Electronic Research Archive

ERA, 32(1): 317–331.
DOI: 10.3934/era.2024015
Received: 20 October 2023
Revised: 21 November 2023
Accepted: 03 December 2023
Published: 25 December 2023

http://www.aimspress.com/journal/ERA

Research article

Estimation of spatiotemporal travel speed based on probe vehicles in mixed traffic flow

Jongho Kim[1], Woosuk Kim[2], Eunjeong Ko[1], Yong-Shin Kang[3] and Hyungjoo Kim[2,*]

[1] Cho Chun Shik Graduate School of Mobility, Korea Advanced Institute of Science and Technology CCS Graduate School of Mobility, Yuseonggu, Daejeon 193, Republic of Korea
[2] Department of Intelligent Transportation System Lab, Advanced Institute of Convergence Technology, Suwon-si, Gyeonggi-do 16229, Republic of Korea
[3] Industrial Intelligence Laboratory, Advanced Institute of Convergence Technology, Suwon-si, Gyeonggi-do 16229, Republic of Korea

* **Correspondence:** Email: hyungjoo@snu.ac.kr.

Abstract: Conventional fixed traffic detectors are limited to their installed locations and are unable to collect general traffic information or monitor microscopic traffic flows. Mobile detectors overcome spatial constraints by allowing the vehicle to act as a detector and can observe microscopic traffic flows by collecting high-resolution trajectory data from individual vehicles. The objective of this study is to estimate spatiotemporal traffic information based on the autonomous driving sensor headway distance and to calculate the appropriate spatiotemporal interval according to the sampling rate. First, individual vehicle trajectory data was collected, and a traffic information estimation was established. Travel speed was calculated based on generalized definitions, and its estimation and errors were analyzed. In addition, the appropriate spatiotemporal interval according to cell size, time interval, and sampling rate was analyzed. The analysis demonstrated that the estimation accuracy was improved by cell size, time interval, and sampling rate. Based on this, the appropriate time and space to minimize the error rate were calculated considering the sampling rate. When the sampling rate was 40% or more, the error rate was 5% or less in all time and space; however, error rate differences occurred in several cases at sampling rates below 40%. These results are anticipated for efficient management of collecting, processing and providing traffic information.

Keywords: spatiotemporal travel speed; probe vehicle; NGSIM; mixed traffic flow

Abbreviations: GPS: global positioning system; MAPE: mean absolute percentage error; NGSIM: next generation simulation; RMSE: root mean square error

1. Introduction

Traffic conditions need to be accurately perceived in real-time to maintain efficient traffic flow and ensure safety. Travel speed is an important indicator for both road users and network managers because it allows traffic conditions to be intuitively identified and is directly related to traffic congestion. It has a significant impact on road users' choice of travel modes and routes and helps network managers establish policies using real-time monitoring of overall traffic flow and history data.

However, most travel speed information has been collected and generated by fixed detectors based on spot or section measurements. Limitations exist in collecting traffic information for spots or sections where detectors are not installed. Spot measurement calculates the instantaneous speed information of vehicles passing through a specific spot and converts it into the travel speed information of the section unit through the average value of the instantaneous speeds based on simplicity and scalability [1,2]. The traffic information collected through section measurements generates speed information through the time that vehicles take to travel from the start to the end of the section. However, traffic information can be generated only when the section is completely traversed from the start to the end, and travelers receive only macroscopic traffic information as the length of the section increases [3–7]. Therefore, the traffic information collected through fixed detectors generates inaccurate information during traffic congestion [1], suffers from a time lag, and has limitations in generating precise travel speed information at the lane level.

Probe vehicles equipped with global positioning system (GPS) sensors can act as mobile detectors. They can overcome the limitations in areas where there are no fixed detectors because they serve as mobile detectors. However, GPS sensors commercialized at the consumer level have limitations in precise lane-level positioning [8]. Nevertheless, numerous studies have proposed the possibility of precise lane-level positioning by combining GPS and other sensing equipment data [9–12]. In particular, Peng et al. [10] proposed a method of accurately estimating relative trajectories at the lane level through dedicated short-range communication with nearby vehicles using GPS receivers. These results show the possibility of generating spatiotemporal traffic information at a more microscopic level than that generated through fixed detectors. In addition, as autonomous vehicles become more commercially viable, they will share the road with conventional vehicles, and traffic safety issues in the mixed traffic flow can be investigated [13–15]. Additionally, one study designed DGCRIN as a tool to impute missing traffic data [16]. Another study proposed MPGCN, a multi-pattern passenger flow prediction framework based on a graph convolutional network, to analyze detailed traffic patterns and learn knowledge about human mobility [17]. In this study, mixed traffic flow refers to a situation where autonomous vehicles and regular vehicles are mixed, and the probe vehicle was assumed to be an autonomous one. In particular, the frequency of conflicts will increase owing to interactions between vehicles under traffic congestion conditions. Thus, precise traffic information at a microscopic level is required for safe and efficient traffic monitoring under such conditions.

Previous literature on travel speed estimation indicated that spatiotemporal intervals are closely related to the reliability of traffic information, thus making it necessary to calculate appropriate intervals. In numerous studies, appropriate time intervals were calculated; however, spatial intervals were not considered. Studies that use high-resolution traffic data have been conducted, as the data from

autonomous vehicles will be used in addition to GPS data in the mixed traffic flow [18–22]. In such studies, traffic information was estimated by detecting the movement of nearby vehicles using onboard devices such as advanced driver assistance systems, radar, light detection and ranging, and cameras; however, appropriate spatiotemporal intervals could not be calculated.

Additionally, in previous literature, travel speed was calculated based on AI algorithms. This approach can only produce good results in traffic environments such as specific geometry, presence and shape of ramps, traffic volume, and number of lanes. This means that the AI model's results may depend on specific geometry or traffic conditions. If road conditions can be managed using AI techniques on all roads, more advanced traffic management may be possible. However, due to financial limitations, it is only applicable to certain geometries. The generalized definition of Edie can accurately collect road condition information without being limited to road characteristics [23]. Based on Edie research, we described a mechanism for calculating appropriate space-time intervals that can be universally applied to all roads, not limited to specific roads.

Therefore, we aim to investigate different sampling rates of probe vehicles in mixed traffic flow to estimate spatiotemporal travel speed. First, spatiotemporal cell-level traffic information at a microscopic level is defined using GPS sensor data. The travel speed is then estimated, and its accuracy is analyzed according to spatiotemporal intervals. Finally, appropriate spatiotemporal intervals are calculated according to the sampling rate of probe vehicles, depending on estimation accuracy. Additionally, next generation simulation (NGSIM) US-101 data that records actual traffic flow in the form of location information is used and processed to estimate travel speed. The speed information is calculated using the travel duration and distance of vehicles that occupied the preset cell (space-time) area, with the estimation accuracy verified through the mean absolute percentage error (MAPE) and root mean square error (RMSE).

The structure of this article is as follows: Section 2 explains the analysis methods used, including data descriptions, cell-level traffic information definition, travel speed estimation scenario settings, travel speed estimation, and error analysis. Section 3 interprets the results, including travel speed estimation as a function of the sampling rate, cell size, and time interval, as well as appropriate spatiotemporal interval calculations. Finally, in Section 4, the results are summarized and the implications are discussed briefly.

2. Materials and methods

2.1. Data description

The spatial background of the NGSIM data used in this study is the US-101 highway located in Los Angeles, CA, USA, as illustrated in Figures 1 and 2. It has a total analysis length of approximately 640 m and comprises five main lanes and one additional lane that connects to the on- and off-ramps. For the temporal background, 45 min of data from 07:50 to 08:35 on June 15, 2005, were used. The data column included basic vehicle, vehicle movement, and relationship information with other vehicles, as listed in Table 1. The vehicle speed, location and acceleration were collected every 0.1 s. The characteristics of the US-101 dataset include traffic oscillation (referred to as the stop-and-go phenomenon), which negatively affects the overall traffic flow, including traffic congestion, capacity reduction, delays, safety degradation, excessive fuel consumption, and exhaust gas emissions [24,25]. In the verified data, traffic oscillation is repeated at approximately 2 min intervals [26], with an

oscillation wave speed of 16 km/h (4.44 m/s) transferred to the upstream section [26,27]. Therefore, empirical studies related to traffic monitoring in a congested state that reflects more diverse traffic patterns than a free-flow speed state are required [27].

Table 1. NGSIM data description.

Category	Subcategory	Column	Unit
Basic vehicle information	Individual vehicle	Vehicle ID	Number
	Time	Global time	ms
		Frame ID	0.1 s
		Total frames	0.1 s
	Vehicle information	Vehicle length	m
		Vehicle width	Number
		Vehicle class	Text
Vehicle movement information	Location information	Local X	m
		Local Y	m
		Global X	m
		Global Y	m
		Lane identification	Number
	Speed information	Vehicle velocity	m/s
		Vehicle acceleration	m/s²
Relationship information with other vehicles	Front/rear vehicle information	Preceding Vehicle ID	Number
		Following Vehicle ID	Number
	Headway information	Space Headway	m
		Time Headway	s

2.2. Cell-level traffic information definition

In this study, precise cell-level traffic information, in terms of time and space, was defined for the preparation of mixed traffic flow with autonomous vehicles. The vehicle trajectory data in Figure 3 were classified at more microscopic spatiotemporal intervals, with cell-level traffic information generated based on existing vehicle trajectory information in the cell area. For the trajectory information of the travel time and distance of all vehicles moving in the cell area, the traffic volume, density, and travel speed were calculated using the generalized definition of Edie [23], using Eqs (1)–(3), where q (A), k (A), and v (A) are the traffic volume, density, and travel speed, respectively, of the cell A area. $d_n(A)$ and $t_n(A)$ are the travel distance and travel time of vehicle n in the cell area. $N(A)$ is the total number of vehicles that pass through the cell A area. $|A|$ is the range of the cell area, meaning the product of the time interval and the spatial interval of the cell area.

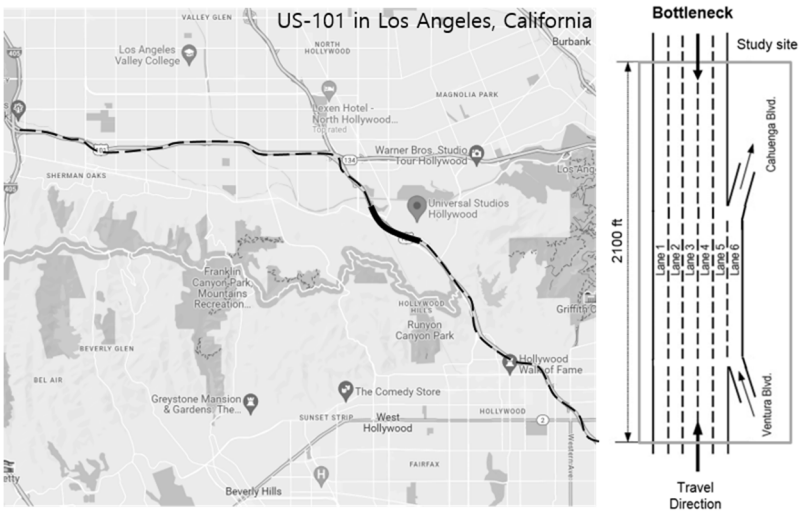


Figure 1. Spatial scope of research.

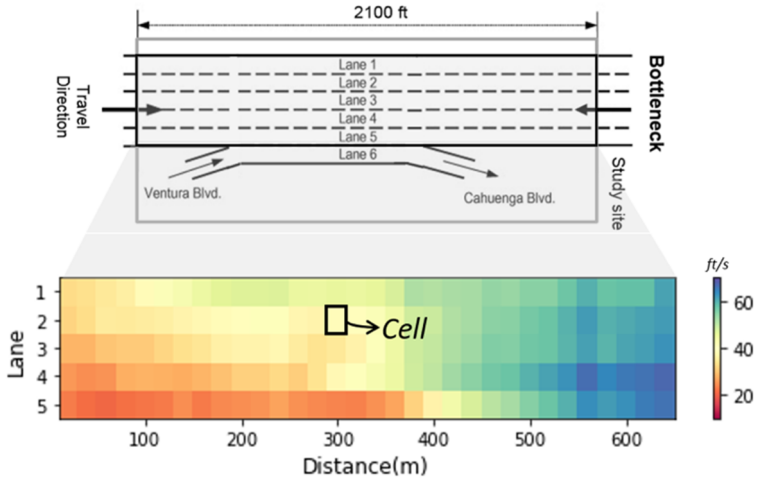


Figure 2. Cell-level traffic information.

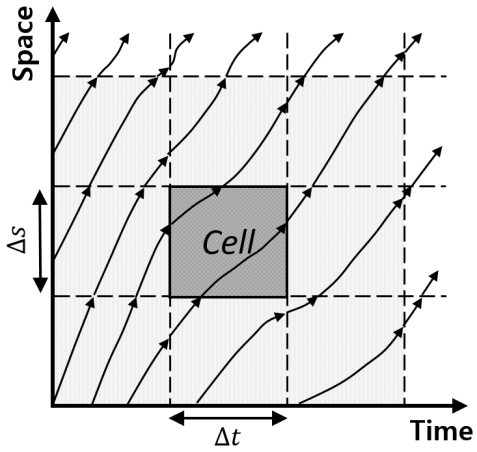


Figure 3. Trajectory data classification.

$$q(A) = \frac{\sum_{n \in N(A)} d_n(A)}{|A|}, \tag{1}$$

$$k(A) = \frac{\sum_{n \in N(A)} t_n(A)}{|A|}, \tag{2}$$

$$v(A) = \frac{\sum_{n \in N(A)} d_n(A)}{\sum_{n \in N(A)} t_n(A)}. \tag{3}$$

In the Mobile Millennium project, an empirical experiment was conducted to calculate road traffic speed based on a probe vehicle equipped with GPS [28]. The study stated that the presence of only 5% of probe vehicles is sufficient to calculate travel speed. However, since the space-time interval was set to 6-mile and 5-min, it was possible to estimate with high accuracy with only 5% of probe vehicles. In other words, if the space-time gap is large enough, it can be estimated based on the actual situation. However, in situations where the space-time interval is low, there is a trade-off relationship where the estimation accuracy drops significantly. Therefore, in this study, to analyze differences in accuracy, they were classified according to space-time intervals in detail.

2.3. Scenario setting

The traffic conditions of all vehicles in one cell area must be made as homogeneous as possible [27]. If they are not homogeneous, the information that road users receive may differ significantly from the actual information, owing to the different travel patterns of each vehicle in the cell area. The degree of homogeneity is related to the size of the cell area, and homogeneous traffic conditions are grouped as the size of the cell area decreases. In this study, the size of space-time was varied, with 144 ($9 \times 4 \times 4 \times 1$) scenarios created, as listed in Table 2. The time interval range was set to 30, 60, 180 and 300 s, while the spatial range was set to 20, 40, 80 and 160 m. In addition, for probe vehicles equipped with GPS sensors, random sampling without replacement was performed by varying the sampling rate from 10–90% in intervals of 10%, with only the travel speed considered for the content range. Here, the traffic volume and density were not considered because they were correlated with each other by the size of the sampling rate and estimation accuracy [29]. The travel speed estimated through probe vehicles was calculated through Eq (4). $P(A)$ indicates a probe vehicle that passes through the cell A area.

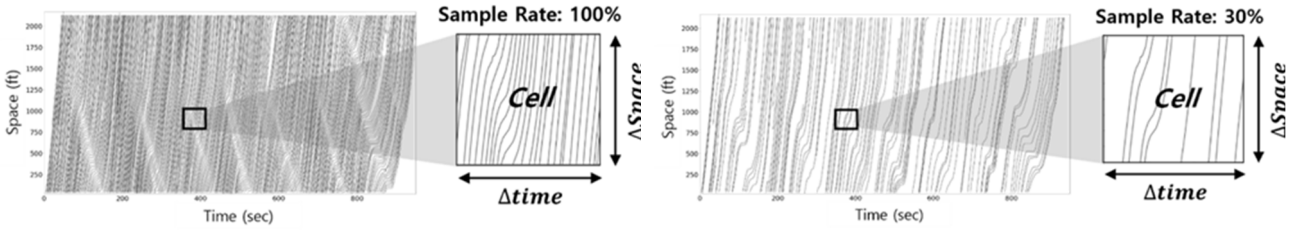
$$v(A) = \frac{\sum_{n \in P(A)} d_n(A)}{\sum_{n \in P(A)} t_n(A)}. \tag{4}$$

Table 2. Indicator definition.

Category	Content	Variable(s)	Variable unit
Data acquisition range	Sampling rate	10–90%	%
Spatial range	Cell	20, 40, 80, 160	m
Temporal range	Time interval	30, 60, 180, 300	s
Content range	Efficiency	Travel speed	

2.4. Error analysis

A large amount of spatiotemporal cell-level traffic information was generated by varying the sampling rate, time interval, and spatial interval in the trajectory information of randomly extracted probe vehicles. The estimation error was calculated by comparing the cells of the same spatiotemporal area with the entire trajectory and sample trajectory data, as illustrated in Figure 4. Figure 4(a) presents the entire vehicle trajectories at a sampling rate of 100%. Figure 4(b) depicts estimated values representing vehicle trajectories at a sampling rate of 30%. In addition, the intervals of $\Delta Space$ and $\Delta Time$ indicate the spatiotemporal intervals of preset cells, with (27,000 s/$\Delta Time$) × (640 m/$\Delta Space$) × 5 (number of lanes) cells generated in a portion of spatiotemporal cell-level traffic information. MAPE and RMSE calculations, as shown in Eqs (5) and (6), were used to analyze the estimation errors of all generated cells. MAPE represents the absolute error between the actual speed and estimated speed as a percentage; a value close to 0% indicates high estimation accuracy without bias. As we could not obtain the actual error value because the absolute error was calculated as a percentage based on the true value, further analysis was conducted using RMSE. RMSE can represent the error from the actual value. Here S_{I} is the actual speed of cell *i*, I is the estimated speed of cell *i,* and N is the total number of cells.



(a) (b)

Figure 4. Spatiotemporal diagram in lane 5 of the US-101: (a) Entire trajectory (sampling rate = 100%); (b) Estimated trajectory (sampling rate = 30%).

$$MAPE = \frac{100}{N}\sum^{N}\left|\frac{S_i-\hat{S}_i}{S_i}\right|, \tag{5}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\hat{S}_i-S_i)^2}{N}}. \tag{6}$$

The section detector calculates travel speed as the difference in travel time between two fixed points. The travel speed is calculated as the ratio of the distance between two points and the time difference between the vehicles passing two points. Moreover, Edie's generalized definition utilized in this study is calculated as the ratio of the time and distance traveled by all vehicles occupied within a specific space-time region. The difference between the two methods is that in the section detector, only vehicles that passed both points are used to calculate the travel speed, and in Edie's generalized definition, vehicles that occupy only a portion of a specific space-time region are also included in the calculation of the travel speed. Additionally, because the section detector is based on the time difference between two points, time-lag occurs. Finally, if the speed is estimated based on the section

detector, the movement trajectory of the vehicle changing lanes will not be fully detected, and the travel speed will be inaccurate in sections where lane changes occur regularly.

3. Results and discussion

In this study, cell-level traffic information was defined, and different scenarios were created by varying the sampling rate and spatiotemporal intervals. The travel speed estimation accuracy for each scenario was calculated using MAPE and RMSE. In this section, (i) the estimation error was analyzed as a function of the sampling rate (%), spatial interval (ΔCell), and time interval (ΔTime), based on the travel speed estimation results; and (ii) appropriate spatiotemporal intervals were calculated based on the estimation error as a function of the sampling rate.

3.1. Travel speed as a function of the sampling rate (%)

The travel speed estimation accuracy improved as the sampling rate of probe vehicles increased, as illustrated in Figure 5 and Table 3. Figure 5 presents the spatiotemporal travel speed estimation results for each sampling rate and true values at ΔCell = 20 m and ΔTime = 180 s. The numbers 1–5 on the x-axis indicate lanes, while the y-axis represents the distance. In addition, the repeated lanes on the x-axis represent the travel speed information that changed with the preset time interval.

Table 3. Estimation errors by sampling rate when ΔTime = 180 s.

Sampling Rate (%)		10	20	30	40	50	60	70	80	90
20 m	MAPE (%)	12.843	8.712	6.548	5.414	4.472	3.500	2.742	2.087	1.292
	RMSE (m/s)	1.415	3.208	2.455	2.073	1.722	1.358	1.068	0.800	0.504
40 m	MAPE (%)	11.641	7.448	5.59	4.717	3.846	2.992	2.368	1.782	1.144
	RMSE (m/s)	1.301	2.807	2.094	1.796	1.493	1.165	0.919	0.683	0.425
80 m	MAPE (%)	9.909	6.193	4.672	4.024	3.296	2.519	1.953	1.503	0.953
	RMSE (m/s)	1.125	2.303	1.738	1.508	1.260	0.979	0.759	0.577	0.349
160 m	MAPE (%)	8.031	4.584	3.516	3.003	2.520	2.043	1.595	1.214	0.767
	RMSE (ft/s)	2.998	1.720	1.322	1.158	0.952	0.768	0.612	0.463	0.278

The travel speed estimation error rate was inversely proportional to the sampling rate and cell size. First, in the case of 20 m cell units, the error rate was 12.84% for a sampling rate of 10%, 4.47% for a sampling rate of 50%, and 1.29% for a sampling rate of 90%. When the cell size was kept constant, the error rate appeared to decrease as the sampling rate increased. In addition, when the sampling rate was 10%, the error rate was 12.843% when the cell size was at 20 m, 11.64% at 40 m, 9.91% at 80 m, and 8.03% at 160 m. When the sampling rate remained constant, the error rate appeared to decrease as the cell size increased. In summary, when the cell unit was 20 m and the sampling rate was 10%, the error rate was the highest at 12.843%, and when the cell unit was 160 m and the sampling rate was 90%, the error rate was the lowest at 0.278%. Therefore, it can be inferred that the accuracy of the travel speed improved with the increasing sampling rate of the probe vehicle and also with increasing cell-unit spacing.

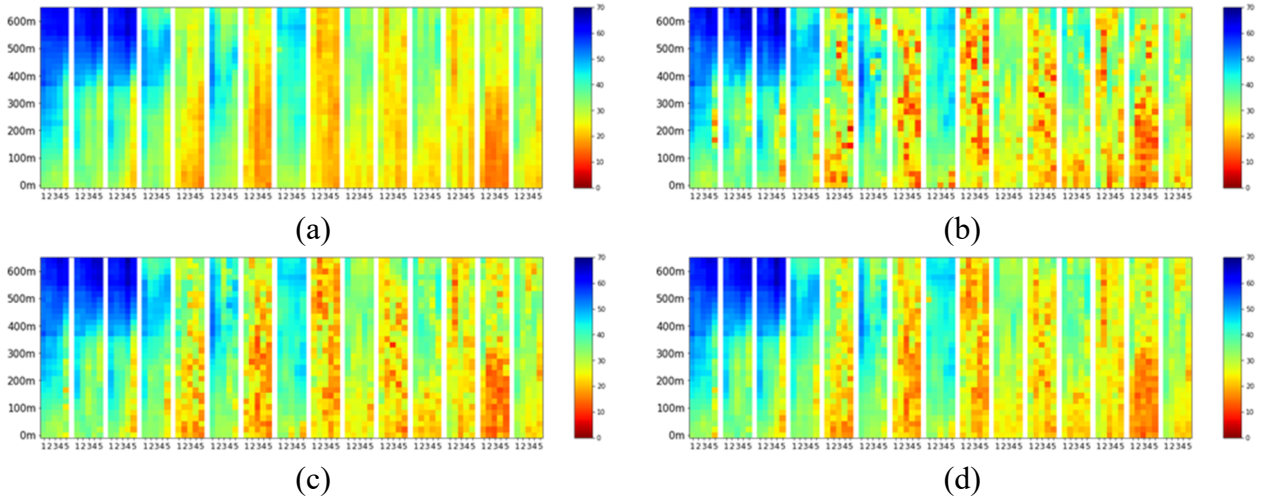


Figure 5. Spatiotemporal travel speed by sampling rate at ΔCell = 20 m and ΔTime = 180 s: (a) Ground truth; (b) Sampling rate 10%; (c) Sampling rate 20%; (d) Sampling rate 30%.

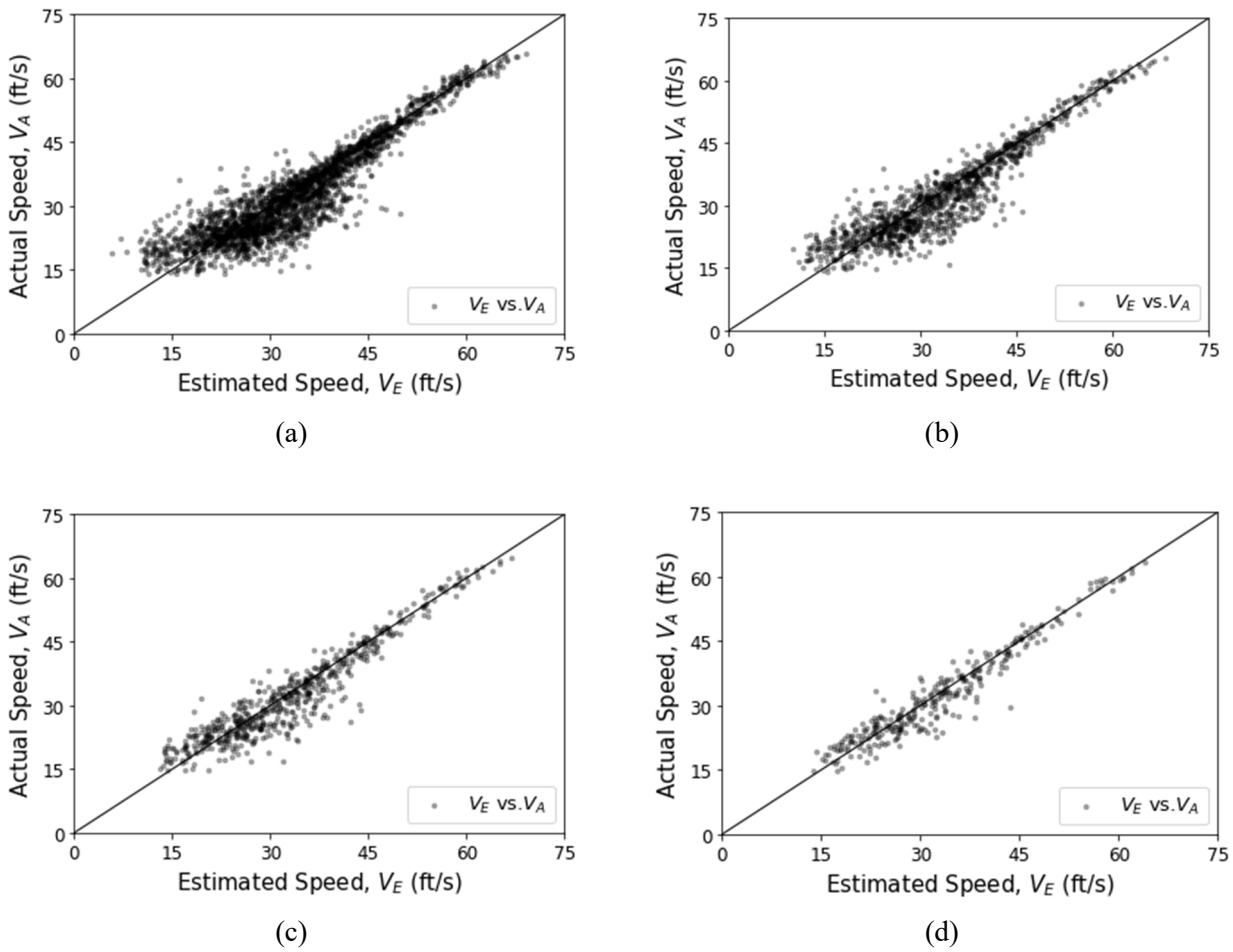


Figure 6. Scatter plot at sampling rate of 10% and ΔTime = 180 s (MAPE, RMSE): (a) ΔCell = 20 m (12.84%, 1.414 m/s); (b) ΔCell = 40 m (11.64%, 1.302 m/s); (c) ΔCell = 80 m (9.91%, 1.125 m/s); (d) ΔCell = 160 m (8.03%, 0.914 m/s).

Figure 6 illustrates a scatter plot of the estimated speed and actual speed when the sampling rate was 10%. When the cell unit was 20, 40, 80, and 160 m, the accuracy was approximately 87.1, 88.3, 90.9 and 91.9%, respectively. The standard deviation of the estimated speed was large; however, it decreased as the cell unit increased in size. Therefore, it can be inferred that the estimation error decreased as the cell unit increased in size.

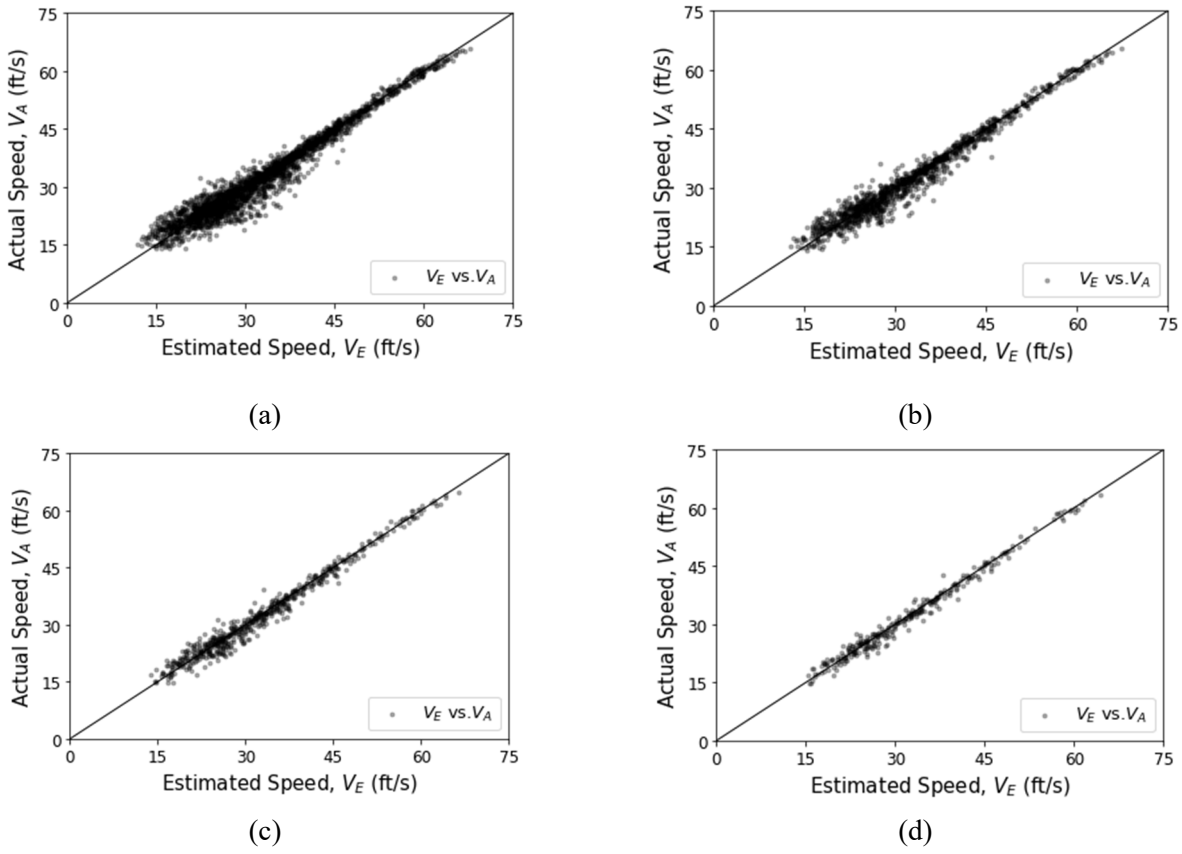


Figure 7. Scatter plot at sampling rate of 30% and ΔTime = 180 s (MAPE, RMSE): (a) ΔCell = 20 m (6.55%, 0.75 m/s); (b) ΔCell = 40 m (5.60%, 0.637 m/s); (c) ΔCell = 80 m (4.67%, 0.53 m/s); (d) ΔCell = 160 m (3.52%, 0.402 m/s).

Figure 7 shows a scatter plot of the estimated speed and actual speed when the sampling rate was 30%. When the cell unit was 20, 40, 80 and 160 m, the accuracy was approximately 93.4, 94.4, 95.3 and 96.4%, respectively. Therefore, it can be inferred that the estimation error decreased as the cell unit increased in size. In addition, the estimation error was found to be within 6.5% at a cell size of 20–160 m.

3.2. Travel speed as a function of time interval

Figure 8 presents the result of estimating travel speed over time. First, as illustrated in Figure 8(a), when the cell unit is 20 m, the travel speed estimation error is inversely proportional to the time interval. Additionally, when the sampling rate was 30–90%, the error was found to be less than 10% regardless of the time interval. It was classified into Mα with an MAPE of 10% or more at 30, 60, 180 and 300 s

and a sampling rate of 10%, and at 30 and 60 s and a sampling rate of 20%. Mβ was classified as having an MAPE of 5–10% at 180 and 300 s and a sampling rate of 20%, at 30, 60, 180 and 300 s and a sampling rate of 30% at 30, 60, and 180 s with a sampling rate of 40%. MAPE was found to be Mγ when less than 5% at 300 s at a sampling rate of 40% and 30, 60, 180 and 300 s at a sampling rate of 50–100%. As illustrated in Figure 8(b), when the sample cell unit is 40 m, the travel speed error rate was found to be 10% or less regardless of time, except for 30 s at a sampling rate of 20–90%. Mα was analyzed at a sampling rate of 10% at 30, 60, 180 and 300 s and at 20% at 30 s. Mβ was found for 60, 180 and 300 s with a sampling rate of 20%, 30, 60, 180 and 300 s with a sampling rate of 30%, and 30 and 60 s with a sampling rate of 40%. 180 and 300 s with a sampling rate of 40% and 30, 60, 180 and 300 s with a sampling rate of 50 to 100% were investigated as Mγ. As in Figure 8(c),(d), when the cell unit was 80 and 160 m, the error rate was Mβ and Mγ of less than 10%, except for 30, and 60 s with a sampling rate of 10%. In summary, when the sampling rate was 30–90%, the estimation error was < 10% for all time intervals, and when the sampling rate was 10–20%, the estimation error was > 10%, and the variation in estimation error over time was large. It can be inferred that as the cell unit increased, the effect of sampling rate and time on the error rate decreased.

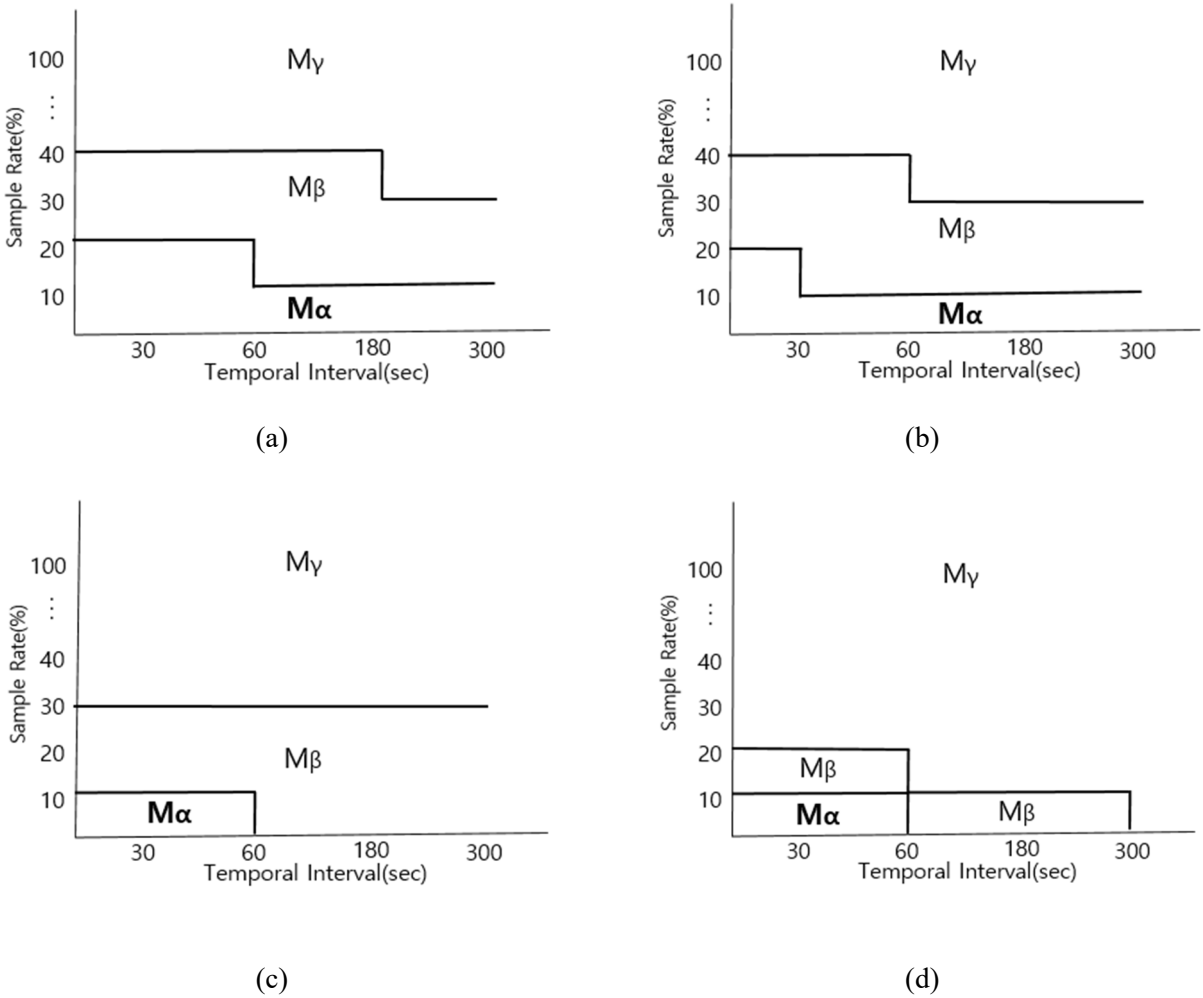
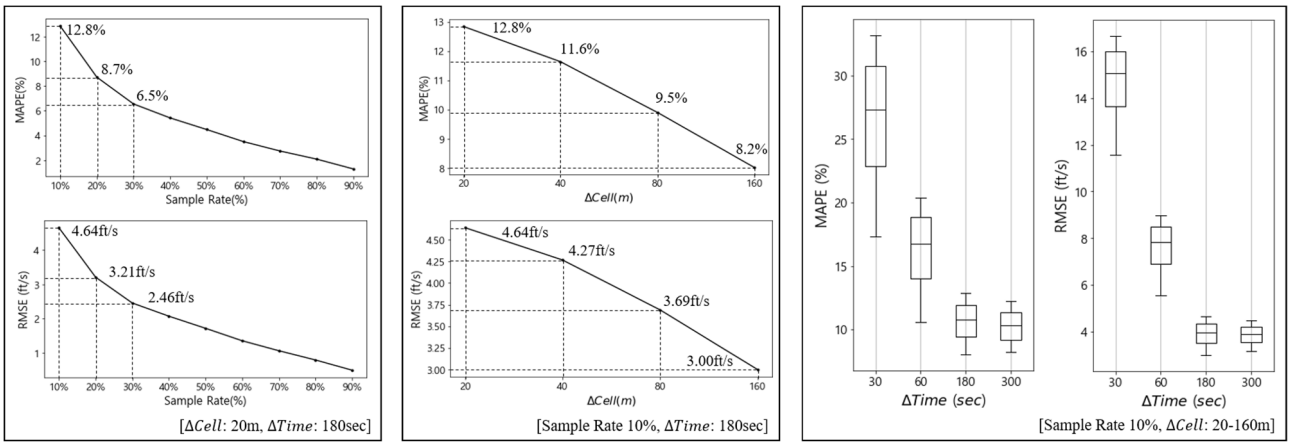


Figure 8. Traffic speed estimation results as a function of time interval for each cell unit: (a) ΔCell = 20 m; (b) ΔCell = 40 m; (c) ΔCell = 80 m; (d) ΔCell = 160 m. (Mα: MAPE > 10%, Mβ: MAPE 5–10%, Mγ: MAPE < 5%)

3.3. Calculation of spatiotemporal intervals

The estimation accuracy results were analyzed by sampling rate, cell size, and time interval, as shown in Figure 9, which demonstrates that the MAPE decreased as the sampling rate increased. In addition, the RMSE decreased as the sampling rate increased. Figure 9(b) shows the error rate as a function of cell unit size, which shows that the MAPE decreases as the cell unit size increases. In addition, the RMSE was found to decrease as the cell unit size increased. Figure 9(c) is a box plot of the MAPE and RMSE for different time intervals, and it can be inferred that the accuracy of the estimation increased as the time interval increased. In summary, the accuracy of the estimation increased as the time interval, cell size, and sampling rate increased.



(a) (b) (c)

Figure 9. MAPE and RSME changes as a function of sampling rate, cell size, and time interval: (a) MAPE, RMSE as a function of sampling rate; (b) MAPE, RMSE as a function of cell size; (c) MAPE, RMSE as a function of time interval.

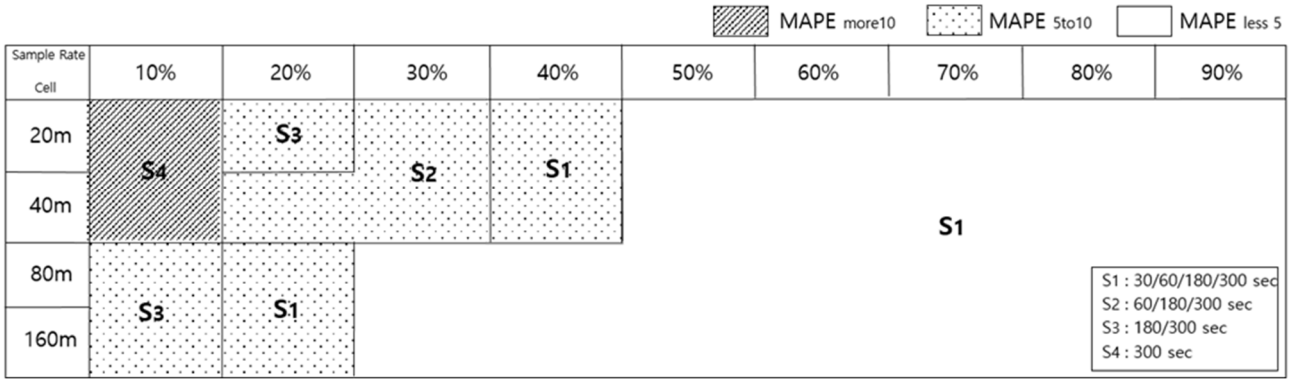


Figure 10. Appropriate space-time interval according to sampling rate.

Figure 10 shows the estimation accuracy as a function of sampling rate, cell unit size, and time interval, categorized as good (< 5%), average (5–10%), or bad (> 10%). Additionally, S1 was classified into 30, 60, 180 and 300 s; S2 was classified into 60, 180 and 300 s; S3 was classified into 180, 300 s

and S4 was classified into 300 s. First, 80 and 160 m with a sampling rate > 50% or 30–40% were classified as good and corresponded to S1. Cell sizes of 20 and 40 m with a sampling rate of 10% were classified as poor and corresponded to S4. The remaining sections were classified as normal and fell into S1, S2 and S3. In summary, when the sampling rate was 10–20%, the estimation error was classified as poor or average depending on the time interval and cell size. When the sampling rate was 30–40%, the estimation error was classified as average or good, depending on the time interval and cell size. For sampling rates > 50%, the estimation error was classified as good, depending on the time interval and cell size. Therefore, in this study, a sampling rate of 10%, a cell size of 80 or 160 m, and a time interval of 180 or 300 s/cell were judged to be appropriate space-time intervals. In addition, when the sampling rate was 20–40%, the appropriate space-time interval depending on the section is as shown in Figure 10, and when it is 50% or more, it is inferred that all space-time intervals are appropriate.

4. Conclusions

The results of this study can be summarized as follows: First, cell-level traffic information in mixed traffic flow considering lanes, unit sections, and time intervals was defined using GPS sensor data. Second, low estimation accuracy was observed at low sampling rates because the value was overestimated or underestimated owing to the insufficient amount of data in the cell area. However, when the cell size and spatiotemporal intervals were large, a high estimation accuracy of 92.8% or higher was confirmed from a sampling rate of 10%. Third, appropriate spatiotemporal intervals were calculated considering the sampling rate, time interval, cell size, and estimation error. Based on these experimental results, the possibility of estimating spatiotemporal cell-level traffic information using the location information of probe vehicles equipped with GPS sensors was presented. This information promises to be at a more microscopic level than the traffic information of the conventional node/link system. These findings will contribute to efficient and safe traffic control operations in mixed traffic flow with conventional and autonomous vehicles in the future. Microscopic traffic information is needed for vehicles to travel safely and quickly on the road. Therefore, results on the appropriate space-time interval with high accuracy to estimate cell-level traffic speed can provide necessary information to vehicles and promote efficient operation by managers. In this study, the estimation accuracy was significantly varied depending on the spatiotemporal intervals at low sampling rates (10–30%) because the travel speed was estimated using the generalized travel speed. Therefore, further studies should investigate measures to improve the estimation accuracy from a low sampling rate through the travel speed calculated by applying different algorithms.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work was supported by the Korea Institute of Police Technology (KIPoT) grant funded by the Korea government (KNPA) (No. 092021C28S01000, Development of integrated road traffic control system and operation technology when autonomous driving is mixed with normal vehicles).

Conflict of interest

The authors declare that there are no conflicts of interest.

References

1. H. Kim, Y. Kim, K. Jang, Systematic relation of estimated travel speed and actual travel speed, *IEEE Trans. Intell. Transp. Syst.*, **18** (2017), 2780–2789. https://doi.org/10.1109/tits.2017.2713983
2. G. Kloot, Melbourne's arterial travel time system, in *International Conference of Australia, 4th*, Adelaide, South Australia, 1999.
3. R. J. Haseman, J. S. Wasson, D. M. Bullock, Real-time measurement of travel time delay in work zones and evaluation metrics using Bluetooth probe tracking, *Transp. Res. Rec*., **2169** (2010), 40–53. https://doi.org/10.3141/2169-05
4. S. Gao, I. Chabini, Optimal routing policy problems in stochastic time-dependent networks, *Transp. Res. Part B Methodol.*, **40** (2006), 93–122. https://doi.org/10.1016/j.trb.2005.02.001
5. D. D. Puckett, M. J. Vickich, Bluetooth-based travel time speed measuring systems development, No. UTCM 09-00-17, Texas Transportation Institute, University Transportation Center for Mobility, 2010.
6. A. Haghani, M. Hamedi, K. F. Sadabadi, S. Young, P. Tarnoff, Data collection of freeway travel time ground truth with Bluetooth sensors, *Transp. Res. Rec. J. Transp. Res. Board*, **2160** (2010), 60–68. https://doi.org/10.3141/2160-07
7. S. Carrese, E. Cipriani, U. Crisalli, A. Gemma, L. Mannini, Bluetooth traffic data for urban travel time forecast, *Transp. Res. Procedia*, **52** (2021), 236–243. https://doi.org/10.1016/j.trpro.2021.01.027
8. M. G. Wing, A. Eklund, L. D. Kellogg, Consumer-grade global positioning system (GPS) accuracy and reliability, *J. For.*, **103** (2005), 169–173. https://doi.org/10.1093/jof/103.4.169
9. J. Du, M. J. Barth, Next-generation automated vehicle location systems: Positioning at the lane level, *IEEE Trans. Intell. Transp. Syst.*, **9** (2008), 48–57. https://doi.org/10.1109/tits.2007.908141
10. Z. Peng, S. Hussain, M. I. Hayee, M. Donath, Acquisition of relative trajectories of surrounding vehicles using GPS and SRC based V2V communication with lane level resolution, in *Proceedings of 3rd International Conference on Vehicle Technology and Intelligent Transport Systems*, (2017), 242–251. https://doi.org/10.5220/0006304202420251
11. Y. S. Li, W. B. Zhang, X. W. Ji, C. X. Ren, J. Wu, Research on lane a compensation method based on multi-sensor fusion, *Sensors*, **19** (2019), 1584. https://doi.org/10.3390/s19071584
12. J. M. Kang, T. S. Yoon, E. Kim, J. B. Park, Lane-level map-matching method for vehicle localization using GPS and camera on a high-definition map, *Sensors*, **20** (2020), 2166. https://doi.org/10.3390/s20082166.
13. J. Kim, D. Lim, Y. Seo, J. J. So, H. Kim, Influence of dedicated lanes for connected and automated vehicles on highway traffic flow, *IET Intell. Transp. Syst.*, **17** (2022), 678–690. https://doi.org/10.1049/itr2.12295
14. H. P. Yu, S. H. Tak, M. J. Park, H. S. Yeo, Impact of autonomous-vehicle-only lanes in mixed traffic conditions, *J. Transp. Res. Board, Transp. Res. Rec.*, **2673** (2019), 430–439. https://doi.org/10.1177/0361198119847475

15. S. L. Lee, C. Oh, S. M. Hong, Exploring lane change safety issues for manually driven vehicles in vehicle platooning environments, *IET Intell. Transp. Syst.*, **12** (2018), 1142–1147. https://doi.org/10.1049/iet-its.2018.5167.
16. J. C. Herrera, D. B. Work, R. Herring, X. J. Ban, Q. Jacobson, A. M. Bayen, Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment, *Transp. Res. Part C Emerg. Technol.*, **18** (2010), 568–583. https://doi.org/10.1016/j.trc.2009.10.006
17. X. Kong, W. Zhou, G. Shen, W. Zhang, N. Liu, Y. Yang, Dynamic graph convolutional recurrent imputation network for spatiotemporal traffic missing data, *Knowl.-Based Syst.*, **261** (2023), 110188. https://doi.org/10.1016/j.knosys.2022.110188
18. S. He, X. Guo, F. Ding, Y. Qi, T. Chen, Freeway traffic speed estimation of mixed traffic using data from connected and autonomous vehicles with a low penetration rate, *J. Adv. Transp.*, **2020** (2020), 1–13. https://doi.org/10.1155/2020/1361583
19. A. Elfar, C. Xavier, A. Talebpour, H. S. Mahmassani, Traffic shockwave detection in a connected environment using the speed distribution of individual vehicles, *Transp. Res. Rec.: J. Transp. Res. Board*, **2672** (2018), 203–214. https://doi.org/10.1177/0361198118794717
20. W. Ma, S. Qian, High-resolution traffic sensing with probe autonomous vehicles: a data-driven approach, *Sensors*, **21** (2021), 464. https://doi.org/10.3390/s2 1020464
21. D. Lim, Y. Seo, E. Ko, J. So, H. Kim, Spatiotemporal traffic density estimation based on ADAS probe data, *J. Adv. Transp.*, **2022** (2022), 5929725. https://doi.org/10.1155/2022/5929725
22. H. K. Kim, Y. Chung, M. Kim, Effect of enhanced ADAS camera capability on traffic state estimation, *Sensors*, **21** (2021), 1996. https://doi.org/10.3390/s21061996
23. Z. He, Y. Lv, L. Lu, W. Guan, Constructing spatiotemporal speed contour diagrams: using rectangular or non-rectangular parallelogram cells?, *Transp. B Transp. Dyn.*, **7** (2017), 44–60. https://doi.org/10.1080/21680566.2017.1320774
24. H. Yao, Q. Li, X. Li, A study of relationships in traffic oscillation features based on field experiments, *Transp. Res. Part A Policy Pract.*, **141** (2020), 339–355. https://doi.org/10.1016/j.tra.2020.09.006
25. R. L. Bertini, T. L. Monica, Empirical study of traffic features at a freeway lane drop, *J. Transp. Eng.*, **131** (2005), 397–407. https://doi.org/10.1061/(asce)0733-947x(2005)131:6(397)
26. J. A. Laval, L. Leclercq, A mechanism to describe the formation and propagation of stop-and-go waves in congested freeway traffic, *Phil. Trans. R. Soc. A*, **368** (2010), 4519–4541. https://doi.org/10.1098/rsta.2010.0138
27. Z. He. L. Zheng, W. Guan, A simple nonparametric-following model driven by field data, *Transp. Res. Part B Methodol.*, **80** (2015), 185–201. https://doi.org/10.1016/j.trb.2015.07.010
28. T. Seo, T. Kusakabe, Y. Asakura, Estimation of flow and density using probe vehicles with spacing measurement equipment, *Transp. Res. Part C Emerg. Technol.*, **53** (2015), 134–150. https://doi.org/10.1016/j.trc.2015.01.033
29. L. C. Edie, *Discussion of Traffic Stream Measurements and Definitions*, New York: Port of New York Authority, 1963

AIMS Press