

## Research Article

# Microscopic Traffic Simulation Calibration Level for Reliable Estimation of Vehicle Emissions

Jinsoo Kim,<sup>1,2</sup> Jae Hun Kim,<sup>3</sup> Gunwoo Lee ,<sup>3</sup> Hyun-ju Shin,<sup>3</sup> and Jahng Hyon Park<sup>2</sup>

<sup>1</sup>Department of Mechanical Convergence Engineering, Graduate School of Hanyang University, 222 Wangsimni-ro, Seongdong-Gu, Seoul 133-791, Republic of Korea

<sup>2</sup>Department of Automotive Engineering, Hanyang University, 222 Wangsimni-ro, Seongdong-Gu, Seoul 133-791, Republic of Korea

<sup>3</sup>Department of International Logistics, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul 06974, Republic of Korea

Correspondence should be addressed to Gunwoo Lee; [gunwoo@cau.ac.kr](mailto:gunwoo@cau.ac.kr)

Received 9 August 2019; Revised 16 October 2019; Accepted 18 November 2019; Published 23 January 2020

Guest Editor: Rongjie Yu

Copyright © 2020 Jinsoo Kim 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.

Vehicle emissions are largely determined by the details of driving behaviours. Accordingly, emissions are often estimated by integrating micro-scale emission models into traffic simulations. Under this approach, it is essential to replicate the actual traffic situation being considered in an emission evaluation using a proper calibration procedure. Most previous research with respect to traffic flow has primarily focused on adjusting the complex combinations of parameters evaluated in these models, but it is not guaranteed that the use of widely used calibration measures can lead more accurate emissions estimates. Accordingly, we propose a systematic guideline for calibration to ensure reliable micro-scale emissions estimates. A calibration procedure is thus established in this paper based on various measure of effect (MOE) compositions (i.e., calibration levels) consisting of aggregated traffic data to identify the level that most reliably estimates micro-scale emissions. Five calibration levels of progressively more detailed measurements are first defined, valid calibration levels are identified, and the reliable calibration level is finally selected based on the available traffic data. The effect of vehicle type (i.e., light vs. heavy vehicles) composition on the estimated emissions is also evaluated for a well-calibrated simulation. We expect that a highly reliable estimation of emissions is possible using this more detailed traffic simulation calibration measurement.

## 1. Introduction

As the number of vehicles on roads increases, the issue of vehicle emissions has grown in importance. To address such emissions-related concerns, it is necessary to establish appropriate emission management strategies by estimating and monitoring vehicle emissions. It has been widely established that vehicle emissions are primarily determined by the details of driving behaviours such as driving speed, acceleration pattern, vehicle type (e.g., heavy vehicles), and responses to weather and road geometry [1, 2]. Among these factors, vehicle emissions are most effected by dynamic driving behaviour features such as frequent acceleration/deceleration, very high or low speeds, idling at stop, and sudden stops due to congestion or anxious driving [2, 3]. As the ability to simulate the microscopic features of vehicle behaviour in a wide range of road networks has become more sophisticated, the estimation of emissions based on simulations is positioned to better

evaluate the impact of the real-world traffic behaviours on emissions and more completely inform the development and management of traffic operation strategies [4, 5].

To accurately capture the effects of dynamic driving behaviours on emissions, micro-scale emission models such as VERSIT+, VT-micro, CMEM, PHEM and MOVES have been found to be more suitable than macro-scale emission models [6–10]. To describe driving profiles and vehicle characteristics in these micro-scale emission models, an extensive set of parameters are required, primarily consisting of vehicle speed profiles that include trajectory data, model year, and fuel type [3]. Among these previously developed models, MOVES contains an operating mode (OpMode) approach that measures emissions by capturing the state of vehicle engines under specific conditions such as braking, idling, coasting, and cruising/accelerating within various speed ranges and vehicle-specific powers (VSPs) using second-by-second trajectory data [11]. These features enable a more detailed estimation of vehicle

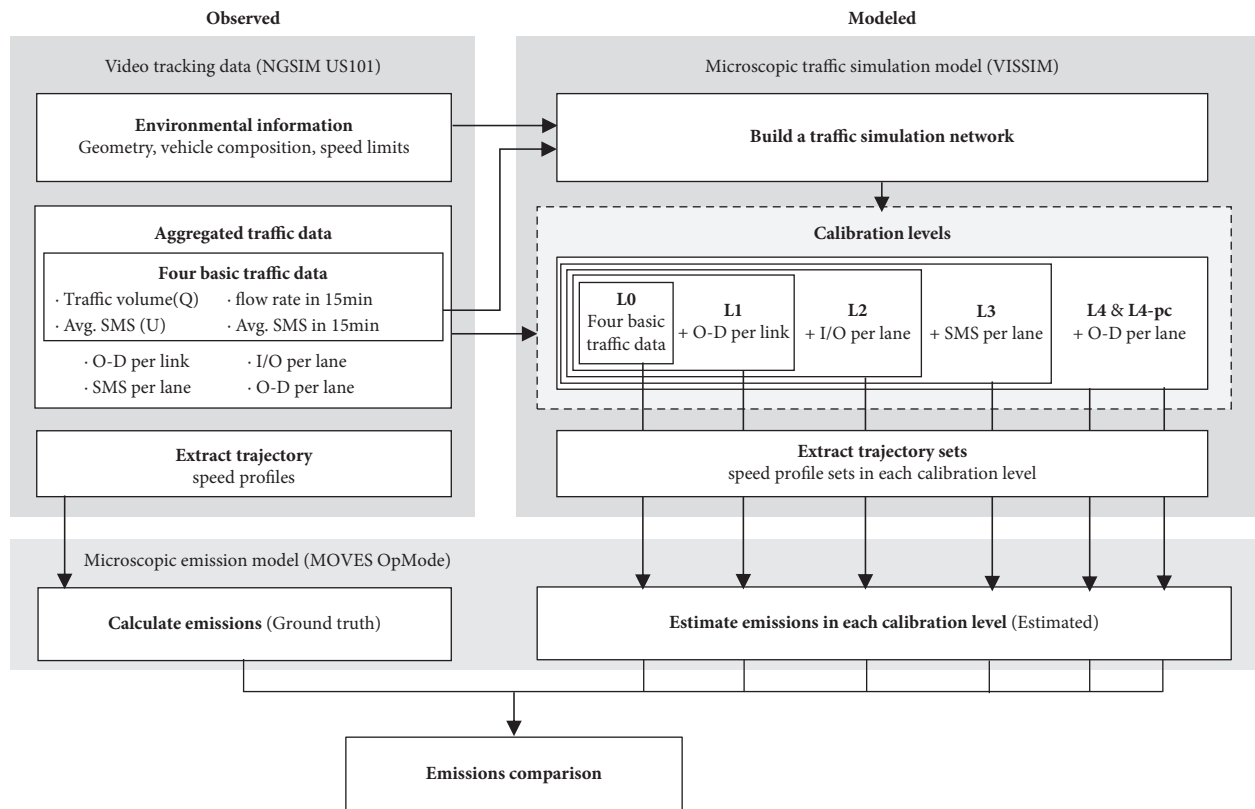


FIGURE 1: Overall framework of study to evaluate the effects of traffic simulation calibration level on micro-scale vehicle emissions estimation.

emissions, but it remains difficult to obtain such trajectory data for an entire vehicle within a given analysis section, so virtual trajectories based on micro-scale traffic simulations are typically used.

Early studies of emissions models focused on determining the differences between micro- and macro-scale emissions models using various types of traffic simulations [6–9]. Recently, there have been some attempts to integrate emission models into traffic simulation models by developing add-on modules based on statistical emission models (e.g., the VERSIT+ based EnViVer modules in the VISSIM software) [12–19]. These studies attempted to determine the sensitivity of emission estimates to road type, geometry, and vehicle type using emission-model-integrated traffic simulations. Though these traffic simulation models dealt with micro-scale emissions, most were conducted under only free-flow traffic conditions in which the impact of the micro-scale emissions model was insufficient [20], and only a few studies discussed a simulation calibration method.

The objective of the calibration is to fine-tune the parameters of the traffic model so that the discrepancy between the observed and simulated traffic flow is minimized. Most studies of microscopic traffic simulation calibration have focused on methods for adjusting complex combinations of behavioural parameters using optimization algorithms such as genetic algorithms (GA) rather than investigating calibration measures of effect (MOEs) [21–25]. Traffic engineers generally use aggregated traffic data as the MOEs such as volume, speed, travel time, and origin/destination volume for calibration of

their models in order to replicate the actual traffic flows [24]. Though these basic MOEs seem to be sufficient to obtain realistic macroscopic traffic flow characteristics (i.e., capacity, average speed, and queuing and travel times), previous studies using them could not guarantee the sufficiently realistic microscopic details of vehicle characteristics (i.e., speed and acceleration) necessary to accurately predict emissions. In order to obtain valid emissions estimates, the simulated vehicle trajectories in the entire network should be similar to the actual trajectories of vehicles in the real world, especially under congested and complex traffic conditions such as weaving sections [26, 27]. In this respect, more effective calibration procedures and MOE compositions need to be studied.

Accordingly, there has been an attempt using raw sample of vehicle trajectory as a calibration MOE for estimating micro-scale emissions [20], showing that such detailed MOEs (i.e., trajectory-based speeds and accelerations) are quite effective in producing models with realistic traffic conditions that reliably estimate emissions. However, the research was conducted under uncongested conditions, and also primarily focused on investigating the calibration parameters themselves. Furthermore, such a trajectory-based calibration approach appears to be difficult to apply in situations in which infrastructure-based traffic data is only available. In other words, it remains essential to investigate methods for capturing the actual traffic situation by utilizing the aggregated traffic data to ensure a realistic micro-scale emissions estimate. As detection technology on a road has continued to develop, it has become possible to obtain more detailed traffic data from

road infrastructure [28–30]. In this regard, it is expected that a more elaborate calibration ensuring reliable emissions estimates can be conducted by utilising data aggregated at the infrastructure-based level.

In this study, we propose a systematic guideline for the calibration of reliable micro-scale emissions estimations. By introducing various MOEs consisting of aggregated traffic data, we establish a calibration procedure and MOE compositions (i.e., calibration levels) to identify the level most capable of reliably estimating emissions. In order to clearly demonstrate the relationship between the micro-scale emission model and the simulation calibration, we focused on a congested weaving section of a highway using video tracking data provided by the Next Generation Simulation (NGSIM) [31]. Vehicle emissions were investigated by applying datasets from NGSIM and VISSIM to a project-level MOVES emissions model. In addition, the effects of heavy vehicle composition on the simulation of estimated emissions were evaluated. Note that we have not attempted to propose any optimization algorithms, but have instead focused on the development of a practical calibration procedure by evaluating a method by which calibration can be conducted using infrastructure-based MOEs to demonstrate general application and transplantability of our work.

The remainder of this paper is organized as follows: Section 2 describes the framework and methodology, Section 3 provides the parameters and results of a micro-scale traffic simulation focusing on calibration, Section 4 evaluates the estimated emissions according to each calibration level, and Section 5 summarizes the conclusions of this study.

## 2. Framework

In this study, we identified the micro-scale traffic simulation calibration level required to minimize the discrepancy between actual and estimated vehicle emissions. The overall framework of the study is shown in Figure 1. First, the environmental information (e.g., geometry, vehicle composition, and speed limit) and eight types of aggregated traffic data (total traffic volume, or  $Q$ , in  $v/h$ ; 15-min flow rate in  $v/15$  min; average space mean speed (SMS), or  $U$ , in  $kph$ ; 15-min SMS in  $kph$ ; origin–destination pair (O–D) per link in  $v/15$  min; input/output volume (I/O) per lane in  $v/15$  min; SMS per lane in  $kph$ ; and O–D per lane in  $v/15$  min) were obtained from NGSIM video tracking data. The aggregated traffic data and vehicle trajectories (i.e., raw data) were taken as the baseline “observed” real-world data. Based on environmental information and four basic aggregated traffic data types (i.e., total traffic volume, 15-min flow rate, average SMS, and 15-min SMS), we built a traffic simulation network using VISSIM software.

After build the network, same types of the aggregated traffic data in accordance with calibration level were exported from the network for comparing with the real-world traffic (i.e., calibrating the simulation). Namely, the eight types of aggregated traffic data were utilized for calibration MOEs. The calibration procedures are conducted by setting the calibration levels to one of five combinations of the MOEs, labelled L0–L4, in which the total MOEs increased with calibration level. The

calibration levels evaluated in this study are defined as follows:

- (i) L0 (default level): Match the four basic aggregated traffic data types (i.e., four MOEs).
- (ii) L1: Satisfy L0 and also match O–D per link (i.e., five MOEs).
- (iii) L2: Satisfy L1 and also match I/O per lane (i.e., six MOEs).
- (iv) L3: Satisfy L2 and also match SMS per lane (i.e., seven MOEs).
- (v) L4: Satisfy L3 and also match O–D per lane (i.e., eight MOEs).

Generally, total traffic volume, 15-min flow rate, average SMS, and 15-min SMS (i.e., the four basic types of aggregated traffic data) have been widely utilized as calibration MOEs in the traffic flow simulation field. Accordingly, we established L0 with these four basic MOEs to serve as the default level that must be satisfied under all other calibration levels. However, even if the simulation calibration is conducted such that the performances of all four basic MOEs show 0% error, it is difficult to be sure that vehicle movements within the entire network are close to those in the real-world. As a result, we adopted four additional MOEs to evaluate their effects on calibration improvement. If all eight MOEs were to be evaluated in all possible combinations, too many combinations would be generated, making it difficult to meaningfully compare the outcome of each calibration.

Therefore, in this study, as the calibration level increases, one additional MOE is introduced to those in the previous levels. The additional MOEs are less aggregated as the calibration level increases, so it is expected that more sophisticated simulation calibration can be conducted, resulting in a simulated traffic flow that is increasingly similar to that in the real world. However, obtaining these decreasingly aggregated data types in the real-world is very difficult. For example, the O–D per lane data in L4 is the most difficult to obtain, as it must be determined by refining the entire vehicle trajectories. In this respect, the five calibration levels evaluated in this study were established not only to reflect an increasingly involved step-by-step calibration procedure (i.e., order of calibration) but also to determine the most reasonable calibration level (i.e., total composition of calibration MOEs) for reliable emissions estimation. In each level of calibration, the corresponding MOEs within the entire network from simulated VISSIM data were compared to those from observed NGSIM data were compared, then the calibration parameters were iteratively adjusted until all performance measurements of all MOEs in each calibration level were within their target ranges. Further details describing the order of calibration, performance measurements, and target ranges according to each MOE are discussed in Section 3.3.

Once each calibration level was complete, we extracted the vehicle trajectory sets. The generated vehicle trajectory sets from NGSIM and VISSIM were input into the MOVES model using OpMode approach, which is able to analyse vehicle emissions in micro-scale. Vehicle emissions were thus estimated according to calibration level by applying each

TABLE 1: Vehicle type composition in NGSIM data.

Vehicle type	1 <sup>st</sup> 15-min interval		2 <sup>nd</sup> 15-min interval		3 <sup>rd</sup> 15-min interval	
	Vehicles	Percentage	Vehicles	Percentage	Vehicles	Percentage
Motorcycle (Gasoline)	30	1.4%	10	0.5%	5	0.3%
Passenger car (Gasoline)	2,086	96.2%	1,963	97.3%	1,870	97.6%
Truck and bus (Diesel)	53	2.4%	44	2.2%	40	2.1%

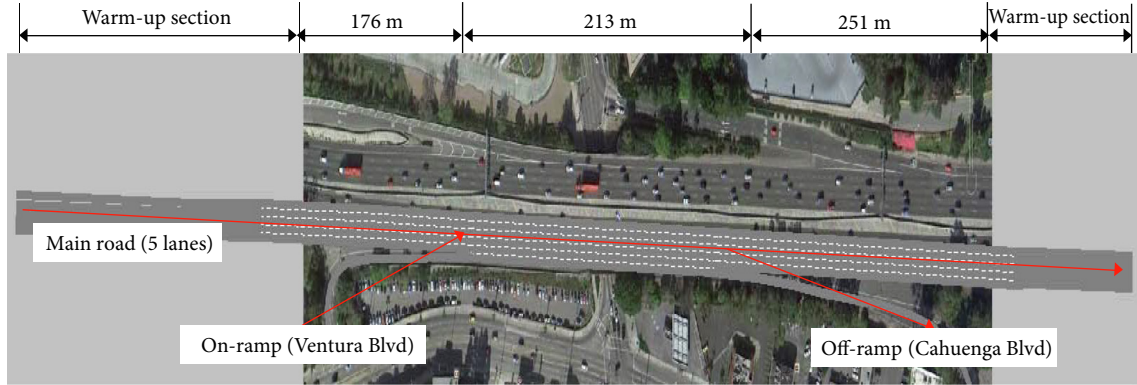


FIGURE 2: Schematic of test site (weaving section in US101).

trajectory set, and then compared with calculated emissions from the NGSIM trajectories, which were defined to be the “ground truth” real-world values.

The driving characteristics are not the only important factor in ensuring a reliable emissions estimation: the vehicle composition is also critical. Even if a model is well-calibrated and realistic traffic flow is simulated, emissions estimates can be inaccurate if the actual composition of vehicles is not reflected in the simulation. In order to determine the effect of vehicle composition on the accuracy of emissions estimation, we evaluated an additional trajectory set excluding heavy vehicles at the highest calibration level (L4-pc). Note that we did not evaluate any such “passenger car only” cases in the lower calibration levels because their effects are not only included in L4-pc, but they are also less calibrated than L4, so evaluating the effect of vehicle composition when estimating emissions at lower levels is not a critical problem.

### 3. Traffic Simulation and Calibration

**3.1. Study Area.** The test network site evaluated in this study was the weaving section of US101 in Los Angeles, CA, as shown in Figure 2. The test section was approximately 640-m long and consisted of five mainline lanes throughout the section, with a 213-m long auxiliary lane for merging between the on-ramp and Ventura Boulevard and the off-ramp at Cahuenga Boulevard. To simulate the traffic flow, calibrate the simulation and calculate emissions, we first obtained the traffic environmental information, aggregated traffic data and individual vehicle trajectories (i.e., speed and acceleration) from the NGSIM data, in which all vehicle movements on the test site were recorded by multiple video cameras and each

vehicle’s trajectory was generated using image processing techniques. The trajectory data provided the precise location of each vehicle within the study area every one-tenth of a second (i.e., each vehicle was tracked in every video frame), resulting in detailed vehicle lane positions and locations relative to other vehicles. A total of 45 minutes of data were available in the dataset for peak-time traffic, segmented into three 15 minute periods: 07:50–08:05, 08:05–08:20, and 08:20–08:35, representing the build-up of congestion, transition between uncongested and congested conditions, and full congestion during the peak period, respectively. In the given time period, a traffic volume of 6,101 vehicles was recorded passing through the test section at an average speed of 10.58 m/s (38.10 kph). The entire vehicle composition is shown in Table 1.

**3.2. Simulation Setup.** Based on environmental information and four basic aggregated traffic data which was explained at Section 2, we built a traffic simulation network using VISSIM which is one of most popular micro-scale traffic simulators for modelling individual vehicle interactions under given traffic conditions. Because of the stochastic nature of VISSIM, we applied a 15 min warm-up and 45 min simulation period, running multiple simulations using random seed numbers. The average result of several simulation runs is reported as the average traffic condition for the given calibration levels.

The scope of this study is considerably microscopic and the total evaluated time frame is very short at 45 min, so the ability to implement and calibrate the traffic network using the standard one-hour interval  $U$  and  $Q$  is limited. Thus, 15-min interval  $U$  and  $Q$  data were extracted from the standard aggregated data and coded accordingly in our traffic network. Furthermore, the network was designed so that the volume could be set for each lane in order to implement the higher calibration levels (i.e., L2–L4) that consider I/O per lane, SMS

TABLE 2: Calibration order, MOEs, performance measurements, and target ranges according to calibration level.

Order	MOE	Performance measurement	Target range	Calibration level					
				L0	L1	L2	L3	L4	L4-pc
1	Total volume	ER	$\pm < 5\%$	<b>-0.31</b> (-19 veh)	<b>-0.03</b> (-2 veh)	<b>0.10</b> (6 veh)	<b>0.20</b> (12 veh)	<b>0.25</b> (15 veh)	<b>0.49</b> (30 veh)
2	15-min flow rate	RMSPE*	<5	<b>1.54</b>	<b>1.45</b>	<b>1.52</b>	<b>0.88</b>	<b>0.85</b>	<b>0.60</b>
3	Avg SMS	ER	$\pm < 5\%$	<b>3.69</b> (1.4 kph)	<b>-2.62</b> (-1.0 kph)	<b>0.52</b> (0.2 kph)	<b>2.83</b> (1.1 kph)	<b>0.85</b> (0.3 kph)	<b>-0.05</b> (0.0 kph)
4	15-min SMS	RMSPE*	<5	<b>2.87</b>	<b>2.34</b>	<b>1.99</b>	<b>1.30</b>	<b>0.75</b>	<b>0.68</b>
5	O-D per link	GEH	<5	max 10.82	<b>4.99</b>	<b>4.00</b>	<b>2.93</b>	<b>3.29</b>	<b>2.75</b>
				avg 4.98	<b>4.02</b>	<b>2.57</b>	<b>1.29</b>	<b>1.30</b>	<b>1.15</b>
6	I/O per lane	RMSPE*	<5	min 0.69	<b>2.51</b>	<b>1.25</b>	<b>0.22</b>	<b>0.25</b>	<b>0.09</b>
				9.07	8.26	<b>3.14</b>	<b>2.35</b>	<b>1.62</b>	<b>1.84</b>
7	SMS per lane	RMSPE*	<5	26.23	22.23	8.35	<b>4.32</b>	<b>3.79</b>	<b>0.44</b>
				max 13.27	11.93	4.00	6.25	<b>3.30</b>	<b>4.24</b>
8	O-D per lane	GEH	<5	avg 6.80	6.20	3.52	1.80	<b>1.32</b>	<b>1.22</b>
				min 0.24	0.29	0.24	0.1.	<b>0.0</b>	<b>0.0</b>

Cells with bold text: calibration achieved.

per lane, and O-D per lane. This is required because this specific test section included complex traffic weaving under congested conditions, and the nature of the simulation setup has a considerable effect on the implementation of such conditions with frequent lane changing. To address this complexity, we established warm-up sections on each end of the test section as shown in Figure 2. Note that while input volumes per lane were based on the NGSIM data for calibration levels L2-L4, we coded the same proportion of input volumes in each lane for calibration levels L0 and L1. Additionally, we evaluated a simulated network composed of only passenger cars for the L4-pc scenario to identify the effect of vehicle composition on the estimated emissions.

**3.3. Simulation Calibration.** The objective of this calibration process is to fine-tune the parameters of the microscopic traffic model so that the discrepancy between the observed and simulated traffic behaviours is minimized. In this respect, the calibration is a critical step in ensuring accurate estimation of vehicle emissions. As described in previous section, we setup of the calibration levels in five as increasing the number of MOEs reflecting order and composition of calibration MOEs assuming that simulated macroscopic traffic flow is more similar to actual as the level increased by sophisticated calibration. After implement the network, same types of the aggregated traffic data in accordance with compositions of calibration MOEs were exported from the simulated network in each level for calibrating the simulation. Specific performance measurement of each MOE, its target range and order of calibration are presented in left side of Table 2.

A number of goodness-of-fit equations (e.g. MAPE, ER, RMSE, RMSPE, GEH, and etc.) have been used to quantify estimation or prediction error and serve as performance measurements in various calibration-related research. Such research has focused heavily on numerical algorithms that minimize error by applying appropriate target values or fitness functions on their own [21-25]. It is, however, widely known that better performance can be expected when the measurement target

range is set smaller, but it is hard to regularize the boundary under such conditions due to general problems such different scales or units (i.e., there is no rule-of-thumb standard). In other words, the goodness-of-fit tests usually employed to assess the effectiveness of calibration do not provide sufficient information for assisting the user in identifying weaknesses encountered during the course of the calibration [21]. To address this shortcoming, we adopted the well-known error rate (ER), root mean squared percentage error (RMSPE), and GEH fitness functions as performance measurements to identify the differences between each MOE and to more easily set its target ranges. The target ranges of these measurements must be satisfied in each calibration level; as the calibration level increases, so does the number of MOEs, maintaining the target ranges from the previous level by definition.

For the default calibration level (i.e., L0), we used four performance measurements: the error rates (ERs) of the total volume and average SMS, and the root mean squared percentage errors (RMSPEs) of the flow rate and average SMS in a 15 min period. Using GEH statistics which is most widely used measurements in O-D data [21], the O-D per link was evaluated to provide better calibration in L1, and four O-D pairs existed in the test network (main-to-main, main-to-off, on-to-main, and on-to-off). After satisfying the calibration requirements of L1, the RMSPE of the I/O per lane was considered in L2. In L3, the RMSPE of the SMS per lane was considered. Finally, we used the GEH of the O-D per lane to provide the highest calibration level (L4), with 36 pairs existing between lanes. If the GEH values resulting from a simulation run at each measurement location for all O-D pairs evaluated were less than 5, the simulation was considered to be acceptably calibrated. The ER target ranges were set to be within  $\pm 5\%$  and the RMSPE\* target ranges were set to below 5 (RMSPE\* indicates that the number has been multiplied by a factor 100 because of the small values resulting from the RMSPE equation). As mentioned previously, smaller ranges are better for sophisticated calibration, but result in a more intricate and time-consuming process. In this study, each of the target ranges were reasonably established by

trial-and-error as it is difficult to theoretically define their upper bounds. However, the established ranges still indicate a fitness performance of approximately 95%.

Because multiple MOEs are configured in each level, determining the order of calibration is a critical problem addressed in this study. At each calibration level, the simulation was iterated, adjusting the calibration parameters until their performance measurements fell within their target ranges. This means that the MOEs contained in each level must simultaneously satisfy all target ranges. However, it is very difficult to simultaneously calibrate several MOEs and to design a single fitness function integrating them that guarantees a global maximum and local minimum in the optimization problem. In this paper, we accordingly adopted the practical and thoroughly discussed systematic general calibration procedure presented in [32]. The order of calibration MOEs evaluated was as shown in the leftmost column in Table 2. As previously mentioned in Section 2, the data in each additional MOE are less aggregated than in previous MOEs. Typically, I/O and SMS per lane can be considered to be the same, hierarchically speaking, but we evaluated the I/O per lane in an earlier level (i.e., L2) in this study because volume-related MOEs have a greater affect than speed-related MOEs [32].

As the calibrations were conducted by order of calibration MOE defined according to level, the re-calibration problem must be considered. At the default level, the total traffic volume and flow rate were calibrated first, followed by the average SMS, then the 15-min SMS. After completing the earlier levels, each additional MOE was calibrated sequentially according to the defined order of levels. During this process, the error of previously calibrated MOEs can change when calibrating an additional MOE. If the resulting error is higher than the target range, the previously calculated MOEs must be re-calibrated until all MOEs satisfy their tolerances. For example, in L3, we adjusted the calibration parameters for the SMS per lane after completing L2. However, the calibration of MOEs in L2 can be broken when calibrating the additional MOE, so the previously calibrated MOEs must be iteratively re-calibrated. Note that because the order of MOEs is somewhat inversely proportional to the aggregation level, this re-calibration is likely to be less difficult at lower levels. This step-by-step calibration procedure can thus be more effective and practical in real-world applications.

In our study, we selected the driving behaviour parameters of Wiedemann74 model in VISSIM which are related to vehicles' car-following and lane-changing, as it is more suitable to modelling urban traffic and weaving sections [18]. These parameters were used as calibration parameters. The speed limit of the network is set to be the same as the actual traffic without calibration. Furthermore, we also adjust the input flow of main lanes to exceed the roadway capacity because the point-measured traffic volume and speed were lower than the capacity and the free flow speed both (i.e., congested traffic conditions). Note that if measured low traffic volume is coded in the simulation, the congested condition never occurs. The ranges (with default values) and calibrated values of the calibration parameters are summarized in Table 3. Most of adjusting boundaries of the parameters were determined as being around 30% of the default values in order to prevent unrealistic

vehicle driving in simulation [23]. In each MOE calibration, the calibration parameters were adjusted automatically by increasing the value within the adjusting boundary, but the target calibrated or re-calibrated values of the MOEs were manually selected in each iteration to check the deviations of each MOE.

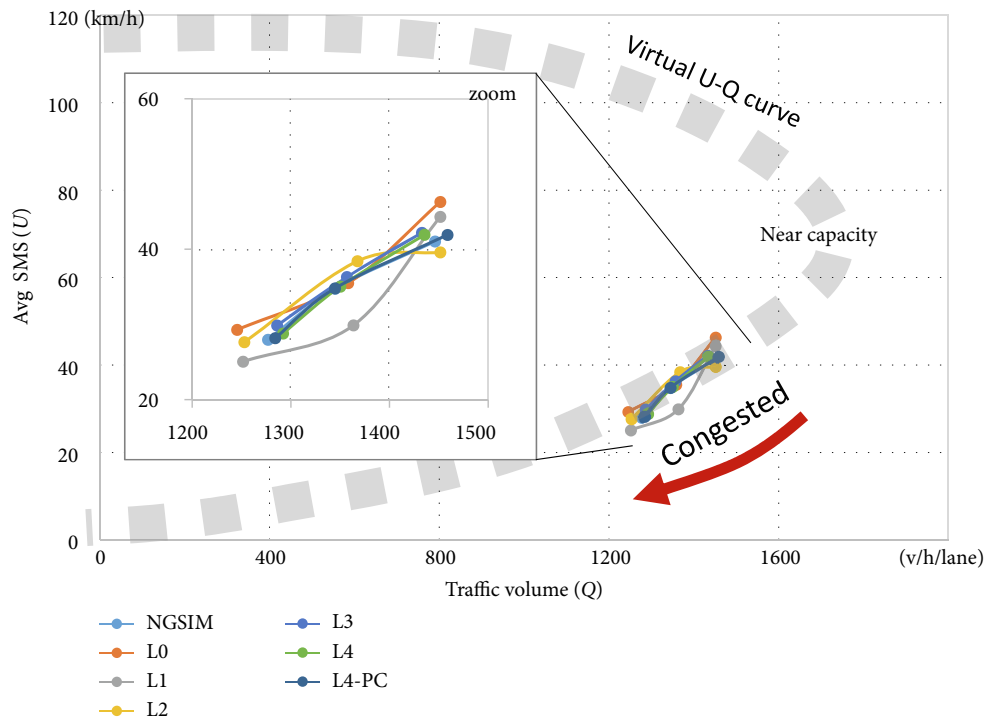
*3.4. Results of Calibration.* The result values in each of the five calibration levels are summarized in right side of Table 2. To compare the results of each level calibration, three performance measurements were considered in this section. First, as shown in the *U-Q* curves in Figure 3, the observed and simulated traffic conditions agree well under the congested condition, and we confirmed that the all calibration levels were successfully completed in L0 level.

To identify the differences between each calibration level and the observed traffic environment in detail, we also compared the average SMS in each lane and generated a heat map of O-D in the three 15 min periods considered in Figures 4 and 5, respectively. In Figure 4, it can be observed that there is a clear difference between the simulated and observed values for L1 and L0. The L0 calibration presents a higher speed than the observed data for Lane 1 through Lane 3, indicating that the traffic flow in these lanes is more fluid for this calibration level than for the others. However, the speed can be seen to dramatically decrease in Lane 4, which exhibits the largest deviation from the observed data, indicating that vehicles mostly change lane at Lane 4. In the case of L1 calibration, vehicles are able to find their own routes, but because fitting the inflow and outflow is a calibration objective at this level, the resulting weaving is significant, especially in Lane 5, as indicated by the fact that it exhibits the slowest speed, showing a considerable deviation from the observed data. According to the results under L2, a minor deviation from the observed data is exhibited in Lanes 4 and 5, but the shape of the speed curve is similar to that of the observed data. In the remaining cases, the shape of curves is increasingly similar to that of the observed data with almost no deviation, indicating that these calibration levels reflect the observed data quite well. The heat map of the O-D flow rates per lane is also compared as shown in Figure 5. As the calibration level increases, the simulated traffic flow becomes more similar to the observed data and exhibits relatively low average GEH values.

Examining the results in Table 2 in detail, it can be observed that the performance measurements of all MOEs of the previous calibration levels satisfied their tolerances (i.e., target ranges). Note that in Table 2, utilizing only the typical MOEs in L0 seems to be insufficient to simulate the actual traffic flow, as indicated by the relatively high values of the performance measurements. However, higher levels of calibration MOEs appear to more accurately reflect actual vehicle driving behaviours. Indeed, in Table 2 the performance measurements that were satisfied in the previous calibration levels mostly improve (decrease) as the calibration level increases because more elaborate MOEs are included. This indicates that calibration levels are correlated and not independent. Accordingly, to validate the proposed calibration procedure, we conducted the calibrations in levels L1-L4 in various orders without completing L0 first (i.e., ignored our proposed

TABLE 3: Calibration parameters and calibrated values.

Calibration parameters	Lower bound	Default value	Upper bound	Calibrated value in each level					
				L0	L1	L2	L3	L4	L4-pc
Coded traffic flow of mainlines (v/h)	9000	250	13500	11900	12850	11050	11450	10900	12700
Look ahead distance (m)	150	250	350	210	260	280	290	310	290
Average standstill distance (m)	1	2	3	1.2	1.4	1.5	1.1	1.2	1.5
Additive part of desired safety distance	1	2	3	1.3	1.7	1.6	1.3	1.7	1.8
Multiple part of desired safety distance	2	3	4	2.6	2.1	2.7	2.2	2.2	2.3
Max Decel. of subject vehicle ( $m/s^2$ )	-7	-4	-1	-6.8	-5.9	-6.6	-6.7	-5.6	-6.4
Max Decel. of trailing vehicle ( $m/s^2$ )	-7	-3	-1	-6.7	-6.7	-6.5	-6.1	-6.1	-6.3
Accepted Decel. of subject vehicle ( $m/s^2$ )	-5	-1	-0.5	-3.1	-4.3	-2.9	-3.3	-3.7	-3.6
Accepted Decel. of trailing vehicle ( $m/s^2$ )	-5	-1	-0.5	-3.4	-3.7	-3.5	-2.3	-2.7	-2.9

FIGURE 3: Comparison of  $U$ - $Q$  curve for each calibration level.

calibration procedure). As a result, it was very difficult to minimize the errors of the default MOEs due to the global maxima, and reliable traffic flows were not simulated even when the less aggregated MOEs were well calibrated. In Section 4.3, we validated the proposed calibration procedure in terms of emissions estimates in additional cases that were not able to satisfy the L0 calibration level. To summarize, it is thus expected that traffic simulation-based micro-scale emissions estimation will be affected by the sophistication of the calibration level, as discussed in the following section.

## 4. Emissions Estimation

**4.1. Vehicle Activity Characterization.** As mentioned in Section 2, each vehicle trajectory set collected from NGSIM and VISSIM for each calibration level was input into MOVES to convert the trajectories into vehicle activity characterizations.

In the MOVES software, the first input step is to create a project-level database in which the imported data are stored. Input files include meteorology data, traffic composition, the percentage of trucks, network length, traffic volume, average speeds, grade, vehicle age distribution, operating mode distribution for running emissions, link drive schedules, and fuel information (i.e., gasoline or diesel). Because the fuel type and model year of all vehicles are difficult to clearly determine from the NGSIM data, we assumed that passenger cars and motorcycle are gasoline engines, and buses and trucks are diesel engines in this study.

The OpMode approach in MOVES is a modal binning type estimation that allows for the definition of the amount of travel time spent in various operating modes, including braking, idling, coasting, and cruising/accelerating within various speed ranges for various ranges of vehicle specific power (VSP) [11], based on the collected second-by-second trajectory data. In other words, the OpMode is a measure of the state of each

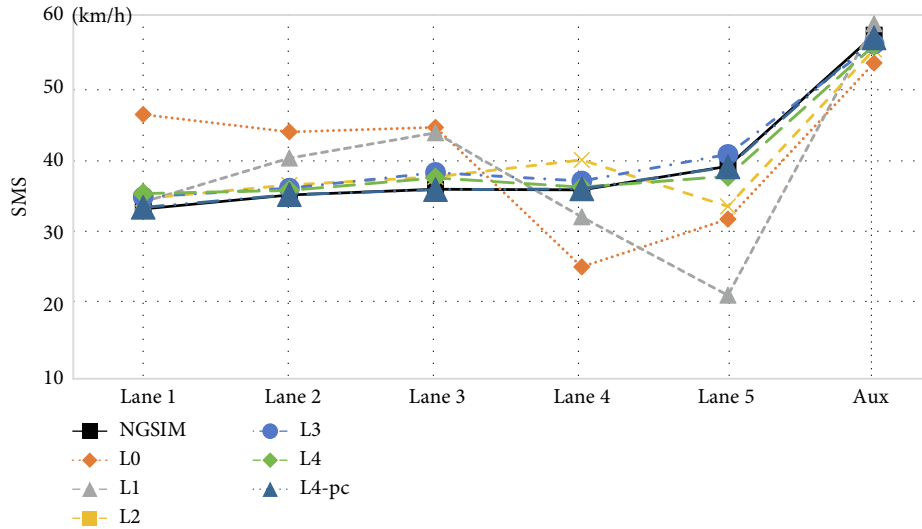


FIGURE 4: Comparison of SMS per lane for each calibration level.

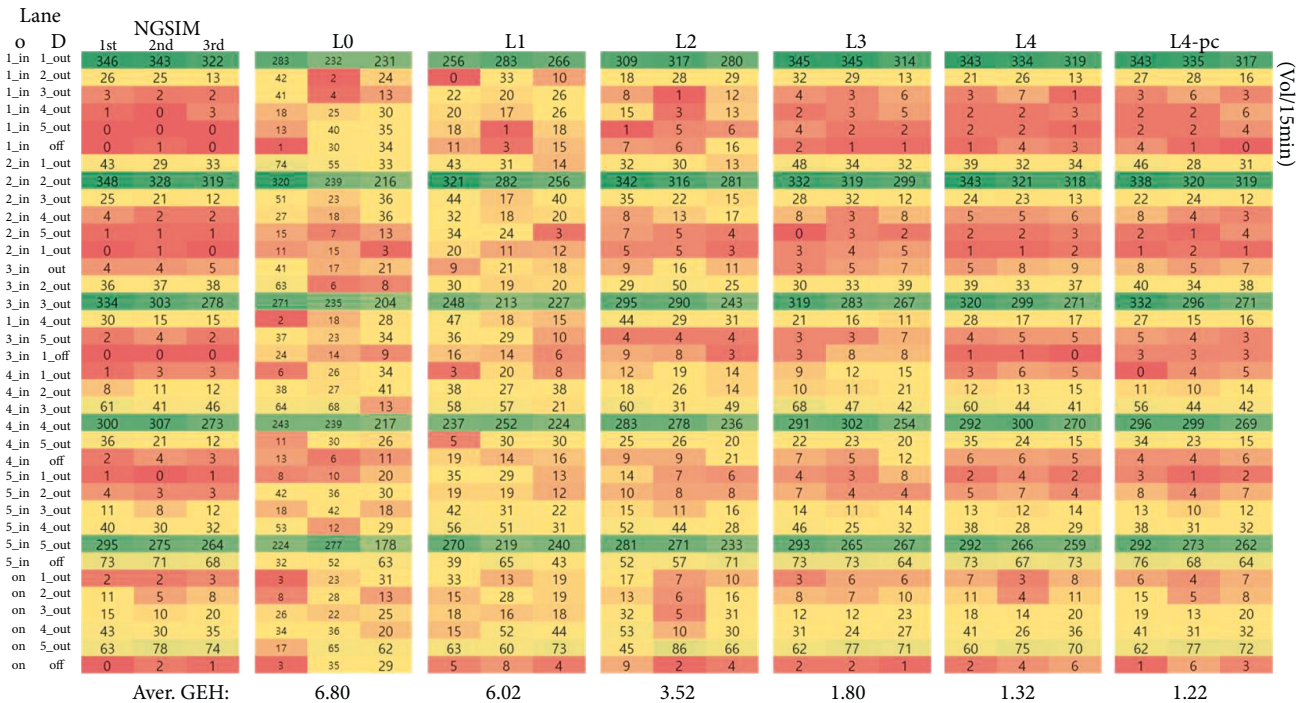


FIGURE 5: Comparison of GEH heat map of O-D per lane in 15 min for each calibration level.

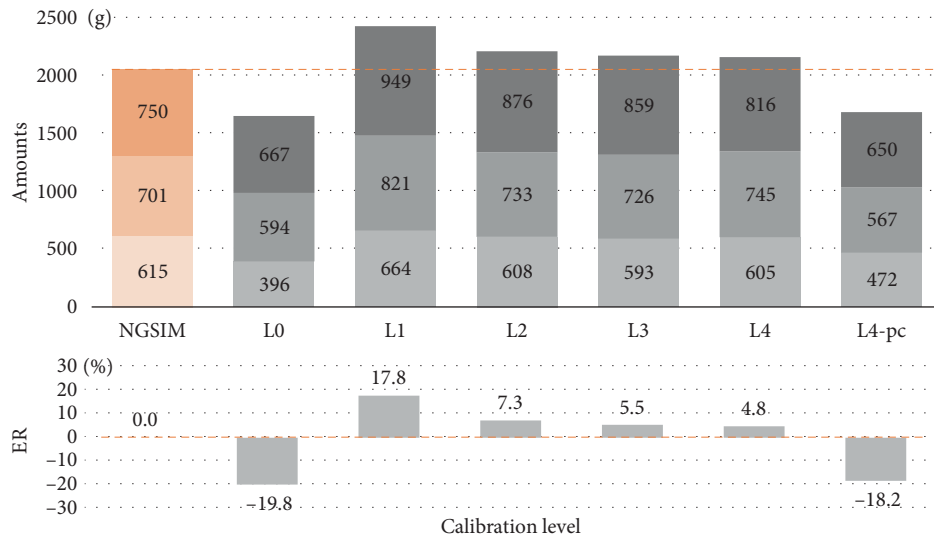
vehicle’s engine at any particular moment. This function produces operating mode fractions for each bin that are then used as one of several inputs to compute the base emission rates. In order to estimate emissions on a micro-scale, we focused solely on the OpMode approach for the project-level data in this study.

**4.2. Comparison Results of Emissions.** We estimated the quantities of CO, CO<sub>2</sub>, and NO<sub>x</sub> emissions using the different calibration levels evaluated in this study. We assumed that the emissions calculated using the NGSIM trajectory data were the “ground truth” real-world values that we compared

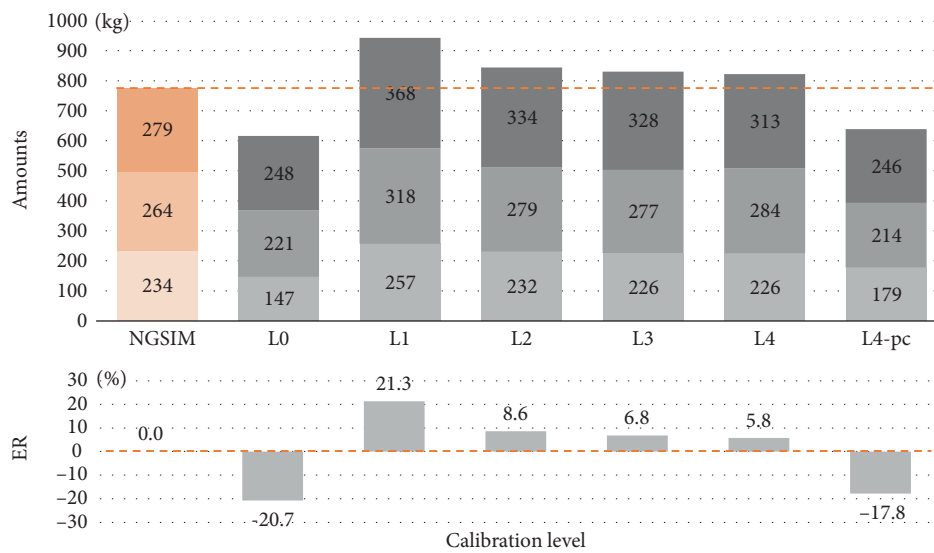
to the vehicle emissions estimated according to trajectory sets generated under the respective calibration levels. Using this comparison, we identified the degree to which the simulation calibration affected the accuracy of the estimated emissions to determine reliable calibration level. Additionally, we examined the effect of the vehicle type composition (mixed vs. small vehicles only) on the emissions estimate when the simulated traffic flow was well-calibrated with the observed real-world situation.

Figure 6 and Table 4 show comparisons of the emissions estimated under each calibration level with the actual amounts from VISSIM and NGSIM in the three 15 min periods

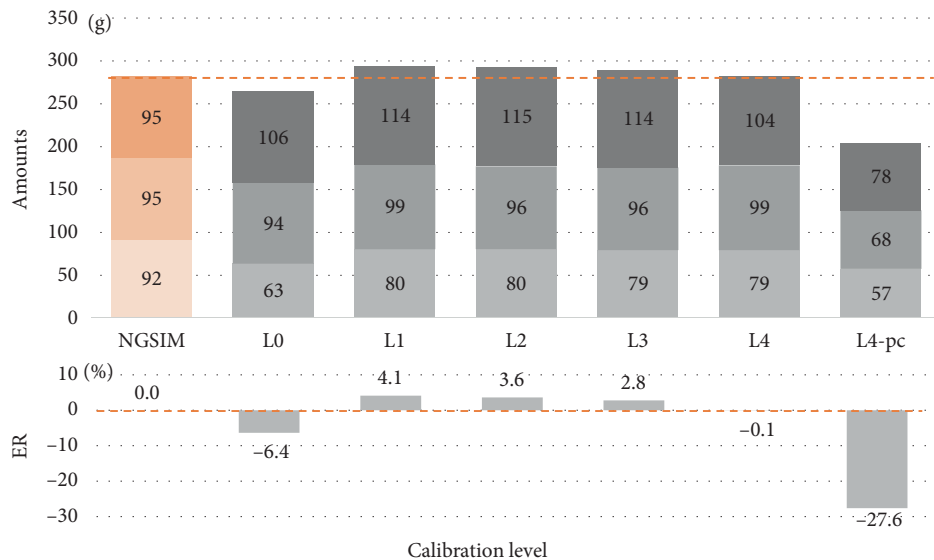




(a)



(b)



(c)

FIGURE 6: Comparisons of emissions from each calibration level and ground truth emissions in the three 15-min periods considered and totals for: (a) CO, (b) CO<sub>2</sub>, and (c) NO<sub>x</sub>.

TABLE 4: Total vehicle emissions and error rates (%) between ground truth and estimated emissions according to calibration levels.

Emissions	Ground truth	L0	L1	L2	L3	L4	L4-pc
CO (kg)	2.06	1.66 (-19.8%)	2.43 (17.8%)	2.22 (7.3%)	2.18 (5.5%)	2.17 (4.8%)	1.69 (-18.2%)
CO <sub>2</sub> (kg)	778	616 (-20.7%)	943 (21.3%)	845 (8.6%)	831 (6.8%)	823 (5.8%)	639 (-17.8%)
NO <sub>x</sub> (g)	281	264 (-6.4%)	293 (4.1%)	292 (3.6%)	289 (2.8%)	281 (-0.1%)	204 (-27.6%)

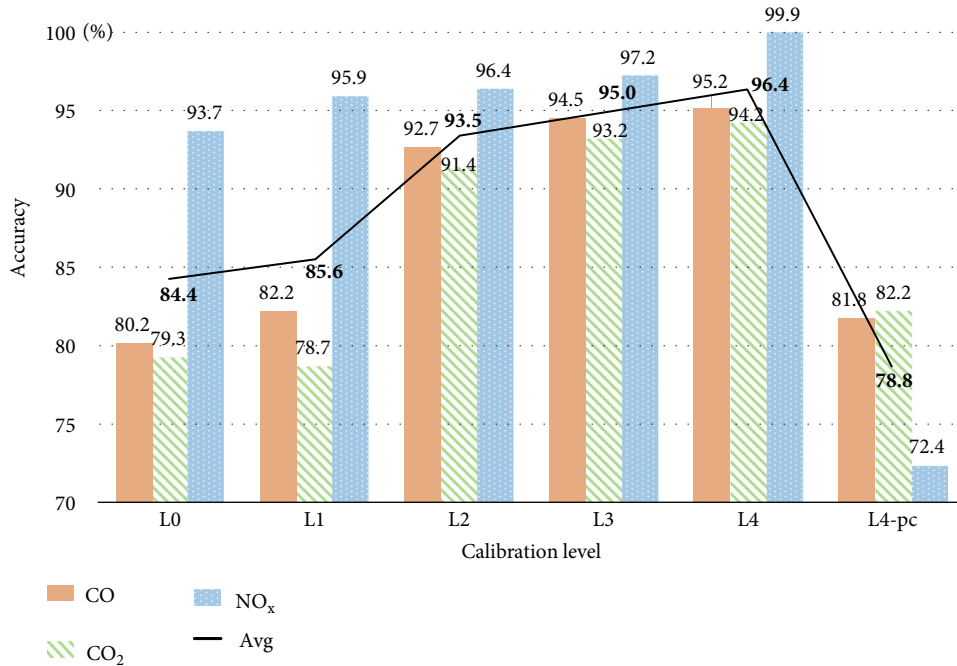


FIGURE 7: Accuracy of estimated emissions for each calibration level.

considered and their summations. Comparing the estimated emissions according to calibration level, it can be found that the ERs of CO and CO<sub>2</sub> emissions are similar. The simulated emissions in L0 are mostly underestimated (i.e., -19.8% for CO, -20.7% for CO<sub>2</sub>, and -6.4% for NO<sub>x</sub>). This is likely because due to the high average speed in Lane 1 through Lane 3, vehicle lane changing for appropriate route searching is less reflected in the simulated case than the observed real-world situation. On the other hand, the simulated values in L1 are mostly overestimated (i.e., 17.8% for CO, 21.3% for CO<sub>2</sub>, and 4.1% for NO<sub>x</sub>), likely because the vehicle routes in the observed real-world traffic situation were considered using the O-D per link MOE in the calibration, so the lane changing in Lane 5 and the auxiliary lane are over-implemented in the simulation. In other words, it can be deduced that vehicle acceleration and deceleration for securing gap acceptance during lane changing behaviour are sensitive to the level of calibration, and in turn effect the estimated CO and CO<sub>2</sub> emissions. In L2, L3, and L4 calibrations, vehicle emissions are slightly overestimated, by -0.1% to 8.6%, but the values are close to the ground truth quantities for each emission. For NO<sub>x</sub> emissions, the deviations are the smallest under all calibration levels that

considered heavy vehicles, but it can be identified in Table 4 that the deviation rate in L4-pc is -27.6%, indicating that the calibrated value is considerably different from the actual amounts. This suggests that although a traffic simulation model may be well calibrated, the estimated vehicle emissions can be considerably underestimated if the presence of heavy vehicles is not accurately reflected in the model.

To sum up, Figure 7 shows the accuracy of the estimated emissions for each calibration level. Overall, a calibration level of L2 or higher can be observed to provide an estimated emissions average accuracy above 90% when considering the presence of heavy vehicles. A comparison of the performance of L0, L1, and L2 as shown in Table 2 indicates that simply fitting the I/O per lane to the observed data leads to a reduction in the RMPSE of SMS per link MOE. As a result, a calibration level of L2 or higher is likely sufficient for a micro-scale traffic simulation-based emissions estimation. Additionally, we determined that accurately accounting for the composition of heavy vehicles has a considerable effect on the accuracy of the estimated emissions. Considering the difficulty of obtaining detailed traffic data for calibration in the real world such as O-D per lane, the use of the L2 calibration level is therefore

TABLE 5: Effect of calibration sequence on estimated emissions accuracy for calibration level L2.

Order	MOE	Performance measurement	Target range	Calibration level		
				L2_v	L2_s	L2
1	Total volume	ER	$\pm < 5\%$	-7.10	<b>0.43</b>	<b>0.10</b>
2	15-min flow rate	RMSPE*	$< 5$	8.12	<b>1.79</b>	<b>1.52</b>
3	Avg SMS	ER	$\pm < 5\%$	<b>1.01</b>	6.79	<b>0.52</b>
4	15-min SMS	RMSPE*	$< 5$	<b>3.17</b>	8.32	<b>1.99</b>
5	O-D per link	GEH	$< 5$	<b>3.97</b>	<b>2.13</b>	<b>2.57</b>
6	I/O per lane	RMSPE*	$< 5$	<b>4.42</b>	<b>3.64</b>	<b>3.14</b>
Avg accuracy of estimated emissions (%)				61.3	82.9	93.5

Cells with bold text: calibration achieved.

most reasonable for micro-scale traffic simulation based emissions estimation (i.e., I/O per lane is a significant factor for calibrating vehicle emissions estimation simulations).

**4.3. Validation.** In Section 4.2, we determined that the calibration MOEs in L2 were reasonable for estimating emissions. In order to validate our proposed calibration procedure, we attempted to complete the L2 calibration regardless of the order of the calibration MOEs. However, we could not complete the calibration in some typical, lower-level MOEs even if additional, less-aggregated MOEs were well calibrated. Two of these cases that were not able to satisfy level L0 in the L2 calibration were accordingly evaluated in terms of the accuracy of their emissions estimates. Case L2\_v includes volume-related MOEs (i.e., total volume and 15-min flow rate) that did not satisfy the calibration criteria in L0. Case L2\_s includes speed-related MOEs (i.e., average SMS and 15-min SMS) that did not satisfy the calibration criteria in L0. In these cases, more re-calculation was required to approach the target ranges. We stopped the calibration iterations when the MOEs were slightly beyond the target ranges (i.e., -7.10% and 8.12% for total volume ER and 15-min flow rate RMSPE\*, respectively, in L2\_v, and 6.79% and 8.32% for average SMS ER and 15-min SMS RMSPE\*, respectively, in L2\_s). We then extracted vehicle trajectories from each of these simulations and estimated the emissions in the same manner as before.

The calibration results of each MOE and the accuracy of estimated emissions are summarized in Table 5. Clearly, L2\_v shows the least accurate estimated emissions at 61.3%, whereas L2\_s shows an accuracy of 82.9%, just lower than that of the L0 calibration (84.4%), even though the MOEs included in L2 were well calibrated. Therefore, it can be confirmed that the calibration of traffic volume-related MOEs is more critical in ensuring accurate emissions estimates. When comparing the performance measurements, the traffic volume and speed are indeed correlated as discussed in Section 3.4, so if the error in one type of MOE is large, the other type of MOE may also include a slight error. Typically, since the added MOEs, O/D per link and I/O per lane, are related to traffic volume, it can be seen that a larger error occurs than the original L2 at the additional MOEs in L2\_v (i.e., from 2.57 to 3.97 at O/D per link, and from 3.14 to 4.42 at I/O per lane). Therefore, our proposed calibration guideline (i.e., especially considering the additional MOE while maintaining the MOE calibration of L0)

was found to be very helpful for more effective calibration in reliable traffic flow implementation and emission estimation.

## 5. Conclusions

In this study, we identified reliable calibration level for microscopic traffic simulations that minimized the discrepancy between observed real-world and estimated vehicle emissions. At first, we introduced the step-by-step calibration procedure including the MOE which is aggregated data such as five calibration levels. At this stage, we demonstrated that widely used calibration level (L0) is not enough to simulate the actual vehicle movements in detail, and our calibration procedure is valid to implement vehicles driving closer to actual. After each level calibration, we generate trajectory data, then determined the quantities of CO, CO<sub>2</sub>, and NO<sub>x</sub> emissions as compared with the calculated ground truth emissions from NGSIM trajectory. Consequently, we demonstrated that the accuracy of emissions is reliant on how well the acceleration and deceleration behaviours are reflected in the simulation trajectories through the application of an appropriate level of calibration.

It was identified that higher calibration levels also resulted in improved estimation accuracy, but the detailed traffic data such as the O-D per lane required for such higher-level calibration is very difficult to obtain in the real world. Additionally, it is more time consuming to apply more sophisticated calibration levels to the simulation. In this respect, determination of a reliable calibration MOEs by comparing the estimation results of emissions between each calibration level is a critical problem. As a result, we figured out that calibrating not only traffic flow rate and speed but also the O-D per link and I/O per lane (i.e., calibration MOEs in L2) in microscopic traffic simulations are critical MOEs for reliable vehicle emissions estimation since the MOEs can be readily employed in the field using recently improved traffic data detection technology. Through validation, we also determined that the calibration of traffic volume-related MOEs should be given priority. In order to identify the effect of vehicle composition on the emissions estimation, we evaluated an additional case excluding heavy vehicles at the highest calibration level (i.e., L4-pc). This evaluation confirmed that the composition of heavy vehicles has a considerable effect on the accuracy of the resulting emissions estimate even when traffic flow is realistically simulated through effective traffic model calibration. Accordingly, when

conducting traffic simulation-based emission estimations, not only is the calibration of the traffic model important, but also the accuracy of vehicle type composition.

As traffic data detection technology continues to develop, it will become increasingly possible to obtain more of the data, such as specific vehicle information, required for the effective use of increased calibration levels. In this regard, we expect that more reliable traffic simulation-based emissions estimations are currently possible by using the findings of this study, which can be used as a guideline for emissions-related organizations and researchers. Furthermore, because our proposed calibration procedure is described in a practical way, it has the advantage of being generally applicable and easily transplanted into other studies.

In future research, we will develop an optimization algorithm by applying a multi-integrated fitness model to our procedure, replacing the time-consuming iterative trial-and-error process required to determine calibration performance measurement ranges. Additionally, we plan to validate the proposed calibration procedure through the application of other site traffic data and other traffic models. We will accordingly strengthen the findings of this research by methodically establishing clearer target ranges through sensitivity analysis and further validation. Also, the effects of various roadway levels of service reflecting more elaborate and varied vehicle type compositions, and their interactions with various levels of calibration, will be evaluated with respect to the resulting emissions estimates. Additionally, further research into the influences of vehicle composition on PM<sub>2.5</sub> emissions should be conducted to further elucidate the effects of collected vehicle data on estimated emissions.

## Data Availability

The [US101 vehicle trajectory] data used to support the findings of this study have been deposited in the [Next Generation Simulation (NGSIM) Vehicle Trajectories and Supporting Data] repository ([<https://data.transportation.gov/Automobiles/Next-Generation-Simulation-NGSIM-Vehicle-Trajectory/8ect-6jqj>]).

## Conflicts of Interest

There is no conflict of interest regarding the publication of this paper.

## Acknowledgments

This work was supported by the National Research Foundation of Korea grant funded by the Republic of Korea government (MSIT) (No. 2018R1C1B6006330) and also supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant 19TLRP-B148659-02), Republic of Korea.

## References

- [1] G. Lee, *Integrated Modeling of Air Quality and Health Impacts of a Freight Transportation Corridor*, Ph.D. Dissertation in the University of California, Irvine, 2011.
- [2] M. Barth and K. Bobiboonsomsin, "Energy and emissions impacts of a freeway-based dynamic eco-driving system," *Transportation Research Part D: Transport and Environment*, vol. 14, no. 6, pp. 400–410, 2009.
- [3] K. Ahn, H. Rakha, A. Trani, and M. Van Aerde, "Estimating vehicle fuel consumption and emissions based on instantaneous speed and acceleration levels," *Journal of Transportation Engineering*, vol. 128, no. 2, pp. 182–190, 2002.
- [4] C. Kun and Y. Lei, "Microscopic traffic-emission simulation and case study for evaluation of traffic control strategies," *Journal of Transportation Systems Engineering and Information Technology*, vol. 7, no. 1, pp. 93–99, 2007.
- [5] G. Lee, S. You, S. G. Ritchie, J.-D. Saphores, M. Sangkapichai, and R. Jayakrishnan, "Environmental impacts of a major freight corridor: a study of the I-710 in California," *Transportation Research Record*, vol. 2123, no. 1, pp. 119–128, 2009.
- [6] M. Claggett and J. Houk, "Comparing MOBILE 6.2 and Emfac 2007 emissions factors for mobile source air toxics," in *Proceedings of the Transportation Research Board 87th Annual Meeting*, Transportation Research Board, Washington, DC, USA, 2008.
- [7] E. K. Nam, C. A. Gierczak, and J. W. Butler, "A comparison of real-world and modeled emissions under conditions of variable driver aggressiveness," in *Proceedings of the Transportation Research Board 82nd Annual Meeting*, Transportation Research Board, Washington, DC, USA, 2003.
- [8] H. Rakha, K. Ahn, and A. Trani, "Comparison of MOBILE5a, MOBILE6, VT-MICRO, and CMEM models for estimating hot-stabilized light-duty gasoline vehicle emissions," *Canadian Journal of Civil Engineering*, vol. 30, no. 6, pp. 1010–1021, 2003.
- [9] S. Bai, D. Eisinger, and D. Niemeier, "MOVES versus EMFAC: comparison of greenhouse gas emissions using Los Angeles County," in *Proceedings of the Transportation Research Board 88th Annual Meeting*, Washington DC, USA, 2009.
- [10] iCET, "Calculating urban, transportation emissions: private vehicles," *Lessons from Existing Tools' Development, Design and Implementation*, Innovation Center for Energy and Transportation, 2015.
- [11] U. S. EPA, "Motor vehicle emission simulator (MOVES): user guide for MOVES2010," <http://www.ngsim.fhwa.dot.gov>.
- [12] H. Abou-Senna and E. Radwan, "VISSIM/MOVES integration to investigate the effect of major key parameters on CO<sub>2</sub> emissions," *Transportation Research Part D: Transport and Environment*, vol. 21, pp. 39–46, 2013.
- [13] H. Abou-Senna, E. Radwan, K. Westerlund, and C. D. Cooper, "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*, vol. 63, no. 7, pp. 819–831, 2013.
- [14] Z. Yao, H. Wei, H. Perugu, H. Liu, and Z. Li, "Sensitivity analysis of project level MOVES running emission rates for light and heavy duty vehicles," *Journal of Traffic and Transportation Engineering (English Edition)*, vol. 1, no. 2, pp. 81–96, 2014.
- [15] H. Rakha and K. Ahn, "Integration modeling framework for estimating mobile source emissions," *Journal of Transportation Engineering*, vol. 130, no. 2, pp. 183–193, 2004.

- [16] J. Y. Park, R. B. Noland, and J. W. Polak, "Microscopic model of air pollutant concentrations: comparison of simulated results with measured and macroscopic estimates," *Transportation Research Record*, vol. 1750, no. 1, pp. 64–73, 2001.
- [17] X. Xu, H. Liu, J. M. Anderson et al., "Estimating project-level vehicle emissions with Vissim and MOVES-matrix," *Transportation Research Record*, vol. 2570, no. 1, pp. 107–117, 2016.
- [18] PTV, "VISSIM 10.00 User Manual," Karlsruhe, 2018.
- [19] PTV, 2019, <http://vision-traffic.ptvgroup.com/en-us/products/ptv-vissim/use-cases/emissions-modelling/>.
- [20] L. Jie, H. Van Zuynen, Y. Chen, F. Viti, and I. Wilmink, "Calibration of a microscopic simulation model for emission calculation," *Transportation Research Part C: Emerging Technologies*, vol. 31, pp. 172–184, 2013.
- [21] L. Chu, H. X. Liu, J. Oh, and W. Recker, *A Calibration Procedure for Microscopic Traffic Simulation*, in *Proceedings of the Transportation Research Board 83rd Annual Meeting*, IEEE, Washington, DC, USA, 2004.
- [22] M. V. Chitturi, J. W. Shaw, J. R. Campbell IV, and D. A. Noyce, "Validation of origin-destination data from bluetooth reidentification and aerial observation," *Transportation Research Record*, vol. 2430, no. 1, pp. 116–123, 2014.
- [23] M. Yu and W. D. Fan, "Calibration of microscopic traffic simulation models using metaheuristic algorithms," *International Journal of Transportation Science and Technology*, vol. 6, no. 1, pp. 63–77, 2017.
- [24] S. Srikanth, A. Mehar, and A. Parihar, "Calibration of vissim model for multilane highways using speed flow curves," *Stavební obzor—Civil Engineering Journal*, vol. 26, no. 3, pp. 303–314, 2017.
- [25] B. Park and H. Qi, "Development and evaluation of a procedure for the calibration of simulation models," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1934, no. 1, pp. 208–217, 2005.
- [26] K. Boriboonsomsin, M. Barth, and K. Xu, "Improvements to on-road emission modeling of freeways with high-occupancy vehicle facilities," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2123, no. 1, pp. 109–118, 2009.
- [27] H. C. Frey, "Development and demonstration of advanced methods for quantifying freight truck activity, energy use, and emissions," Tech. Rep., National Transportation Center at Maryland, University of Maryland, 2017, Final Report, Project ID NTC2015-MU-R-04.
- [28] V. Kovvali, "Video-based vehicle trajectory data collection," in *Proceedings of the Transportation Research Board 86th Annual Meeting*, Transportation Research Board, Washington, DC, USA, 2007.
- [29] V. Punzo, M. T. Borzacchiello, and B. Ciuffo, "On the assessment of vehicle trajectory data accuracy and application to the next generation SIMulation (NGSIM) program data," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 6, pp. 1243–1262, 2011.
- [30] H. Wei, C. Feng, and E. Meyer, "Video-capture-based approach to extract multiple vehicular trajectory data for traffic modeling," *Journal of Transportation Engineering*, vol. 131, no. 7, pp. 496–505, 2005.
- [31] NGSIM us101 dataset, 2018, <https://www.opendatane트워크.com/dataset/data.transportation.gov/8ect-6jqj>.
- [32] J. Hourdakis, P. G. Michalopoulos, and J. Kottommannil, "Practical procedure for calibrating microscopic traffic simulation models," *Transportation Research Record*, vol. 4167, no. 1, pp. 130–139, 2003.
- [33] Y. Wei, Y. Yu, L. Xu et al., "Vehicle emission computation through microscopic traffic simulation calibrated using genetic algorithm," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 9, no. 1, pp. 67–80, 2019.
- [34] J. So, N. Motamedidehkordi, Y. Wu, F. Busch, and K. Choi, "Estimating emissions based on the integration of microscopic traffic simulation and vehicle dynamics model," *International Journal of Sustainable Transportation*, vol. 12, no. 4, pp. 286–298, 2018.
- [35] Cambridge Systematics, "NGSIM US-101 Data Analysis, Summary Reports," Tech. Rep., Federal Highway Administration (FHWA), 2005, <https://camsys.com/>.