresis phenomena. The result shows that the proposed multiperiod fundamental diagram can improve the accuracy of flow estimation. The uncertainty of traffic flow estimation at both 10 minutes and 1 hour is also quantified, and the result indicates that the variation of the estimation uncertainty at 1 hour is lower than that at 10 minutes, with an average 7% reduction of the range of 95% confidence interval (CI). But there is no significant difference in magnitudes of the estimation uncertainty at 1 hour compared with that at 10 minutes. Moreover, the uncertainty for congested flows is lower than that for free flows. In the case study, the proposed model is employed to develop the spatial and temporal distributions of flows and emissions for the metropolitan area in Beijing.

1. Introduction

Heavy air pollution days have occurred frequently in some megacities of China in the past decades [1, 2]. In order to model air quality and cope with heavy pollution days, one of the important tasks is monitoring emissions of major pollution sources, including industrial, agricultural, domestic, and mobile pollution. And emissions of these pollution sources should be monitored at a uniform interval, according to the Bulletin on the Second National Census of Pollution Sources released by China's Ministry of Ecology and the Environment [3, 4].

Emissions of stationary pollution sources, such as industrial, agricultural, and domestic pollution sources, can be dynamically monitored at 1 hour on a small scale. However, it is not practically easy to conduct dynamic monitoring of

mobile pollution sources on the road network, which requires the hourly traffic flow and travel speed of each road link [5–7].

And it is difficult to collect reliable traffic flows for the whole road network, which has become one of the problems to be solved in most cities to carry out dynamic traffic emission estimation. This is primarily attributed to the insufficient coverage of traffic detectors on the road network [8]. Conventional traffic monitoring systems use stationary detectors, such as inductive loops, traffic cameras, radio frequency identification, and remote traffic microwave sensor (RTMS) to collect traffic flows [9, 10]. The stationary detectors remain quite expensive to install and maintain. Therefore, these technologies are too costly to cover the traffic flows of metropolitan areas completely at high temporal-spatial resolution.

With the development of mobile traffic monitoring technology, the real-time speed data collection of the large-scale road network has been achieved, covering various types of urban roads, such as expressways, major arterial roads, and minor arterial roads. For example, the floating car system in cities such as Beijing [11, 12], Shanghai, Guangzhou, and Hangzhou, as well as the navigation software of Internet companies such as AutoNavi and Baidu, can provide real-time speed data with high coverage in a simple and cost-efficient way. In this context, floating car data (FCD), also referred to as probe vehicle data, is widely employed to provide traffic information as a supplement to traditional traffic monitoring techniques because of its low cost and high coverage [13, 14]. The FCD consists of position, direction, and speed information of probe vehicles equipped with on-board Global Positioning System (GPS). The data is transmitted to a central processing unit that processes data for all probe vehicles to calculate average speeds of road links [11, 15].

In this context, for the road links without traffic detectors, deriving traffic flows from the FCD speed based on the macroscopic fundamental diagram has important practical significance for the dynamic estimation of vehicle emissions.

However, the accuracy and uncertainty of this method has been questioned in practical applications, which has a great impact on the results of vehicle emission estimation. At a specific speed, the observed flow tends to fluctuate around the estimated flow in a range. The uncertainty of the flow estimation can be illustrated as the difference between the estimated flow and the observed flow [16]. Besides, the estimated traffic flow is often aggregated to 1 hour or even larger time intervals.

Therefore, the object of this study is to address the issue: how much uncertainty is the hourly flow estimation from the FCD for vehicle emission estimation?

2. Literature Review

FCD or probe vehicle data is obtained from GPS-equipped probe vehicles, and it can be provided by private companies. Many cities such as Beijing, Shanghai, and Berlin have established taxi-FCD systems, which have the advantages of the long operation time, wide coverage, and centralized management [17]. Meanwhile, private companies such as AutoNavi, Baidu, INRIX, and TomTom also collect link-based speed and travel time data from private vehicles. Gühnemann et al. [9] reported that taxis could be used as cheap probe vehicles for an extensive FCD base without additional expenses for on-board equipment of vehicles or software at the server site of the taxi headquarters, and their utilization and coverage of cities were much higher than private vehicles. Llorca et al. [18] reported that collecting commercial FCD from private cars arose privacy issues due to possible liability for speed limit violations. Thus, FCD providers typically choose not to report speeds more than posted speed limits to protect their customers, resulting in the speed truncation at the posted speed limits [15]. Sun et al. [17] studied a data fusion framework based on multisource

FCD from the taxi, Baidu, and bus. The method described in this paper utilizes GPS data from the taxi dispatching system of taxi companies. As the penetration rate of probe vehicles has increased and FCD has become more and more available, the proposed method in this paper can also be applied.

FCD has been widely used in traffic analysis, such as travel time estimation [19, 20], traffic state estimation [21], and congestion detection [22]. However, the quality of FCD and its true potential to be used in traffic analysis have been concerned about for a long time, as the FCD only monitors a limited sample of the total traffic [23]. Some studies discussed the penetration of probe vehicles required for FCD use. Sunderrajan et al. [24] concluded that a minimum of 5%–10% probe vehicles was needed to achieve reliable traffic state estimating, depending on the prevailing traffic conditions. Houbraken et al. [25] examined the potential of FCD to be used for real-time traffic management, which was collected by a county-wide FCD system in the Netherland covering 6–8% of all vehicles. The results showed that FCD presented a valuable alternative for loop data in dynamic traffic management systems. The research of Kerner et al. [26] showed that high-quality reconstruction of the actual travel times in the road network could be achieved when the penetration rate of probe vehicles was 1.5% among all vehicles. Talebpour et al. [27] investigated the fundamental diagrams under different penetration rates of connected vehicles.

Some studies attempt to investigate the potential of FCD use by comparing commercial FCD speeds with ground truth (GT) speeds. Altintasi et al. [15] evaluated the quality of FCD speeds provided by Be-Mobile in Turkey, which was compared with video-based GT speeds collected on a specific urban arterial road segment. The results showed that there was a nonlinear relation with a high correlation coefficient of 0.82. However, Hu et al. [28] reported that the magnitude of correlation varied with GT speeds for the arterials, ranging from -0.236 to 0.360 , and the highest correlation was observed when the speed was between 40 and 45 mph. Chase et al. [29] reported that the speed differences were normally distributed with a mean speed difference greater than 5 mph, and the nonsystematic differences were found to be caused by the time lag in recovery from slow speeds, which had an impact on the speed-flow relationship. Kim et al. [30] evaluated INRIX speeds over a two-month period, including recurrent and nonrecurrent events. They found that the reported INRIX speeds tended to lag the loop detector measurements by almost 6 min, and the effective sampling periods were between 3 and 5 min. Based on these findings, FCD speed data needs to be filtered and quality controlled before use to reduce nonsystematic error.

As the FCD data can provide the travel speed of the road link, it is widely used in fundamental diagram derivation. Bussion et al. [31] studied the fundamental diagram for Toulouse based on the flow and occupancy data of the loop detectors and found that the fundamental diagram was affected by the uniformity of the detector distribution, road types, and road congestion. Ambühl et al. [32] developed a methodology for estimating fundamental diagrams by fusing FCD and loop detector data. The results showed that the

fusion algorithm using two data sources simultaneously could reduce the fundamental diagram estimation error to 0.04 with an estimated 25% of all drivers generating FCD. Ji et al. [33, 34] determined a fundamental diagram by combining loop detector data and GPS data from probe vehicles. They found that data from 30% of the links could also generate a fundamental diagram, similar to the one derived from data on the whole network.

There have been many studies that attempt to estimate traffic flows from the FCD typically based on the fundamental diagrams [9, 24, 35–37]. Gühnemann et al. [9] derived traffic flows from FCD speed based on the relationship between the flow and space mean speed on the Berlin networks. Sunderrajan et al. [24] estimated the traffic state constituting average speed, density, and flow using FCD. They also investigated the impact of probe penetration and the sampling period on the estimation accuracy. Anuar et al. [35] estimated traffic flows based on the FCD speed by utilizing four fundamental diagrams including Greenshield, Underwood, Northwestern, and Van Aerde. It was concluded that Van Aerde's fundamental diagram provided the least error compared to the others. Neumann et al. [36] proposed a more detailed representation of the fundamental diagram based on stochastic Bayesian networks for estimating hourly traffic flows from probe vehicle data.

As for the uncertainty of traffic flow estimation, most of the existing studies concentrated on the statistical characteristics of the estimation errors. Anuar et al. [35] investigated the estimation errors from different speed-flow fundamental diagrams and FCD aggregation intervals (5 minutes, 10 minutes, and 15 minutes). The result showed that with increasing aggregation intervals of FCD, traffic flows could be more accurately estimated [35]. Moreover, the estimation of traffic flows during congested traffic conditions was found to be more accurate than that during free-flow conditions [35]. Mulligan et al. [37] analyzed the uncertainty of the traffic flow estimation using the moving-observed method, which was calculated using Monte Carlo simulations for each combination of controlled variables. Deng et al. [38] constructed several linear equations that mapped the traffic measurements as a function of cumulative vehicle counts on both ends of a traffic segment based on multiple data sources. Then, they quantified the uncertainty and value of estimation of traffic state by the mean and variance of estimated cumulative vehicle counts. Davis et al. [39] employed empirical Bayes methods to compute quantiles of the predictive probability distribution of the traffic total at a highway site based on a sample of daily traffic volumes from that site, to quantify the uncertainty of traffic totals estimation. Guo et al. [40] predicted short-term flow rate and quantified the uncertainty as kickoff percentage and width to flow ratio.

In summary, the speed-flow fundamental diagrams are widely used to estimate traffic flows from the FCD. A similar concept is also applied in this study. However, existing work has one limitation since they often employed typical fundamental diagrams without considering the spatial and temporal variations of traffic state. Besides, the distribution characteristics of the uncertainty of hourly traffic flow

estimation for emission estimation have not been investigated thoroughly in the existing studies.

In this study, an improved fundamental diagram considering the spatial and temporal variations of traffic state between the traffic flow and the travel speed is developed for traffic flow estimation from the FCD. The uncertainty of the flow estimation is defined, and the impact of aggregating the FCD and traffic flow data to the 1-hour interval is investigated, by employing bootstrapping to generate the distribution of the estimation uncertainty. Thus, the distribution characteristics of the hourly traffic flow estimation uncertainty from the FCD for vehicle emission estimation are quantified.

3. Materials and Methods

To overcome the limitations of existing studies, the general methodology used in this research includes four steps:

- (1) The FCD speed and the corresponding traffic flow from RTMS are prepared, both of which are matched geographically at 10 minutes
- (2) Based on the field data, a multiperiod fundamental diagram between the FCD speed and the traffic flow is developed with improved accuracy, compared with the typical Van Aerde model
- (3) The uncertainty of the flow estimation is defined and quantified. The effect of aggregation intervals on the uncertainty at different speed bins is also analyzed in this paper
- (4) As a case study, traffic flows that have been extracted from the FCD are further exploited for developing a dynamic emission model in Beijing

3.1. Data Sources and Preparation. In this study, two field data types are collected, including the FCD and RTMS data. The FCD only provides the travel speed without a direct measurement of traffic flows. In this case, a data fusion between FCD and traffic flow data from other sources, e.g., RTMS, can provide the needed information.

The FCD is collected by more than 60,000 taxis equipped with GPS devices in Beijing, which provides the travel speed at 5 minutes for road links. And the FCD system has been proven to be an available source of travel speed data in the previous study, with a penetration rate of 2.5% among all vehicles in Beijing [41–43]. The corresponding RTMS data is collected by 44 detectors on the ring expressways in Beijing, which provides reliable traffic flow data at 2 minutes [41]. In order to ensure the data quality of each detector, a data preparation method is designed, and the outlier data records were deleted. More detailed information about the data quality control and the potential of the FCD used in this paper has been illustrated in previous studies, which are investigated by comparing FCD speeds with RTMS speeds [11, 41].

Then, the FCD data of road links and RTMS data is matched geographically to ensure the temporal-spatial matching of the two data sources [41]. In this study, 8,640

pairs of speed and flows at 10 minutes were obtained from 3 detectors on the second ring expressway for five days (16 June 2011, 30 June 2011, 3 July 2012, 10 July 2012, and 17 July 2012).

3.2. Estimation of Traffic Flows Based on FCD Speed. To estimate traffic flows based on the relationship between the travel speed and the traffic flow for the whole network, it is necessary to develop the macroscopic fundamental diagram, which should be categorized into road types, speed limits, and other factors of traffic flow characteristics [44]. Therefore, the macroscopic fundamental diagram of the second ring expressway in Beijing is developed considering these influencing factors.

Besides, according to the Beijing traffic management policy, the truck can be allowed on the Beijing second ring only from 23:00 to 6:00, which makes the composition of the traffic flows during this period different from that during other time periods. Meanwhile, the traffic state during this period is in a state of free flow, and the speed is nearly constant and independent of the travel density. As a result, it is not reliable to estimate traffic flows from speed between 23:00 and 6:00; thus, this study only covers the period from 6:00 to 23:00 of workdays.

3.2.1. Existence of Hysteresis Phenomena on Fundamental Diagram. This paper analyzes the speed from FCD and corresponding flows from RTMS on the second ring expressway in Beijing, which entails a limiting speed of 80 km/h and six lanes in two directions.

Figure 1(a) shows the temporal variation characteristics of average flow and average speed. It is observed that the average flow remains at a high level between 7:00 and 23:00, and the temporal variation in the average speed is more significant all day. From 6:00 to 7:00, the traffic flow is at the onset of the congestion. From 7:00 to 18:00, the traffic flow is in a state of congestion with high flow and low speed. At 18:00, the traffic congestion reaches a peak at the minimum speed. After 18:00, the traffic flow is at the offset of congestion and reaches another peak at 21:00.

Note that the values of flows can be different although at the same speed during the day. For example, at the speed of 40 km/h, the values of flows are different for the onset and offset of the congestion [45]. To make it more explicit, this study plots the flow-density pairs for the whole network as shown in Figure 1(b). It is observed that the flow-density curves follow clockwise hysteresis loops, and the values of the maximum flow are different at the day and night peaks.

This phenomenon indicates that it is necessary to develop a multiperiod fundamental diagram that can describe the onset and offset of congestion separately. Thus, the traffic flow can be estimated more accurately, compared with typical flow estimation without considering this phenomenon.

Based on the analysis of the phenomenon of the hysteresis loops [45–49], the shape and characteristics of the fundamental diagrams are different in these two time periods. In this case, a multiperiod fundamental diagram is

developed for different time periods (from 6:00 to 18:00 and from 18:00 to 23:00) by averaging the data of all the detectors.

3.2.2. Development of Speed-Flow Fundamental Diagram.

To derive traffic flows from the travel speed, it is necessary to investigate the relationship between the two variables. As discussed earlier, it is necessary to divide the data into two time periods according to the shape and characteristics of fundamental diagrams. However, in practice, the speed-flow relationship is typically developed without considering the difference between different time periods. In order to determine the accuracy improved by the multiperiod fundamental diagram between the speed and the flow, the all-day speed-flow relationship is also developed for comparison purposes.

In the subsequent analysis, the RTMS data at 10 minutes is converted into the number of vehicles per hour per lane. By employing the Van Aerde single-regime model, four essential parameters including free-flow speed (FS), the speed at capacity (SC), traffic capacity (CAP), and jam density (JD) for the road network are calibrated by the regression [41, 50, 51]. Besides, the absolute percentage error (APE) is used as a performance indicator to reflect the deviation between the estimated flows and the ground truth measurements, as shown in

$$APE = \frac{|q_v - Q_v|}{Q_v}, \quad (1)$$

where Q_v is the observed flow at the FCD speed v and q_v is the estimated flow derived from FCD speed v .

3.3. Uncertainty Analysis of Flow Estimation with Bootstrap.

In the subsequent analysis, the uncertainty of the flow estimation, also referred to simply as the estimation uncertainty, is defined as the relative differences between the observed and the estimated flow. The uncertainty of traffic flow estimation can be illustrated as in

$$U_v = \frac{q_v - Q_v}{Q_v}, \quad (2)$$

where U_v is the uncertainty of flow estimation at stated (α), usually at 5%.

To investigate the distribution of estimation uncertainty for different aggregation intervals and speed bins, large samples of the relative differences between the observed and the estimated flows are required. In this study, bootstrapping is taken as a Monte Carlo style simulation method to obtain a 95% CI of the estimation uncertainty [52]. The 95% CI of the estimation uncertainty means that at the same speed, upon repeating to observe the traffic flow a hundred times, only 5% samples of the observed flow are expected to deviate from the estimated flow by more than U_v .

The uncertainty of traffic flow estimation is quantified using the following steps. First, the FCD speed data is divided into bins at 5 km/h interval, such as [5, 10), [10, 15), ..., [60, 65), and bootstrap data sets are developed for each

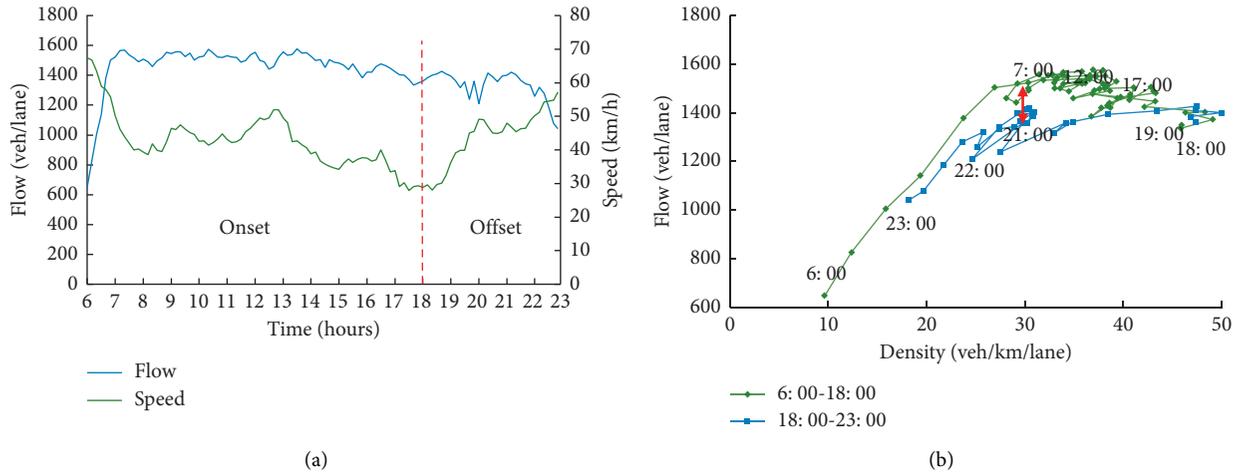


FIGURE 1: FCD speed and corresponding RTMS flows averaged over five days for the whole network. (a) Time-series of speed and flow. (b) Variation of flow with density.

speed bin, containing the speed and flow data of two aggregation intervals including 10 minutes and 1 hour. Second, the relative difference between the observed and the estimated flow of different aggregation intervals is calculated for each iteration. Third, the distribution characteristic of the estimation uncertainty is investigated, including the averages and 95% CI. The lower and upper bounds of the 95% CI are obtained by the 2.5% and 97.5% percentiles of the uncertainties of all iterations [52].

For each speed bin, to generate bootstrap data sets at 10 minutes, the original data pairs of the speed and flow data at 10 minutes of size N are taken as the data pool. Data pairs of size N are randomly sampled from the data pool with a replacement for each time. The routine is repeated 1,000 times. For each speed bin, to generate bootstrap data sets at 1 hour, the aggregated speed and flow data of size M are taken as the data pool. Data pairs of size M are randomly sampled from the aggregated data pool with a replacement for each time. The routine is repeated 1,000 times.

4. Results

4.1. Comparison of Fundamental Diagram for Different Times of the Day. The flow-speed fundamental diagrams of different times for the second ring road expressways in Beijing are shown in Figures 2(a) and 2(b), and the parameters of speed-flow fundamental diagrams for different times of the day are shown in Table 1. It can be observed that the regression curves can fit the speed-flow relationship well. However, when the speed is around the speed at capacity, observed flows are not always close to the theoretical capacity, and this separation of data points is caused by special characteristics of the ring road expressways in Beijing, which has been observed in the previous study [46]. And compared with the multiperiod model, this dispersion of data points is more significant in the all-day model.

In this paper, the estimated flows based on these different fundamental diagrams are compared with the observed flows from RTMS, as illustrated in Figures 3(a) and 3(b). It

can be observed that the estimated errors of the multiperiod model range from -300 to 300 , while the estimated errors of the all-day model range from -600 to 600 . Moreover, Figure 3(c) shows the average and deviation of APE for the different models of fundamental diagrams. According to Figure 3(c), the average APE (corresponding with the symbol “ \times ”) of the multiperiod model has decreased to 11%, and the value of the all-day model is 13%. Besides, the interquartile range and maximum of APE of the multiperiod model are reduced by 3% and 7%, which suggests that the variation of the estimated errors is reduced. From Figure 3(c), a key point that can be inferred is that a better performance appears in using the multiperiod speed-flow fundamental diagram.

4.2. Uncertainty of Traffic Flow Estimation. For each speed bin, the estimation uncertainty is calculated at different aggregation intervals, including 10 minutes and 1 hour. Figure 4 illustrates the distribution of the estimation uncertainty from data sets generated by bootstrapping.

Figure 4 shows that the variation of the estimation uncertainty is reduced as the time aggregation increases from 10 minutes to 1 hour. According to equation (2), the central tendency of the estimation uncertainty at 1 hour can be explained as the offset of positive and negative estimated errors at 10 minutes, resulting in less randomness in the process of aggregating the traffic conditions from 10 minutes to 1 hour. This rule is universal in different speed bins and especially significant in both high- and low-speed bins of ≥ 60 km/h and < 15 km/h.

Moreover, the variation of the estimation uncertainty is different across the speed bins. As seen in Figure 4, the estimation uncertainty presents a smaller size of the boxes and whiskers in the speed bins of 30–50 km/h, which suggests a smaller variation of estimation uncertainty. As Figure 2(a) shows, the data points are clustered around the fitting curve in the speed bins of around the speed at capacity (around 40 km/h). The traffic state in these speed bins is

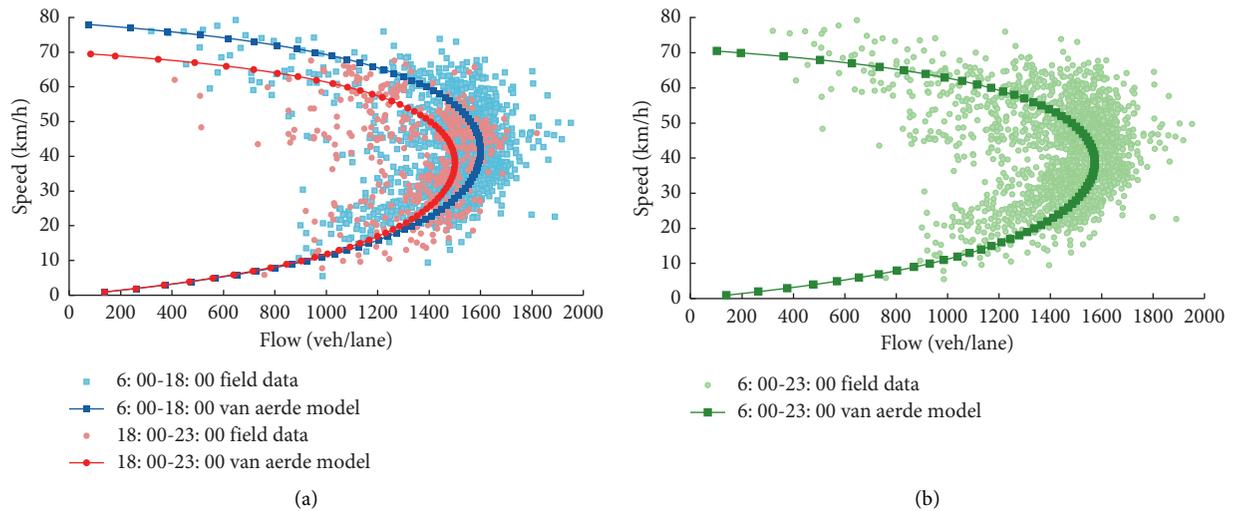


FIGURE 2: The flow-speed fundamental diagram of different times for the second ring expressway in Beijing. (a) Multiperiod model of the fundamental diagram. (b) All-day model of the fundamental diagram.

TABLE 1: Four parameters of the Van Aerdel model at different times are calibrated by the regression.

Time periods	FS (km/h)	SC (km/h)	CAP (veh/hour/lane)	JD (veh/km/lane)
6:00–18:00	78	41	1610	146
18:00–23:00	70	39	1501	145
6:00–23:00	71	39	1572	146

during the congested state. Besides, the variation of the estimation uncertainty of the two aggregation intervals is significantly high in the speed bins of ≥ 60 km/h. The traffic state in these speed bins is in the state of free flow, and the vehicle activities are much more complicated than those in other speed bins. Combined with Figure 2(a), the observed speed fits a wide range of flows in these speed bins.

To quantify the impact of time aggregation intervals for different speed bins on the estimation uncertainty, the averages and 95% CI of the uncertainty are obtained, which are presented in Figure 5 and Table 2.

According to Figure 5 and Table 2, key findings regarding the characteristics of the estimation uncertainty are summarized as follows.

- (1) There is no significant difference in the average of the estimation uncertainty with an increase in the aggregation level of the FCD and flow data.
- (2) For both of the two aggregation intervals, the average of the estimation uncertainty is relatively smaller during the transition from low- to high-speed bins, compared with that in both high- and low-speed bins of ≥ 60 km/h and < 15 km/h.
- (3) In terms of variation of the estimation uncertainty, the range of 95% CI at 1 hour is smaller than 2% for every speed bin, while the range of 95% CI at 10 minutes is higher than 12% in both high- and low-speed bins of ≥ 60 km/h and < 15 km/h.
- (4) The variation of the estimation uncertainty at 1 hour is relatively lower compared with that at 10 minutes for

each speed bin, with a 7% reduction of the range of 95% CI. This reduction of 95% CI is especially significant in the speed bins of ≥ 60 km/h and < 15 km/h, with 22% and 12% respectively.

5. Case Study

The general purpose is to estimate hourly vehicle emissions from transportation networks in Beijing. Therefore, traffic flows are estimated from FCD speeds based on the speed-flow fundamental diagrams.

A case study is conducted for the metropolitan area in Beijing. The traffic flow estimation and the dynamic vehicle emission estimation are developed based on FCD on 20 June 2016. Subsequently, the speed-flow fundamental diagrams of the Beijing road network are developed, which are categorized by road types, speed limits, and other traffic conditions to derive traffic flow of road links. The result of the hourly traffic flows from 8:00 to 9:00 is depicted in Figure 6(a).

FCD speed and the estimated traffic flows are further used to develop a dynamic vehicle emission model. This model is linked to a Geographic Information System (GIS) for visualization and further analysis of spatial effects [9]. The vehicle emissions are estimated as illustrated in the previous studies [53, 54]. Figure 6(b) depicts the spatial distribution of CO emissions from 8:00 to 9:00, which covers the metropolitan area of Beijing. Compared with conventional emission models, the dynamic emission model combined with the temporal resolution of FCD speeds can effectively identify emission hotspots at the high spatial and temporal resolution.

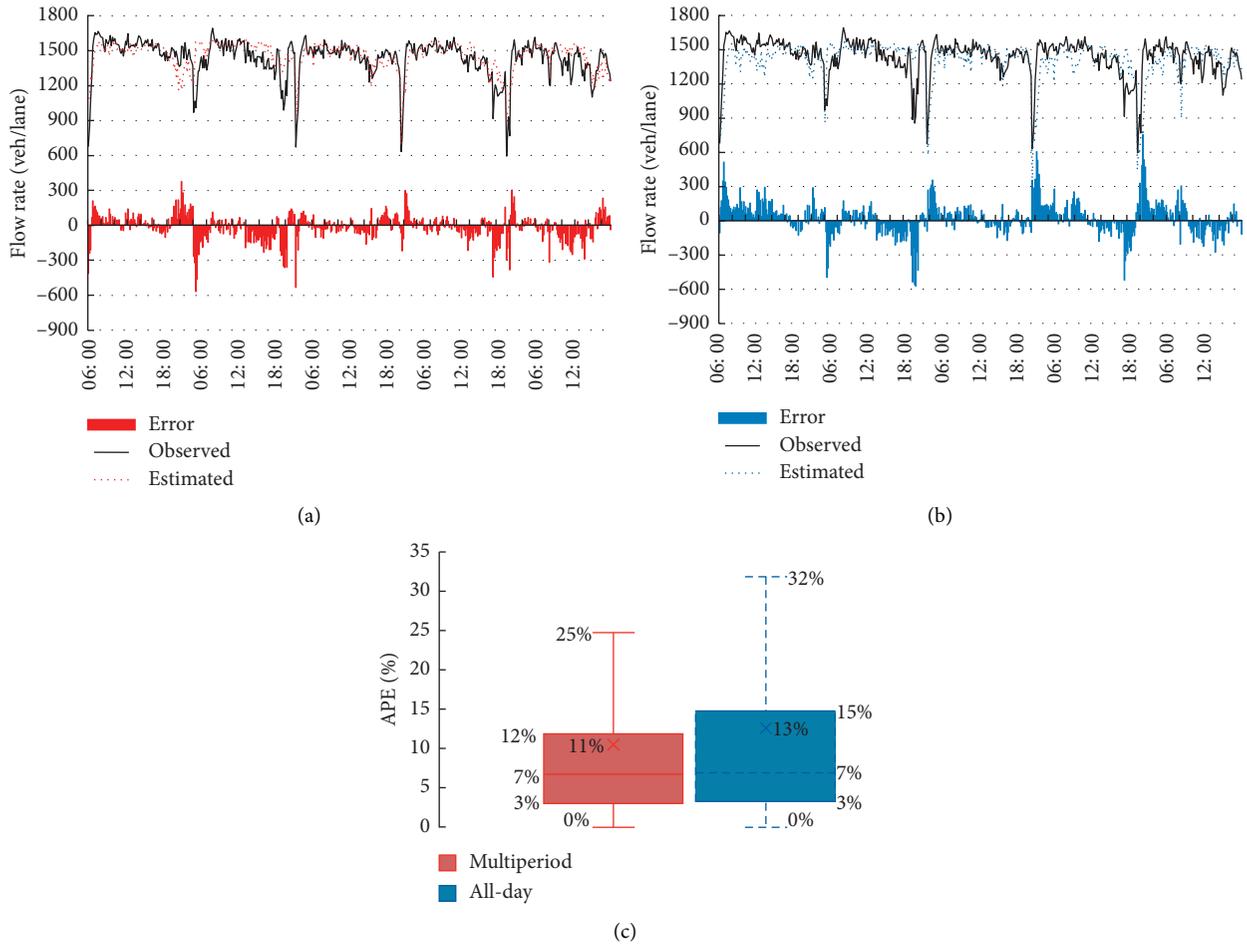


FIGURE 3: Comparison of observed and estimated flows of different models. (a) Multiperiod model. (b) All-day model. (c) Comparison between the APEs of different models.

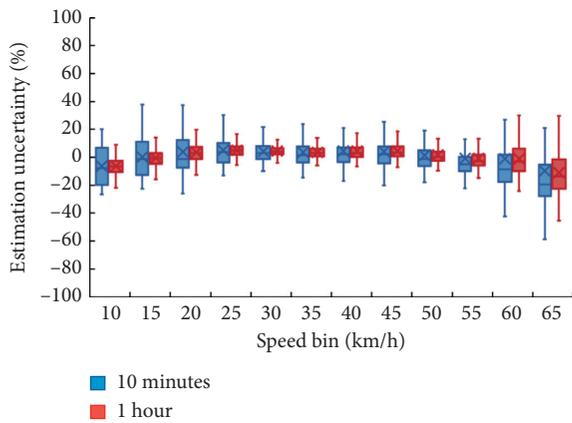


FIGURE 4: Distribution of the estimation uncertainty of different aggregation intervals.

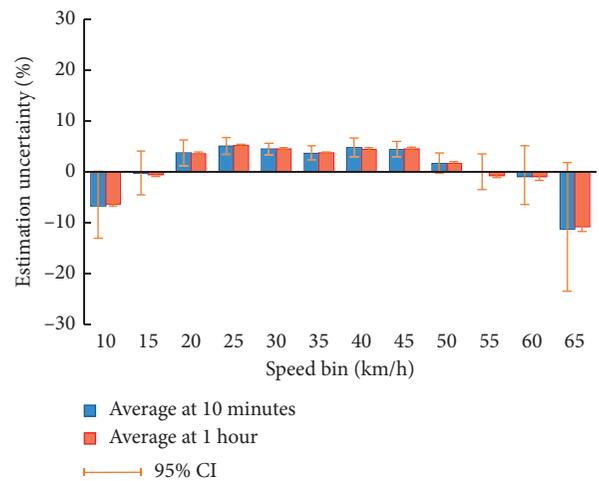
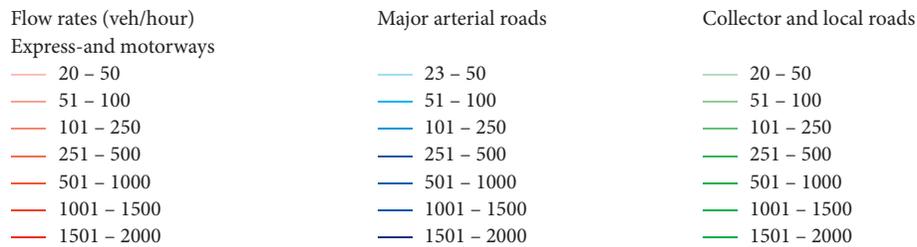
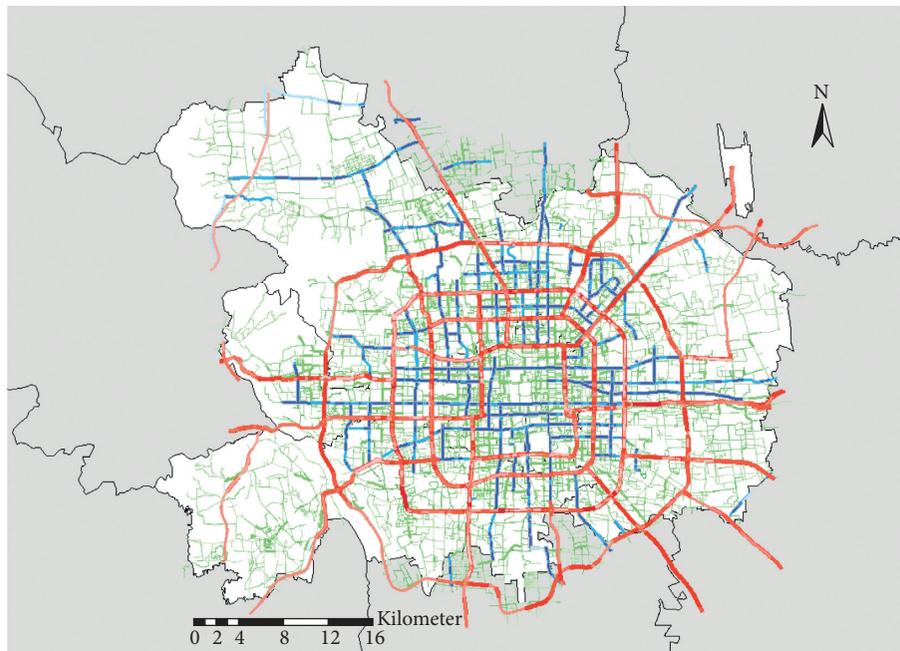


FIGURE 5: The characteristics of the estimation uncertainty based on data sets generated by bootstrapping.

TABLE 2: The average, lower, and upper bound of 95% CI and CV of estimation uncertainty.

Speed bin (km/h)	Uncertainty at 1 hour (%)			Uncertainty at 10 minutes (%)			Difference 1 hour vs 10 minutes (%)	
	Average	95% lower bound	95% upper bound	Average	95% lower bound	95% upper bound	Average	Range of 95% CI
10	-6.4	-6.8	-6.0	-6.7	-13.1	0.1	-0.3	-12.5
15	-0.6	-1.0	-0.2	-0.3	-4.6	4.0	0.3	-7.9
20	3.5	3.1	3.9	3.7	1.1	6.2	0.2	-4.3
25	5.2	4.9	5.4	5.0	3.4	6.7	-0.1	-2.8
30	4.5	4.2	4.7	4.4	3.4	5.6	-0.1	-1.8
35	3.7	3.4	3.9	3.6	2.3	5.1	0.0	-2.3
40	4.4	4.1	4.7	4.7	2.9	6.6	0.3	-3.0
45	4.5	4.1	4.8	4.4	3.0	6.0	-0.1	-2.4
50	1.6	1.3	2.0	1.6	-0.3	3.7	0.0	-3.3
55	-0.7	-1.2	-0.2	-0.1	-3.5	3.5	0.6	-6.1
60	-0.9	-1.7	-0.2	-1.0	-6.5	5.1	0.0	-10.1
65	-10.8	-11.8	-9.9	-11.3	-23.5	1.8	-0.5	-23.4



(a)

FIGURE 6: Continued.

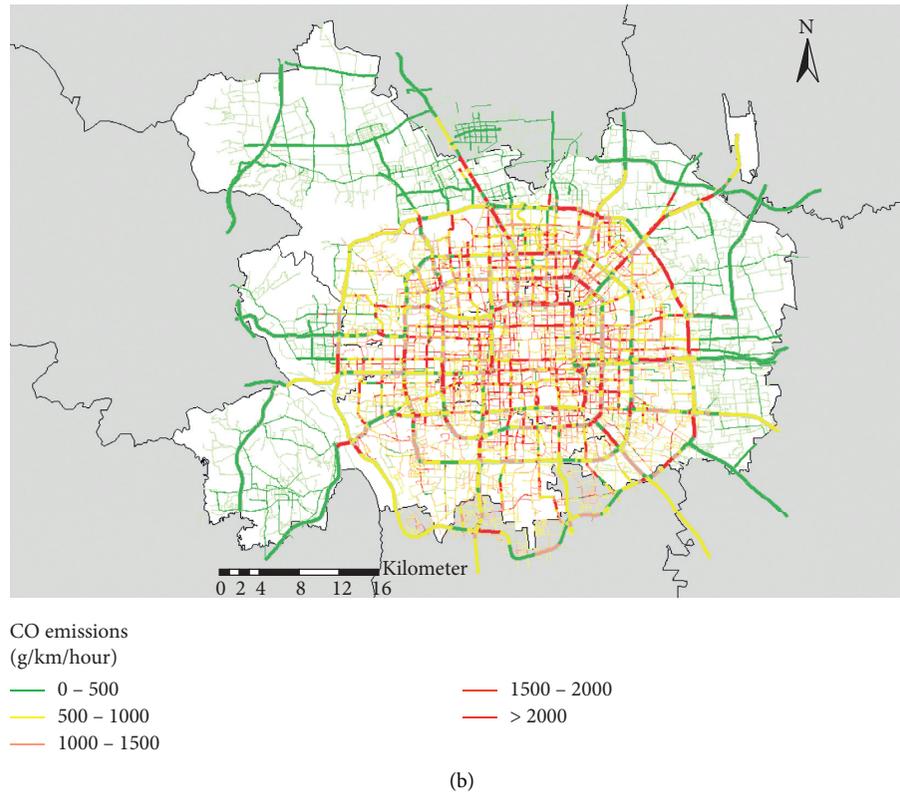


FIGURE 6: Model result for spatial and temporal distribution of flows and emissions in Beijing (8:00–9:00). (a) Traffic flows are estimated from the FCD of different road types. (b) The spatial distribution of CO emissions.

6. Conclusions

Hourly traffic flow estimation from FCD is vital for dynamic estimation of vehicle emissions on the road network, but the uncertainty of hourly flow estimation using a fundamental diagram is often questioned.

To improve the accuracy of traffic flow estimation, this study investigates the existence of hysteresis phenomena on the fundamental diagrams and proposes a multiperiod speed-flow fundamental diagram. Further, the uncertainty of traffic flow estimation using the improved model is quantified, and the impact of aggregating the FCD and traffic flow data to the 1-hour interval is investigated at different speed bins. The main conclusions can be summarized as follows:

- (1) The multiperiod speed-flow fundamental diagram considering the hysteresis phenomena can improve the accuracy of the flow estimation effectively.
- (2) There is no significant difference in magnitudes of the estimation uncertainty at 1 hour compared with that at 10 minutes.
- (3) The variation of the estimation uncertainty is reduced in the process of aggregating the time intervals from 10 minutes to 1 hour, with an average 7% reduction of the range of 95% CI. This reduction of 95% CI is especially significant in the speed bins of ≥ 60 km/h and < 15 km/h, with 22% and 12% respectively. It is

demonstrated that hourly traffic flow estimation presents a better performance compared with that at 10 minutes.

(4) Besides, the traffic flow estimation performs better during congested states (with a higher emission rate) compared with that at the free-flow state (with a lower emission rate). The average and variation of the estimation uncertainty are relatively smaller in the speed bins around the speed at capacity (40 km/h), compared with those in both high- and low-speed bins of ≥ 60 km/h and < 15 km/h.

It should be noted that this research has limitations. First, more data fusion should be collected to validate the accuracy of the proposed method. Second, the study investigates the uncertainty on a ring expressway in Beijing. An uncertainty analysis covering other road types is recommended.

Data Availability

The matched FCD and RTMS data used to support the findings of this study are from the cited literature [41].

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors are thankful to all the personnel who either provided the technical supports or helped in data collection and processing. This research was supported by the National Key R&D Program of China (#2018YFB1600701) and the Natural Science Foundation of China (NSFC) (#71871015).

References

- [1] D. Ji, Y. Wang, L. Wang et al., "Analysis of heavy pollution episodes in selected cities of northern China," *Atmospheric Environment*, vol. 50, pp. 338–348, 2012.
- [2] X. Wang, W. Wei, S. Cheng et al., "Characteristics and classification of PM_{2.5} pollution episodes in Beijing from 2013 to 2015," *Science of The Total Environment*, vol. 612, pp. 170–179, 2018.
- [3] Ministry of Ecology and Environment of the People's Republic of China, *China Mobile Source Environmental Management Annual Report*, Ministry of Ecology and Environment of the China, Beijing, China, 2019.
- [4] J. Edmonds, C. Nichlos, M. Adamantiades et al., "Could congressionally mandated incentives lead to deployment of large-scale CO₂ capture, facilities for enhanced oil recovery CO₂ markets and geologic CO₂ storage," *Energy Policy*, vol. 146, Article ID 111775, 2020.
- [5] Y.-H. Liu, J.-L. Ma, L. Li, X.-F. Lin, W.-J. Xu, and H. Ding, "A high temporal-spatial vehicle emission inventory based on detailed hourly traffic data in a medium-sized city of China," *Environmental Pollution*, vol. 236, pp. 324–333, 2018.
- [6] S. Zhang, T. Niu, Y. Wu et al., "Fine-grained vehicle emission management using intelligent transportation system data," *Environmental Pollution*, vol. 241, pp. 1027–1037, 2018.
- [7] Y. Zhang, L. Wu, C. Zou et al., "Development and application of urban high temporal-spatial resolution vehicle emission inventory model and decision support system," *Environmental Modeling & Assessment*, vol. 22, no. 5, pp. 445–458, 2017.
- [8] J. Liu, K. Han, X. Chen, and G. P. Ong, "Spatial-temporal inference of urban traffic emissions based on taxi trajectories and multi-source urban data," *Transportation Research Part C: Emerging Technologies*, vol. 106, pp. 145–165, 2019.
- [9] A. Günemann, R. P. Schäfer, K. U. Thiessenhusen, and P. Wagner, *Monitoring Traffic and Emissions by Floating Car Data*, Institute of Transport Studies, Leeds LS2 9JT, UK, Working Paper ITS-WP-04-07, 2004.
- [10] K.-H. Bethke, S. Baumgartner, M. Gabele et al., "Air- and spaceborne monitoring of road traffic using SAR moving target indication-Project TRAMRAD," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 61, no. 3-4, pp. 243–259, 2006.
- [11] L. Zhu, H. Wen, and J. Sun, "Floating car based real-time-traffic-info collection system in Beijing," *Urban Transport of China*, vol. 6, no. 1, pp. 77–80, 2008.
- [12] L. Zhu, Y. Hu, J. Sun et al., "A study on transformation practice of real-time probe vehicle system Architecture in Beijing," *Information Technology & Informatization*, vol. 7, pp. 22–25, 2016.
- [13] D. Budimir, N. Jelušić, and M. i Perić, "Floating car data technology," *Pomorstvo*, vol. 33, no. No. 1, pp. 22–32, 2019.
- [14] E. Brockfeld, S. Lorkowski, P. Mieth et al., "Benefits and limits of recent floating car data technology—an evaluation study," in *Proceedings of the Presented at the 11th WCTR Conference*, Berkeley, CA, USA, June 2007.
- [15] O. Altintasi, H. Tuydes-Yaman, and K. Tuncay, "Quality of floating car data (FCD) as a surrogate measure for urban arterial speed," *Canadian Journal of Civil Engineering*, vol. 46, no. 12, pp. 1187–1198, 2019.
- [16] N. Zhao and L. Yu, "Statistical distribution of capacities on ring-road expressways of Beijing," in *Proceedings of the Presented at Transportation Research Board 89th Annual Meeting*, Washington, DC, USA, January 2010.
- [17] C. Sun, H. Zhang, and X. Chen, "Road traffic operation assessment based on multi-source floating car data fusion," *Journal of Tongji University (Natural Science)*, vol. 46, no. No. 1, pp. 46–52, 2018.
- [18] D. F. Llorca, M. A. Sotelo, S. Sánchez et al., "Traffic data collection for floating car data enhancement in V2I networks," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, 13 pages, 2010.
- [19] M. Rahmani, H. N. Koutsopoulos, and E. Jenelius, "Travel time estimation from sparse floating car data with consistent path inference: a fixed point Approach," *Transportation Research Part C: Emerging Technologies*, vol. 85, pp. 628–643, 2017.
- [20] B. Dewulf, T. Neutens, M. Vanlommel et al., "Examining commuting patterns using Floating Car Data and circular statistics: exploring the use of new methods and visualizations to study travel times," *Journal of Transport Geography*, vol. 48, pp. 41–51, 2015.
- [21] S. Axer and B. Friedrich, "Level of service estimation based on low-frequency floating car data," *Transportation Research Procedia*, vol. 3, pp. 1051–1058, 2014.
- [22] L. Xu, Y. Yue, and Q. Li, "Identifying urban traffic congestion pattern from historical floating car data," *Procedia-Social and Behavioral Sciences*, vol. 96, pp. 2084–2095, 2013.
- [23] M. Houbraken, S. Logghe, M. Schreuder et al., "Automated incident detection using real-time floating car data," *Journal of Advanced Transportation*, vol. 2017, 2017.
- [24] A. Sunderrajan, V. Viswanathan, W. Cai, and A. Knoll, "Traffic state estimation using floating car data," *Procedia Computer Science*, vol. 80, 2016.
- [25] M. Houbraken, S. Logghe, P. Audenaert et al., "Examining the potential of floating car data for dynamic traffic management," *IET Intelligent Transport Systems*, vol. 12, no. 5, pp. 335–344, 2018.
- [26] B. S. Kerner, C. Demir, R. G. Herrtwich et al., "Traffic state detection with floating car data in road networks," in *Proceedings of the 2005 IEEE Intelligent Transportation Systems*, pp. 44–49, Vienna, Austria, September 2005.
- [27] A. Talebpour and H. S. Mahmassani, "Influence of connected and autonomous vehicles on traffic flow stability and throughput," *Transportation Research Part C: Emerging Technologies*, vol. 71, pp. 143–163, 2016.
- [28] J. Hu, M. D. Fontaine, and J. Ma, "Quality of private-sector travel time data on arterials," *Journal of Transportation Engineering*, vol. 142, no. No. 4, 2016.
- [29] R. T. Chase, B. M. Williams, N. M. Roupail, and S. Kim, "Comparative evaluation of reported speeds from corresponding fixed-point and probe-based detection systems," *Transportation Research Record*, vol. 2308, no. 1, pp. 110–119, 2012.
- [30] S. Kim and B. Coifman, "Comparing INRIX speed data against concurrent loop detector stations over several months," *Transportation Research Part C: Emerging Technologies*, vol. 49, pp. 59–72, 2014.
- [31] C. Buisson and C. Ladier, "Exploring the impact of homogeneity of traffic measurements on the existence of

- macroscopic fundamental diagrams,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2124, no. 1, pp. 127–136, 2009.
- [32] L. Ambühl and M. Menendez, “Data fusion algorithm for macroscopic fundamental diagram estimation,” *Transportation Research Part C: Emerging Technologies*, vol. 71, pp. 184–197, 2016.
- [33] Y. B. Ji, M. Xu, and J. Li, “Determining the macroscopic fundamental diagram from mixed and partial traffic data,” *Promet-Traffic & Transportation*, vol. 30, no. 3, pp. 267–279, 2018.
- [34] Y. B. Ji, H. Vanzuylen, and S. Lu, “Determining the macroscopic fundamental diagram on the basis of mixed and incomplete traffic data,” in *Proceedings of the Presented at TRB 95th Annual Meeting*, Washington, DC, USA, January 2016.
- [35] K. Anuar, F. Habtemichael, and M. Cetin, “Estimating traffic flow rate on freeways from probe vehicle data and fundamental diagram,” in *Proceedings of the Presented at IEEE 18th International Conference on Intelligent Transportation Systems*, pp. 2921–2926, Las Palmas de Gran Canaria, Spain, September 2015.
- [36] T. Neumann, P. L. Böhneke, and L. C. Touko Tcheumadjeu, “Dynamic representation of the fundamental diagram via bayesian networks for estimating traffic flows from probe vehicle data,” in *Proceedings of the Presented at 16th International IEEE Conference on Intelligent Transportation Systems*, Hague, Netherlands, October 2013.
- [37] A. M. Mulligan and A. Nicholson, “Uncertainty in traffic flow estimation using the moving-observer method,” in *Proceedings of the Presented at the 2002 Meeting of the Institution of Professional Engineers New Zealand (IPENZ) Transportation Group*, Rotorua, New Zealand, September 2002.
- [38] W. Deng, H. Lei, and X. Zhou, “Traffic state estimation and uncertainty quantification based on heterogeneous data sources: a three detector approach,” *Transportation Research Part B: Methodological*, vol. 57, pp. 132–157, 2013.
- [39] G. Davis and S. Yang, “Accounting for uncertainty in estimates of total traffic volume: an empirical Bayes approach,” *Journal of Transportation and Statistics*, vol. 4, no. 1, pp. 27–38, 2001.
- [40] J. Guo, W. Huang, and B. M. Williams, “Adaptive kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification,” *Transportation Research Part C: Emerging Technologies*, vol. 43, no. 1, pp. 50–64, 2014.
- [41] J. Wang, L. Yu, Y. Gao, J. Zhang, and G. Song, “Analysis of speed differences in data from remote traffic microwave sensors and floating car data systems on expressways: Beijing case study,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2643, no. 1, pp. 112–120, 2017.
- [42] G. Song, F. Zhang, J. Liu, L. Yu, Y. Gao, and L. Yu, “FCD based method for detecting flooding incident under grade separation bridges in beijing,” *IET Intelligent Transport Systems*, vol. 9, no. 8, pp. 817–823, 2015.
- [43] S. M. Turner, *Guidelines for Developing ITS Data Archiving Systems*, Report 2127-3, Texas Transportation Institute, Texas A&M University System, College Station, Bryan, TX, USA, 2001.
- [44] N. Zhao, L. Yu, H. Zhao, J. Guo, and H. Wen, “Statistical distribution of capacities on ring-road expressways of Beijing using FCD and remote traffic microwave sensor data,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2124, no. 1, pp. 178–185, 2009.
- [45] N. Geroliminis and J. Sun, “Hysteresis phenomena of a macroscopic fundamental diagram in freeway networks,” *Transportation Research Part A: Policy and Practice*, vol. 45, no. 9, pp. 966–979, 2011.
- [46] K. Wang and D. M. Levinson, “Towards a metropolitan fundamental diagram using travel survey data,” *PLoS One*, vol. 11, no. 2, 2016.
- [47] Z. He, S. He, and W. Guan, “A figure-eight hysteresis pattern in macroscopic fundamental diagrams and its microscopic causes,” *Transportation Letters*, vol. 7, pp. 133–142, 2015.
- [48] M. Amirgholy and H. O. Gao, “Modeling the dynamics of congestion in large urban networks using the macroscopic fundamental diagram: user equilibrium, system optimum, and pricing strategies,” *Transportation Research Part B: Methodological*, vol. 104, pp. 215–237, 2017.
- [49] X. Shi and H. Lin, “Research on the macroscopic fundamental diagram for Shanghai urban expressway network,” *Transportation Research Procedia*, vol. 25, pp. 1300–1316, 2017.
- [50] H. Rakha and B. Crowther, “Comparison and calibration of FRESIM and INTEGRATION steady-state car-following behavior,” *Transportation Research Part A: Policy and Practice*, vol. 37, no. 1, pp. 1–27, 2003.
- [51] H. Rakha and B. Crowther, “Comparison of greenshields, pipes, and van Aerde car-following and traffic stream models,” *Transportation Research Record*, vol. 1802, pp. 248–262, 2002.
- [52] H. Lu, H. Liu, X. Xu et al., “Moves project-level uncertainty analysis: impacts of variability in on-road operating conditions on emission rates,” in *Proceedings of the Presented at Transportation Research Board 98th Annual Meeting*, Washington, DC, USA, January 2019.
- [53] Y. Tang, H. Yin, Z. Huang et al., “Application of cvem emission model in comprehensive transportation planning: a case study on chengdu-chongqing urban agglomeration,” *Transport Research*, vol. 4, no. 4, pp. 31–40, 2018.
- [54] D. Xie, G. Song, J. Guo et al., “Development and application of an online dynamic emission model for traffic networks-A case study of beijing,” in *Proceedings of the Presented at Transportation Research Board 97th Annual Meeting*, Washington, DC, USA, January 2018.