sustainability

MDPI

Article

Heterogeneity Aware Emission Macroscopic Fundamental Diagram (e-MFD)

Mohammad Halakoo [1], Hao Yang [1,*] and Harith Abdulsattar [2]

[1] Department of Civil Engineering, McMaster University, Hamilton, ON L8S 4L8, Canada
[2] McMaster Institute for Transportation and Logistics, Hamilton, ON L8S 4L8, Canada
* Correspondence: haoyang@mcmaster.ca

Abstract: Transportation sector is one of the major producers of greenhouse gases which are responsible for climate change. Finding an appropriate emission estimation tool for large-scale networks is essential for developing efficient emission mitigation strategies. This paper presents an advanced version of the emission macroscopic fundamental diagram (e-MFD) which improves the stability and accuracy of the previous model. A bi-modal function is applied to separate free-flow and congested branches of the e-MFD. The accuracy of the proposed e-MFD is evaluated with both a synthetic grid network and a real-world city-level network. The study also assesses the model's stability under directional traffic demands and road incidents. A comparison with the original e-MFD also verifies the superiority of the proposed model with higher accuracy. Standard deviation of density used in the proposed model to boost the performance. It is worth mentioning the standard deviation can be recorded with the existing hardware, such as loop detectors, and does not impose a considerable computational complexity. The proposed model can be employed for emission measurement in large-scale networks and hierarchical traffic control systems for more homogeneous congestion distribution and emission control.

Keywords: emission estimation; emission macroscopic fundamental diagram (e-MFD); model calibration

Citation: Halakoo, M.; Yang, H.; Abdulsattar, H. Heterogeneity Aware Emission Macroscopic Fundamental Diagram (e-MFD). *Sustainability* **2023**, *15*, 1653. https://doi.org/10.3390/su15021653

Academic Editor: Silvia Fiore

Received: 4 November 2022
Revised: 9 January 2023
Accepted: 12 January 2023
Published: 14 January 2023

1. Introduction

Transportation is one of the major producers of greenhouse gases [1], which are responsible for climate change. Therefore, reducing emissions in the transportation sector seems necessary. Emission reduction can be made by changing traffic composition to less pollutant vehicles such as electric vehicles. However, electric vehicles' wide adoption is hindered by issues such as their charging times, driving range, and charging infrastructures [2–4]. With that in mind, emission reduction through traffic management seems more applicable. Emission control strategies such as re-routing can reduce the emission made by transportation [5]. The first step for developing such strategies is estimating the current emission level to take subsequent actions to reduce it. There are two main approaches to emission estimation: macroscopic [6] and microscopic [7,8]. The former refers to emission estimation at a larger scale, e.g., at the network level, and the latter refers to emissions estimation of each vehicle.

While microscopic models are highly accurate and considered the ground truth for emission estimation in the literature [7,8], they are not suitable for large-scale applications. Microscopic models need second-by-second vehicle trajectory data for emission estimation [9], which is financially expensive to acquire. The fine resolution of microscopic emission models challenges their scalability, as many trajectories are needed. Fitting the microscopic emission model to many vehicle trajectories in the network requires high computational power. Then the estimated emission levels need to be aggregated to get the network emission.

On the other hand, macroscopic models [6,10] are more computationally efficient for large-scale networks but usually offer lower accuracy compared to microscopic methods [11]. Macroscopic model inputs such as network speed and density are less computationally and financially expensive to collect. Loop detectors, already available in many cities, can provide the inputs for macroscopic models. Mesoscopic models stand between microscopic and macroscopic models. For example, VT-Meso [12] constructs synthetic driving cycles for each link instead of tracking individual vehicles and aggregates the emission levels. The driving cycle construction is based on the assumption that all drivers accelerate and decelerate at a constant rate. However, finding a platoon of vehicles with a similar acceleration and deceleration rate to apply the model to them might not be accessible in practice [13].

Previous studies found that emission levels correlate with macroscopic traffic network state characteristics, such as density, speed, and flow [9,13]. Recently, [14] combined MFD, COPERT III model [15], and VT-CPEM model [16] to develop CE-MFD (Carbon Emission MFD). In that study, the electric and regular vehicle flows were calculated by multiplying the network flow from the MFD by each vehicle type's penetration rate. Then, each vehicle type's flow was fed to the related macroscopic emission model (Copert for regular vehicles and VT-CPEM for electric vehicles). Eventually, CE-MFD was developed based on the total emission of vehicles (summation of both electric and regular vehicle emissions). While this approach simplifies the emission calculation, the drawbacks of macroscopic models such as [15] were discussed and presented here are in place. Also, there was no comparison with the microscopic emission models to quantify the accuracy of CEI-MFD.

The emission MFD (e-MFD) concept was recently introduced by Saedi et al. (2020) [13]. The e-MFD shows the emission level as a function of the network traffic state (i.e., density, speed, or flow). These inputs can be collected from the existing infrastructure, like loop detectors, making them more accessible to many cities. The advantage of e-MFD is the simplicity and lower computation costs while offering close to microscopic emission estimation results. A recent real case study validated the existence of the e-MFD relationship [9] using vehicle trajectories recorded by drones. Table 1 shows the summary of the previous literature.

Table 1. Literature Review.

Study	Scope	Notes
Panis et al. [7]	Microscopic	High computational cost for macroscopic applications
Rakha et al. [8]	Microscopic	High computational cost for macroscopic applications, Data requirement (GPS records)
Sun et al. [17]	Microscopic	Data requirement for large scale applications (GPS trajectories)
Ntziachristos et al. (COPERT IV) [6]	Macroscopic	Lower accuracy compared to microscopic models
Barmpounakis et al. [9]	Macroscopic	High accuracy, Data requirements (drone recorded vehicle trajectories)
Jamshidnejad et al. [18]	Mesoscopic	Limited to a small simple theoretical network (5 intersections)
Csikos et al. [11]	Macroscopic	COPERT model and MFD were used, lower accuracy compared to microscopic models
Zegeye et al. [19]	Macroscopic	Limited to isolated networks like freeways
Rakha et al. (VT-Meso) [12]	Mesoscopic	Data requirement (finding platoons of vehicles with similar acceleration and deceleration)
Ji et al. [14]	Macroscopic	Similar estimation accuracy to COPERT
Gori et al. [20]	Mesoscopic	limited application (vicinity of signalized intersections)
Saedi et al [13]	Macroscopic	First e-MFD, no stability analysis was shown

Congestion distribution in the network was found to be an essential factor affecting the overall traffic network performance [21,22]. Knoop et al. (2015) used the standard deviation of density for measuring congestion distribution. They showed that the increase in the standard deviation of density, maintaining the density level, negatively impacts the traffic flow in the traffic network [21]. Recently, Ref. [13] showed that emission rates and flow are highly correlated. Based on what was discussed above, congestion distribution could improve the accuracy and stability of e-MFD in different scenarios. To the authors' knowledge, no study has incorporated congestion distribution. Also, the literature did not study the stability of e-MFD under different circumstances, such as the change in demand and accidents. In this study, we aim to answer the following questions:

- Does network congestion distribution have any relationship with emission rate?
- Will the inclusion of congestion distribution term (as the standard deviation of density) improve the accuracy of the e-MFD model?
- Will the congestion distribution term enable e-MFD to handle exceptional circumstances, such as evacuation and accidents in the network, with high estimation accuracy?

This study comprehensively analyzes e-MFD stability and proposes a new e-MFD model. In addition to what has been proposed in the previous study [13], the proposed model considers the traffic congestion homogeneity using the standard deviation of density [21,22]. It is worth mentioning that this is the first study to consider congestion distribution as an independent variable for emission estimation modelling. The developed model and the previous one [13] were tested on a grid network in three experimental scenarios: day-to-day stability, directional demand to mimic evacuation situations, and blockage on high-density links to mimic accidents. The developed model outperformed the previous e-MFD under the same test conditions. A quantitative comparison between the models was performed. In addition, the models were calibrated for Blacksburg, VA, USA network, and the validation results were compared. Results showed that the new heterogeneity-aware e-MFD outperformed the previous model.

The developed model adds no computational complexity to the previous model, relying solely on existing data collection devices such as loop detectors. The developed model can be used in congestion prevention and energy management programs by homogeneously distributing traffic flow throughout the network.

The rest of the paper is organized as follows: first, in the Section 2, the tools used for emission estimation are introduced; secondly, in Section 3, our modelling setup is introduced; thirdly, in Section 4 our modelling results for our case studies are presented and discussed. Then, a discussion of the results is provided. Finally, the concluding remarks, policy recommendations, and future research directions are provided.

2. Review of Methodology

The emission estimation using fundamental traffic relationships requires the integration and calibration of several models. This section focuses on the methods used to develop the emission macroscopic fundamental diagram (e-MFD). Table 2 shows the variables used in this study and their definition.

Table 2. Notation.

Variable	Definition
$Q(t)$	Network flow at time t
$K(t)$	Network density at time t
$K_{std}(t)$	Network standard deviation of density at time t
$V(t)$	Network speed at time t
N	Total number of links in the network
$q_i(t)$	Flow of link i at time t
l_i	Length of link i
$k_i(t)$	Density of link i at time t
$v_i(t)$	Speed of link i at time t
M	Number of non-zero weights in standard deviation
$E_n(t)$	Emission rate of the pollutant (g/s) for vehicle n at time t
E_0	Pollutant's lower limit
f_j	Model parameter j
$v_n(t)$	speed of the vehicle n at time t
$a_n(t)$	acceleration of the vehicle n at time t
is_{ff}	Dummy variable for separating the free flow and congested branch of the e-MFD
c_k	Estimated model parameter k
p_s	Penetration rate of vehicle type s
α & β	Fixed model parameter

2.1. Macroscopic Fundamental Diagram (MFD)

The macroscopic fundamental diagram was first introduced in the late 2000s by [23]. MFD shows the relationship between network density (accumulation) and flow (production). Field experiments validated the existence of such a relationship between accumulation and production [24]. MFD brought many advantages, such as being independent of demand and being a property of network features, drivers' behaviour, and control policy [23], which make it an excellent tool for traffic state estimation. Network flow, density, speed, and standard deviation of density in MFD are calculated as Equations (1)–(4), respectively. The standard deviation of density of all N links at time t ($K_{std}(t)$) was calculated based on Equation (4) in which M stands for the number of non-zero weights (number of links here). It is worth mentioning that loop detectors are sufficient for calculating the standard deviation of density in practice. In this paper, the MFD is assumed to represent the network traffic state and will have full access to link-level information for all links.

$$Q(t) = \frac{\sum_{i=1}^{N} q_i(t) \times l_i}{\sum_{i}^{N} l_i} \tag{1}$$

$$K(t) = \frac{\sum_{i=1}^{N} k_i(t) \times l_i}{\sum_{i}^{N} l_i} \tag{2}$$

$$V(t) = \frac{\sum_{i=1}^{N} v_i(t) \times l_i}{\sum_{i}^{N} l_i} \tag{3}$$

$$K_{STD}(t) = \sqrt{\frac{\sum_{i=1}^{N} l_i \times (k_i(t) - K(t))^2}{\frac{M-1}{M} \sum_{i}^{N} l_i}} \tag{4}$$

2.2. Emission Estimation

Panis et al. microscopic emission model [7] (the default microscopic emission model in Aimsun traffic simulator [25]) was used to estimate the emissions produced from the simulated vehicles. All vehicles were modelled as gasoline passenger cars for the sake of simplicity. This microscopic model uses each vehicle's instantaneous speed and acceleration

to calculate the emission rate using Equation (5). Please refer to [7] for more information about the calibration process.

$$E_n(t) = max[E_0, f_1 + f_2 v_n(t) + f_3 v_n(t)^2 + f_4 a_n(t) + f_5 a_n(t)^2 + f_6 a_n(t) v_n(t)] \quad (5)$$

Emission level from individual vehicles was aggregated every minute to calculate the network-level emission rate. The network level emission was used for calibrating our model and as the ground truth for evaluating the accuracy of the calibrated model. In this paper, because of limited space, we only focused on CO_2, which is a greenhouse gas responsible for climate change [13]. However, our methodology applies to other pollutants as well.

3. Modelling Framework

The first step in our methodology is acquiring network-level emission and traffic data. The data could be collected from the field or generated through microscopic simulation software to mimic real-life scenarios. Since collecting different traffic congestion levels and incident data in real life is costly, this study used Aimsun [25], a microscopic simulation platform, to model the proposed scenarios used in this study. The traffic demand, traffic control, routing strategies, and emission model [7] of each specific scenario were imported into the simulation software. Accordingly, traffic states (i.e., speed, density, and standard deviation of density) and network pollution were generated by Aimsun and used for e-MFD model calibration. Two networks were used in this study. A synthetic Manhattan-like grid network and Blacksburg, VA, USA, with a realistic demand level and traffic signal setting. Three different scenarios were considered to validate the developed model for the grid network. The first scenario considered a uniform demand pattern, similar to what was used for model calibration but with different random seeds to mimic the daily traffic dynamics variability. The second scenario considered a directed traffic demand to simulate rush hours, evacuation, and special event situations where a large portion of trips are generated from and attracted to specific regions in the network. The last scenario considered a lane closure to resemble a traffic incident situation. Only the similar demand to calibration with different random seeds scenarios (like the first scenario for the grid network) was used to validate the developed model for the Blacksburg network. Figure 1 shows our modelling framework.

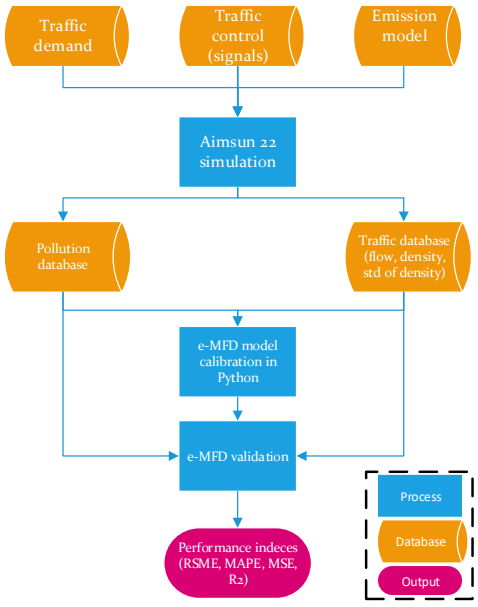

Figure 1. Modelling Framework.

4. Numerical Experiment

4.1. Grid Network

The e-MFD model was developed, calibrated, and validated for a Manhattan-like synthetic network. The network consists of 7 × 7 gird of 200 m blocks with trip generation and attraction points located in the middle of each block of one-way, two-lane roads. This type of network has been widely used in network-level model development, testing, and validation because of the homogeneity of links and simplicity of analysis. Pre-timed traffic signals with 60 s of cycle time and 28 s of green time and 2 s of inter-phase time were used to control the intersections. A spatially uniform demand with temporal increments was used. The grid network and the demand profile in the base scenario are shown in Figure 2. In order to observe e-MFD behaviour in congested and free-flow regimes, the traffic flow was gradually increased until the network reached a gridlock status. Also, to avoid hysteresis phenomena, no unloading time was given to the network [9,26,27]. The base scenario experiment was replicated ten times. Seven replications were used for model calibration, and three were used for model validation. We tested the calibrated model based on the seven base scenario replications in three unseen replication of the same scenario, the accident scenario and the directed demand scenario.

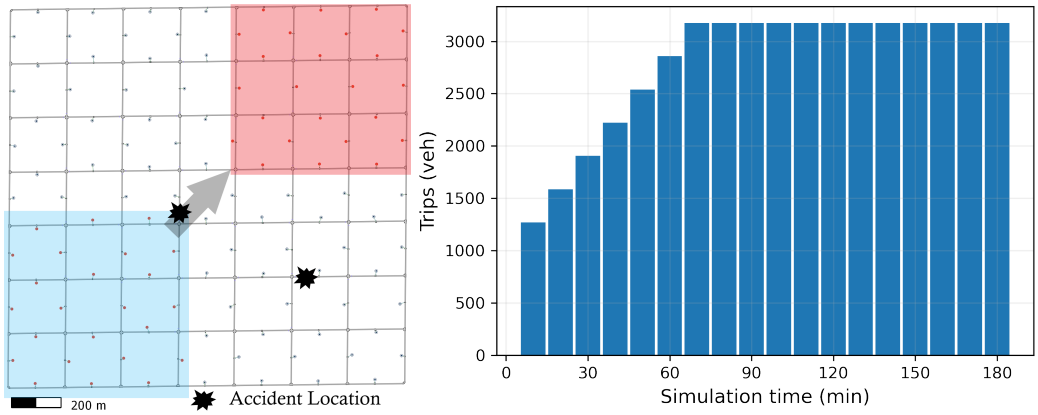

Figure 2. Synthetic grid network (**left**), base scenario demand profile (**right**).

4.1.1. Non-Uniform Traffic Demand Effect

The uniform spatial traffic demand pattern was altered to a directed demand pattern by assigning main traffic generation and attraction zones. Two trip attraction and generation zones were defined in the network as shown in Figure 2. The trips generated from the blue shaded area are attracted by the red shaded area. The traffic demand was increased by 20% in the designated areas. This scenario was replicated three times with different random seeds to add randomness and the results. The outputs were fed into the model developed using the base scenario. Then, the performance of the model was tested.

4.1.2. Accident Effect

Two accidents were created to test the model stability in the presence of disruption in the network. The location of each accident is marked in Figure 2. Each accident closed one lane from the corresponding section during the simulation period. Accident locations were selected based on their density level at the end of the second hour. Previous studies showed that in a grid network, the accidents that occurred on the link closer to the center of the network have more impact on the network performance, such as the trip completion rate and travel time [27,28]. Accordingly, this study followed the same logic for accident location selection. Each accident scenario was replicated three times with different random seeds.

4.1.3. e-MFD Model Development

In this section, the relationship between the network-level traffic variables and the levels of emissions is established. Later, the model calibration and validation are explained.

Figure 3A shows the relationship between the regular MFD (flow and density) and the level of CO_2 in the network (shown in colours). The MFD suggests that the increase in the level of congestion is associated with an increase in the level CO_2. The MFD shows that the critical density is 35 veh/km. Based on this observation, a binary variable is created (Is_{ff}, which takes value one if the density is smaller than 35 veh/km) to separate free flow and congested. The relationship between density and CO_2 is shown in Figure 3B. The Pearson correlation coefficient suggests a high correlation between density and CO_2 (r = 0.93). Figure 3B shows that CO_2 increases at a higher rate in the free flow branch and increases at a lower rate as the flow reduces after reaching capacity. However, the total emission amount is higher in the congested branch despite increasing at a lower rate than in the free-flow branch. Figure 3C shows an inverse relationship between speed and CO_2, which is confirmed by the large and negative Pearson correlation coefficient (r = −0.92). A mild slope (for lower speed values) followed by a steep slope (for higher speed values) were observed. Figure 3D shows a positive relationship between the standard deviation of density and CO_2, which is supported by the large correlation coefficient (r = 0.88). Before reaching the capacity, the standard deviation of density and CO_2 increase at a higher rate than after reaching the capacity. Also, there is no strong relationship between the standard deviation of density and CO_2 for standard deviation values larger than 100 veh/km. Similar results were reported about the relationship between the standard deviation of density and flow at different congestion levels [22]. The increase in standard deviation density at lower density levels reduced the flow. However, the density standard deviation did not impact flow at higher congestion levels. Figure 3E represents the emission MFD (e-MFD) for this grid network which relates network speed and density to the total CO_2 observed in the network. The free flow branch of our e-MFD (density < 35) is similar to what [9] reported for Athens, Greece. However, the congested branch was not observed in realistic cases due to congestion prevention.

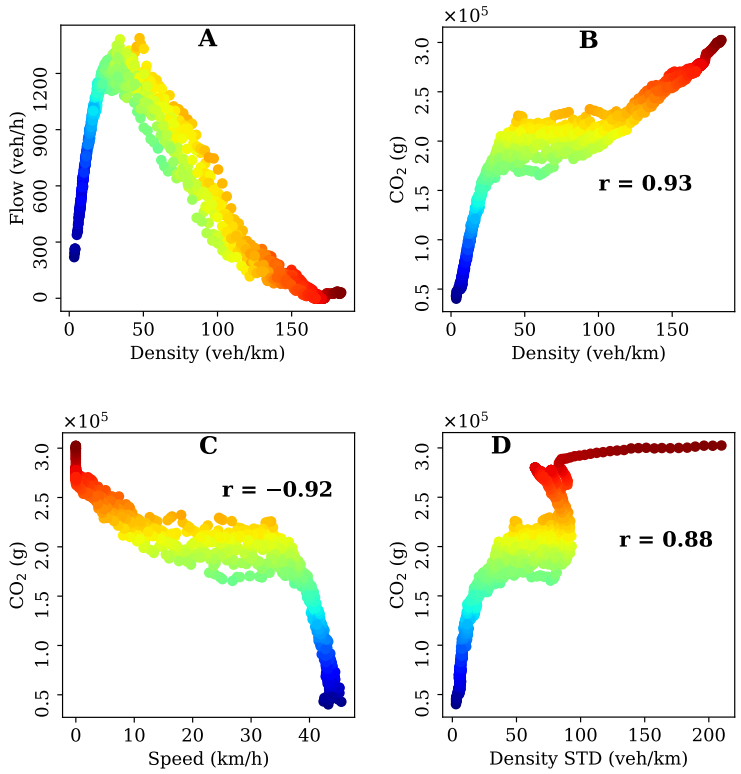

Figure 3. *Cont.*

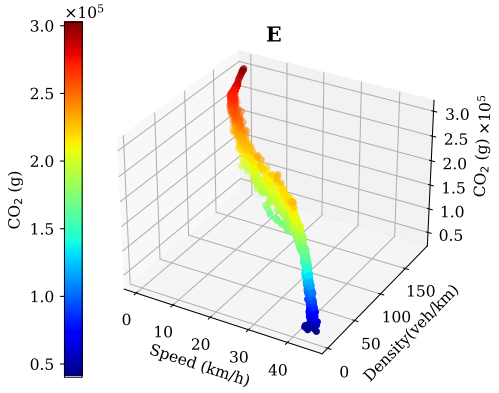

Figure 3. Grid network base scenario results: (**A**) MFD; (**B**) density and CO_2; (**C**) speed and CO_2; (**D**) standard deviation of density and CO_2; (**E**) e-MFD.

4.1.4. Model Calibration

The Ordinary Least Squares (OLS) method was used for model calibration. The non-linear model proposed in [13] is extended for the development of a heterogeneity-aware e-MFD model (We acknowledge that [13] offered two models, an SVR model and a non-linear regression model. However, SVR is a black box approach [29], which is not explainable. This paper only focused on the non-linear regression model for comparison.). The non-linear model [13] is demonstrated in Equation (6) where CO_2 is the network CO_2 level, c_s and c_v are model parameters to estimate, and p_s stands for the penetration rate of vehicle type s. In this paper, we did not consider the summation part in Equation (6) since we only modelled gasoline passenger cars. The extended heterogeneity-aware e-MFD is shown in Equation (7) where V stands for speed, and is_{ff} is a dummy variable for separating the free flow and congested branch of the e-MFD. To gain better estimation results such as higher R^2 and identify significant parameters, we used a transformation on standard deviation density ($K_{std}(t)$), i.e., $(\alpha^{K_{std}(t)})^{\beta}$. We also opted to use the third power of density instead of the first (see [30] for more information about the suitability of a third-degree polynomial for MFD) (These transformations are identified by trial and error.). $\{c_1, c_2, c_3\}$ are the model parameters to calibrate while we fixed $\alpha = 0.8$ and $\beta = 0.3$ (The values of α and β, as well as $\{c_1, c_2, c_3\}$ should be calibrated for each network (similar to [13]).). Table 3 summarized the calibration results, which indicate that all parameters of both models are statistically significant at a 95% confidence interval. We also provided the calibration results for the interaction terms in Table 3. Also, $[1]$ means the dummy variable $is_{ff} = 1$. Table 4 shows that both models offer high R^2 values, which shows the models explained a large portion of the variance in the data used for calibration. However, our model offered a higher R^2. Our model also outperformed the other model in other indices like MSE, RMSE, MAE, and MAPE.

$$CO_2 = (\sum_{s=1}^{S} c_s p_s) K(t)(c_v + V(t)) \tag{6}$$

$$CO_2(t) = c_1 \cdot is_{ff} \cdot (c_2 V(t) + c_3 K(t)^3)(\alpha^{K_{std}(t)})^{\beta} \tag{7}$$

Table 3. Model parameters (grid network).

Model	Parameter	Coef	Std Err	t	*p*-Value
Saedi [13]	K	1566.8498	10.880	144.018	0.000
	V × K	81.1724	1.282	63.303	0.000
Our model	$is_{ff}[0]$	1.961×10^5	1939.747	101.107	0.000
	$is_{ff}[1]$	-2.345×10^5	5.05×10^4	−4.645	0.000
	V	−680.1761	63.875	−10.649	0.000
	$is_{ff}[1] \times$ V	9402.2533	1225.045	7.675	0.000
	K^3	0.0176	0.000	40.801	0.000
	$is_{ff}[1] \times K^3$	1.6356	0.166	9.846	0.000
	$(0.8_s^K td(t))^{0.3}$	6.168×10^6	5.3×10^5	11.634	0.000
	$is_{ff}[1] \times (0.8_s^K td(t))^{0.3}$	-5.489×10^6	5.35×10^5	−10.253	0.000
	V $\times (0.8_s^K td(t))^{0.3}$	-1.614×10^5	1.45×10^4	−11.127	0.000
	$is_{ff}[1] \times$V $\times (0.8_s^K td(t))^{0.3}$	1.43×10^5	1.46×10^4	9.784	0.000
	$K^3 \times (0.8_s^K td(t))^{0.3}$	−1.4151	0.114	−12.457	0.000
	$is_{ff}[1] \times K^3 \times (0.8_s^K td(t))^{0.3}$	16.2529	0.842	19.292	0.000

Table 4. Training evaluation results (grid network).

Model	Scenario	MSE	RMSE	MAE	MAPE	R^2	R^2_{adj}
Saedi [13]	Base	1.280×10^9	35,772.217	30,674.165	0.215	0.765	0.764
Our model	Base	3.133×10^7	5597.562	3947.204	0.028	0.994	0.994

4.1.5. Model Validation

The performance of the developed models was tested in different scenarios, and the results were visualized and numerically compared. Figure 4 shows the model validation results. In all subplots in Figure 4 an underestimation for lower CO_2 values (<2×10^5 g) and an overestimation for higher than 2×10^5 grams by Saedi model was observed. Conversely, our model does not have this issue. As Figure 3 shows, it can be found that 2×10^5 g CO_2 corresponds to the critical density (35 veh/km). As Figure 4A,D show, the estimated/observed scatter plot is better aligned around a 45-degree line for our model. Therefore, our model offers more consistency with microscopic model estimates for the three unseen replication with the same demand. Figure 4B,E show the directed demand scenario validation results. Saedi et al. model shows inconsistencies with the ground truth for certain CO_2 values. Contrarily, our estimations are closer to the 45-degree trend line. Figure 4C,F show the results of the accident scenario validation. In this scenario, once again, our model showed superiority in terms of estimation accuracy. In the accident and directed demand scenarios, congestion quickly builds up due to the blockage and capacity lost. It pushes the network to the complete gridlock state where network speed is close to zero, and higher CO_2 levels are predicted for that speed range by Saedi et al. model. However, our model using the congestion distribution term and the dummy variable for free-flow could handle these scenarios better. Table 5 shows the numerical validation results of testing both models in different scenarios. As Table 5 shows, our model offers a lower Mean Square of Error (MSE), Root Mean Square of Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage of Error (MAPE) and higher R^2 and higher adjusted R^2, which all indicate the superior performance and the flexibly of our model in handling different scenarios compared to the previous model. Please note these R^2 and $R^2_{adjusted}$ are calculated based on the model's performance in explaining the variance of unseen data, which is different from the R^2 reported for explaining the variance of seen data in Table 4.

Table 5. Model validation results for the grid network.

Model	Senario	MSE	RMSE	MAE	MAPE	R^2	R^2_{adj}
Saedi [13]	Base	1.307×10^9	36,157.204	31,879.340	0.221	0.748	0.747
	Directed demand	1.087×10^9	32,972.903	28,570.864	0.186	0.821	0.821
	Accident	1.193×10^9	34,538.044	29,483.985	0.205	0.784	0.783
Our model	Base	6.992×10^7	8361.546	5774.478	0.036	0.987	0.986
	Directed demand	6.992×10^7	14,314.160	9579.152	0.051	0.966	0.966
	Accident	6.229×10^7	7892.216	5192.642	0.039	0.989	0.988

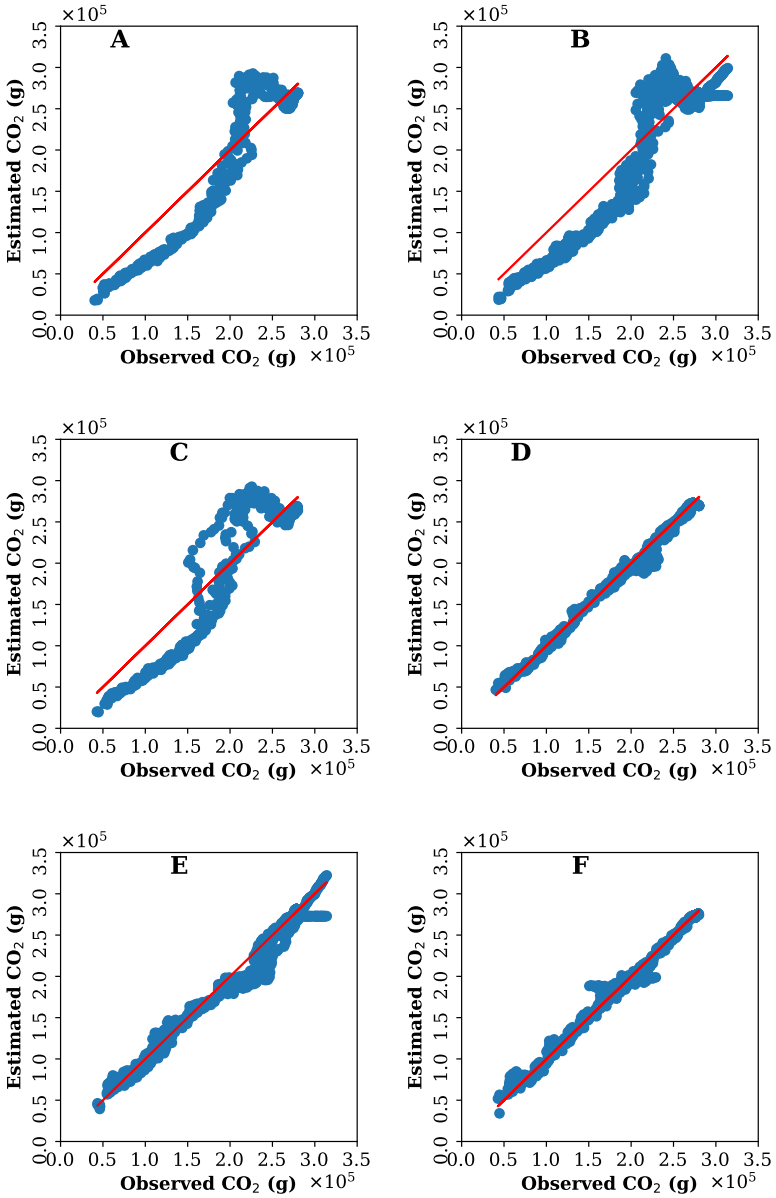

Figure 4. Validation results: (**A**) with unseen balanced scenario using Saedi model; (**B**) unseen directed demand scenario using Saedi model; (**C**) unseen accident scenario using Saedi model; (**D**) with unseen balanced scenario using our model; (**E**) unseen directed demand scenario using our model; (**F**) unseen accident scenario using our model.

4.2. Blacksburg, VA Network

Blacksburg covers an area of around 19.77 sq mi (51.20 km^2) and populates around 44,826 people [31]. The transportation network includes US 460, a high-speed 4-lane highway. We simulated this network in Aimsun with a realistic demand level based on traffic surveys done by the city in 2016. Figure 5 shows Blackburg network and the demand profile we used for loading this network. The simulation duration was four hours from 5–9 AM, which generated 96,702 vehicles. The model includes 2022 links and 665 nodes. Pre-timed traffic signals based on realistic timing plans acquired from the city were used in Aimsun to control traffic at the major arterial corridors. For the sake of simplicity, we replaced all vehicles with gasoline passenger cars. Due to the long simulation run time, we replicated the simulation four times (mimicking four different days) with different random seeds. We used three replications for model calibration and one for model validation.

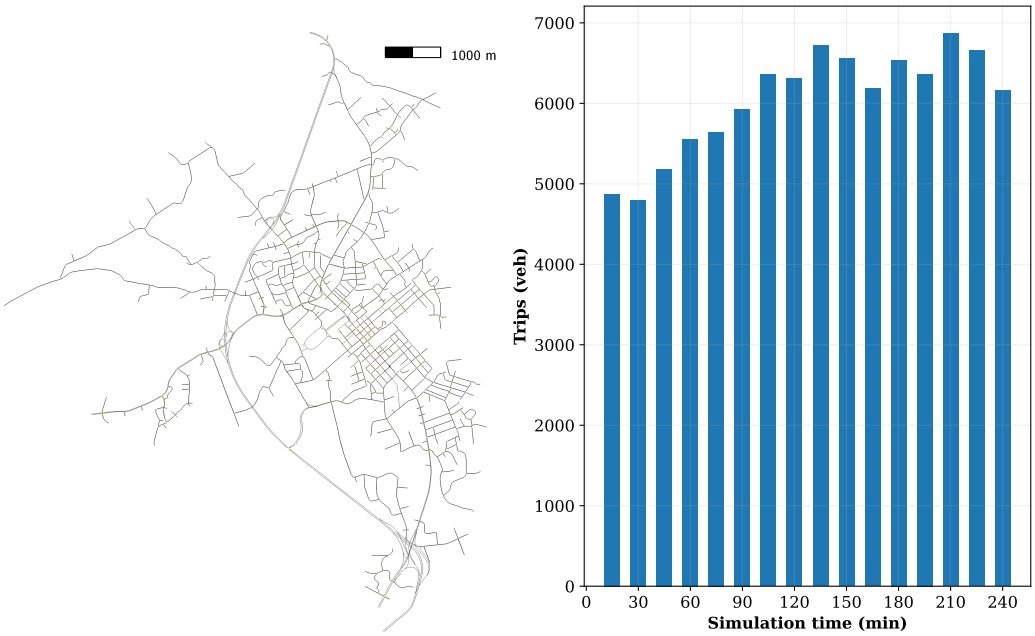

Figure 5. Blacksburg, VA: Network (**left**), demand (**right**).

4.2.1. Simulation Results: Blacksburg

The simulation results for Blacksburg, VA network are shown in Figure 6. The results show that the network quickly transfers from the free flow regime to the capacity state at 5 veh/km density. Between the density of 5–20 veh/km, the network maintains approximately the same flow, which remarks the capacity state in the network. For density values larger than 20 veh/km, network flow reduces, which shows the start of the congested branch. Please note that due to the limited routing options, the grid network immediately transferred from the free-flow state to the congested state. On the other hand, due to more routing options in the Blacksburg network, the network experience a capacity state for a large range of density values similar to what [32] reported for Brisbane, Australia. We observed high Pearson correlation values for the relationship between density, speed, and standard deviation of density with CO_2, shown in Figure 6. The e-MFD is shown in Figure 6. Similar to the grid network e-MFD, we see two branches here. However, the free-flow branch for Blacksburg includes a smaller range of densities than what we observed for the grid network. Also, as Figure 6B–D showed, the congested and capacity branch look similar in terms of the relationship between CO_2 and the traffic state variables. Therefore, we opted to keep our e-MFD model bi-modal (free-flow and congested).

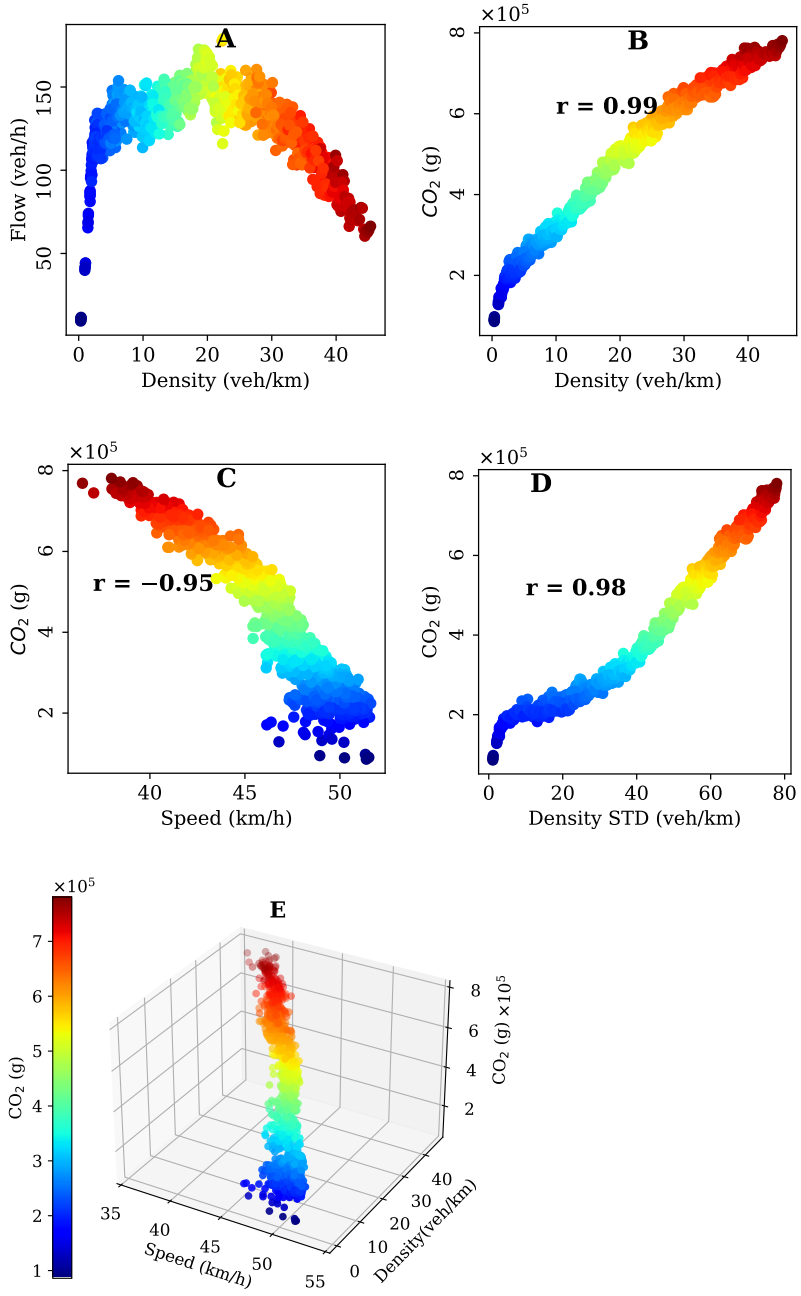

Figure 6. Simulation results for Blacksburg, VA: (**A**) MFD; (**B**) density and CO_2; (**C**) speed and CO_2; (**D**) standard deviation of density and CO_2; (**E**) e-MFD.

4.2.2. Model Calibration

The same equations as the grid network parameter estimation were used for the Blacksburg network. The only difference between them is the change we made in the standard deviation term in Equation (7) to $(0.8^{K_{std}(t)})^{0.86}$ for better fitting results. Calibration results are present in Table 6. Results show that all parameters of the Saedi model were statistically significant at a 95% confidence interval. Similarly, all parameters of our model were statistically significant at 95% interval except for the dummy variable coefficient for free-flow condition ($is_{ff} = 1$), which was very close to 0.05 (0.056). This observation could be attributed to the fewer observation for the free-flow branch. Table 7 shows model training evaluation measures. As Table 7 shows, our model offered far better results compared to Saedi et al. [13] model.

Table 6. Model parameters (Blacksburg network).

Model	Parameter	Coef	Std Err	t	*p*-Value
Saedi [13]	K	-3.252×10^4	1559.521	−20.856	0.000
	K × V	1266.8478	36.489	34.719	0.000
Our model	$is_{ff}[0]$	9.093×10^5	4.52×10^4	20.095	0.000
	$is_{ff}[1]$	1.784×10^5	9.34×10^4	1.910	0.056
	V	−7318.9981	997.966	−7.334	0.000
	$is_{ff}[1] \times$V	7209.3231	2091.202	3.447	0.001
	K^3	2.0439	0.111	18.423	0.000
	$is_{ff}[1] \times K^3$	436.9824	77.302	5.653	0.000
	$(0.8_s^K td(t))^{0.86}$	-6.09×10^7	1.97×10^7	−3.096	0.002
	$is_{ff}[1] \times (0.8^{K_{std}(t)})^{0.86}$	6.108×10^7	1.97×10^7	3.105	0.002
	$V \times (0.8^{K_{std}(t)})^{0.86}$	1.539×10^6	3.94×10^5	3.906	0.000
	$is_{ff}[1] \times V \times (0.8^{K_{std}(t)})^{0.86}$	-1.545×10^6	3.94×10^5	−3.920	0.000
	$K^3 \times (0.8^{K_{std}(t)})^{0.86}$	-2.728×10^5	5663.503	−48.161	0.000
	$is_{ff}[1] \times K^3 \times (0.8^{K_{std}(t)})^{0.86}$	2.877×10^5	6201.035	46.393	0.000

Table 7. Training evaluation results (Blacksburg network).

Model	MSE	RMSE	MAE	MAPE	R^2	R^2_{adj}
Saedi [13]	3.557×10^9	59,640.855	44,775.987	0.159	0.985	0.888
Our model	2.785×10^8	16,686.847	12,518.739	0.033	0.991	0.991

4.2.3. Model Validation

The validation results are provided in Table 8. The validation results show the superiority of our model for all validation metrics compared to Saedi model. Please note these R^2 and $R^2_{adjusted}$ are calculated based on the model's performance in explaining the variance of unseen data, which is different from the R^2 reported for explaining the variance of seen data. Since the models were validated against the same demand pattern, higher R^2 than observed for the grid network. The similar performance of our model and Saedi model estimates is because of the validation of the model against a similar scenario as the calibration process. In contrast, more challenging scenarios such as accidents and directional demand would produce a more considerable difference between our and Saedi model. Still, our model offered around 10 percent higher R^2. Nevertheless, since we planned to keep this case study similar to the realistic demand, we did not create a directional demand pattern scenario for Blacksburg. Figure 7A,B show the validation results for different values of CO_2. As Figure 7A,B show, our model estimates CO_2 levels with higher accuracy for all ranges of CO_2 values compared to Saedi model. Similar to what we observed for the grid network, Saedi model underestimates lower CO_2 values, while our model does not show any underestimation for any CO_2 value. However, the differences in validation results that we observed here are not as considerable as the grid network. Figure 7C,D show the estimation results compared to microscopic model results over time. Figure 7E,F show the cumulative CO_2 rate for different density levels for our model and Saedi model. As the figures show, Saedi model underestimates the CO_2 level for the densities lower than 30 veh/km, which is the density range before the start of the congested branch (see Figure 6A). On the other hand, our model estimations are close to microscopic estimates for all density values. Including the dummy variable for the free flow branch and the standard deviation of density made our model more flexible since we tested the performance of our model without those terms, and we observed similar results to Saedi model (i.e., underestimation). However, our model slightly overestimates CO_2 compared to the microscopic model. This overestimation is not as considerable as Saedi model. On the bright side, makes our model slightly more conservative than the microscopic model, which could be beneficial for emission control [5].

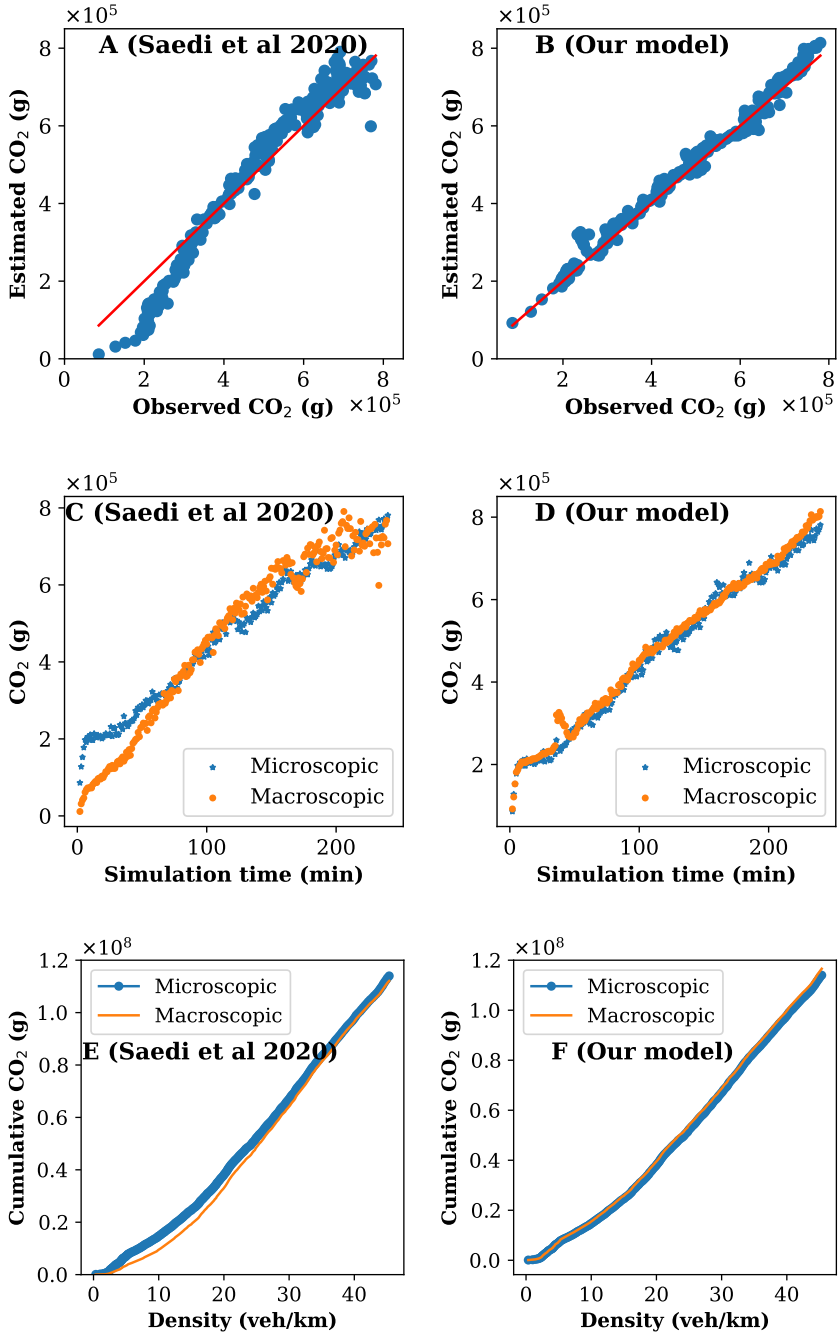

Figure 7. Validation results for Blacksburg Network: (**A**) Saedi model [13]; (**B**) Our model; (**C**) CO_2 rate (Saedi [13]); (**D**) CO_2 rate (Our model); (**E**) Cumulative CO_2 rate (Saedi [13]); (**F**) Cumulative CO_2 rate (Our model).

Table 8. Validation results for Blacksburg network.

Model	MSE	RMSE	MAE	MAPE	R^2	R^2_{adj}
Saedi [13]	3.338×10^9	57,771.680	46,061.040	0.144	0.900	0.899
Our model	5.287×10^8	22,992.509	17,630.351	0.043	0.984	0.983

5. Discussion

Our results showed that separating the free-flow and congestion branches and including the standard deviation of density could improve the estimation results of the e-MFD.

Previous literature showed the standard deviation of density impacts the flow [21,22] and the inclusion of this term could lead to calibrating a more accurate MFD. However, standard deviation of density impact depends on the level of density [22]. Moreover, similar to [22,30] suggestion for MFD, we used the third power of density in our e-MFD. However, we opted to drop the first and the second power from our e-MFD since they were not statistically significant in our calibration trials. Saedi model showed underestimation for the free-flow branch of e-MFD and overestimation for the congested branch of e-MFD in both numerical experiments. Interestingly, when we solely imported the standard deviation of density without the dummy variable for free-flow, we observed the same behaviour for the free-flow branch of our model. Therefore, we suggest incorporating the dummy variable and the standard deviation of density in the e-MFD. This suggestion is in harmony with the dependence of standard deviation of density impact on flow to the congestion level [21,22]. As [13] stated, macroscopic emission rates correlate highly with the flow. However, this relationship is not strictly linear and can have different slopes for free-flow and congested phases. While the original single modal e-MFD was not able to capture those different slopes, we showed that by accounting for the traffic state (free-flow and congested) and including the standard deviation of density, the calibrated model stayed more flexible in different scenarios and across different congestion levels (for example, Figure 7F).

We observed more considerable differences between the performance of our model and the previous model for the grid network compared to Blacksburg network. This less considerable difference in model performance can be attributed to the availability of alternative routes in Blacksburg network (US 460 highway that passes by Blacksburg). Contrarily, in a grid network, density's standard deviation grows quickly due to the limited routing options and the nucleation effect [21]. While the grid network results highlight our model's superiority for urban networks with more homogeneous structures (CBD, downtown alike structures), our model stays relevant and offers better estimation results compared to Saedi model. Around 10% improvement in R^2 without the need for any new hardware installation or data collection) for the less homogeneous networks (CBD + highways). We speculate that partitioning the network [33] into more homogeneous sub-networks could improve the estimation accuracy and make emission control more feasible. Future research work can explore the potential benefits of network partitioning in improving e-MFD performance.

6. Conclusions

Previous studies showed that congestion distribution plays a critical role in traffic state estimation [22,30]. However, the potential benefits of this important variable in improving the emission estimation model's accuracy were not explored. In this paper, we presented an enhanced version of the emission macroscopic fundamental diagram (e-MFD), which included the standard deviation of the density as an independent variable, along with density and speed. Our observations suggested e-MFD is bi-modal (free flow and congested branch). With that in mind, we added a dummy variable for separating the free-flow and congested branches to our e-MFD model.

The original e-MFD showed promising, and accurate results [13]. However, the stability of the e-MFD was not explored. Also, we showed the e-MFD accuracy could be improved by adding the standard deviation of density, which was missing in previous studies. We tested the performance of our model in two case studies, a synthetic grid network and Blacksburg, VA network. Also, three challenging scenarios were selected to test the stability of the model in the grid network. First, the day-to-day stability was done by testing the developed model against unseen simulated data with different random seeds. The stability in directional demand pattern scenarios and accident scenarios were the two other experimental scenarios considered for testing the stability of the proposed model. The contributions of the current study are twofold: Offering an advanced version of e-MFD, which performs better than the original e-MFD in emission estimation without

requiring new data collection. Also, we tested the stability of the original e-MFD and our e-MFD and showed the higher stability of our model. Moreover, this study highlights the critical role of network congestion homogeneity, which is crucial in maintaining network performance close to the desired state. Our model can be used in realistic cases for emission by importing the speed, density, and standard deviation of density, which are all easily accessible from loop detectors. We achieved higher emission estimation accuracy by adding only one single variable to the existing model, which can be obtained from loop detectors. This research highlighted the critical role of navigation in emission reduction as well. Our correlation analysis showed lower emission rates could be expected by keeping density's standard deviation low through better navigation and more homogeneous congestion distribution. Furthermore, the model could be used for emission reduction in hierarchical traffic controllers [22,34], which were proven to be very effective in improving network performance. Such controllers use the congestion level and congestion homogeneity of each region presented in the proposed model.

In this paper, we considered some simplifying assumptions which could be relaxed in future research, such as: adding different vehicle types and conducting a sensitivity analysis of our model for different vehicle combinations. This analysis will help the network managers to optimize the network modal share to achieve minimum pollution. For example, previous studies showed that the effect of public transportation lines on regional emission levels could be considerable [35]. Another interesting research frontier could be removing the full-coverage assumption we made in this paper. In real life, a limited number of network links are equipped with loop detectors, and we acknowledge this partial coverage could impact our model's performance. However, finding the critical links (similar to [36] about MFD) to capture network state with limited coverage could solve this issue. Furthermore, the e-MFD of a network can be utilized to develop advanced emission mitigation systems. In such a system, the network's emissions level is monitored and predicted with e-MFD to search for the most efficient traffic management strategies, such as cooperative signal control, dynamic routing, and variable speed limit, to minimize the total network-level emissions.

Author Contributions: Conceptualization, M.H., H.Y. and H.A.; Methodology, M.H.; Software, M.H.; Validation, M.H.; Formal analysis, M.H.; Investigation, M.H.; Writing—original draft, M.H.; Writing—review & editing, M.H., H.Y. and H.A.; Visualization, M.H.; Supervision, H.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. EPA. Inventory of U.S. Greenhouse Gas Emissions and Sinks. 2022. Available online: https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks (accessed on 27 July 2022).
2. Ullah, I.; Liu, K.; Yamamoto, T.; Zahid, M.; Jamal, A. Prediction of electric vehicle charging duration time using ensemble machine learning algorithm and Shapley additive explanations. *Int. J. Energy Res.* **2022**, *46*, 15211–15230. [CrossRef]
3. Ullah, I.; Liu, K.; Yamamoto, T.; Al Mamlook, R.E.; Jamal, A. A comparative performance of machine learning algorithm to predict electric vehicles energy consumption: A path towards sustainability. *Energy Environ.* **2022**, *33*, 1583–1612. [CrossRef]
4. Ullah, I.; Liu, K.; Yamamoto, T.; Shafiullah, M.; Jamal, A. Grey wolf optimizer-based machine learning algorithm to predict electric vehicle charging duration time. *Transp. Lett.* **2022**. [CrossRef]
5. Ingole, D.; Mariotte, G.; Leclercq, L. Minimizing network-wide emissions by optimal routing through inner-city gating. *Transp. Res. Part D Transp. Environ.* **2020**, *86*, 102411. [CrossRef]
6. Ntziachristos, L.; Gkatzoflias, D.; Kouridis, C.; Samaras, Z. COPERT: A European road transport emission inventory model. In *Information Technologies in Environmental Engineering*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 491–504.

7. Panis, L.I.; Broekx, S.; Liu, R. Modelling instantaneous traffic emission and the influence of traffic speed limits. *Sci. Total. Environ.* **2006**, *371*, 270–285. [CrossRef] [PubMed]
8. Rakha, H.; Van Aerde, M.; Ahn, K.; Trani, A. Requirements for evaluating traffic signal control impacts on energy and emissions based on instantaneous speed and acceleration measurements. *Transp. Res. Rec.* **2000**, *1738*, 56–67. [CrossRef]
9. Barmpounakis, E.; Montesinos-Ferrer, M.; Gonzales, E.J.; Geroliminis, N. Empirical investigation of the emission-macroscopic fundamental diagram. *Transp. Res. Part D Transp. Environ.* **2021**, *101*, 103090. [CrossRef]
10. CARB. EMFAC2014 Volume III—Technical Documentation. 2015. Available online: https://www.sandiegocounty.gov/content/dam/sdc/pds/ceqa/JVR/AdminRecord/IncorporatedByReference/Appendices/Appendix-P---Greenhouse-Gas-Emissions-Report/CARB%202015B.pdf (accessed on 27 July 2022).
11. Csikós, A.; Tettamanti, T.; Varga, I. Macroscopic modeling and control of emission in urban road traffic networks. *Transport* **2015**, *30*, 152–161. [CrossRef]
12. Rakha, H.; Yue, H.; Dion, F. VT-Meso model framework for estimating hot-stabilized light-duty vehicle fuel consumption and emission rates. *Can. J. Civ. Eng.* **2011**, *38*, 1274–1286. [CrossRef]
13. Saedi, R.; Verma, R.; Zockaie, A.; Ghamami, M.; Gates, T.J. Comparison of Support Vector and Non-Linear Regression Models for Estimating Large-Scale Vehicular Emissions, Incorporating Network-Wide Fundamental Diagram for Heterogeneous Vehicles. *Transp. Res. Rec.* **2020**, *2674*, 70–84. [CrossRef]
14. Ji, Y.; Dong, J.; Jiang, H.; Wang, G.; Fei, X. Research on carbon emission measurement of Shanghai expressway under the vision of peaking carbon emissions. *Transp. Lett.* **2022**. [CrossRef]
15. Ntziachristos, L.; Samaras, Z.; Eggleston, S.; Gorissen, N.; Hassel, D.; Hickman, A.; Joumard, R.; Rijkeboer, R.; White, L.; Zierock, K.-H. COPERT III. *Computer Programme to Calculate Emissions from Road Transport, Methodology and Emission Factors (Version 2.1)*; European Energy Agency (EEA): Copenhagen, Denmark, 2000.
16. Rakha, H.; Wang, J.; Abdelmegeed, M. *Develop Multi-Scale Energy and Emission Models*; Technical Report; National Institute for Advanced Transportation Technology, University of Idaho: Moscow, ID, USA, 2016.
17. Sun, Z.; Hao, P.; Ban, X.J.; Yang, D. Trajectory-based vehicle energy/emissions estimation for signalized arterials using mobile sensing data. *Transp. Res. Part D Transp. Environ.* **2015**, *34*, 27–40. [CrossRef]
18. Jamshidnejad, A.; Papamichail, I.; Papageorgiou, M.; De Schutter, B. A mesoscopic integrated urban traffic flow-emission model. *Transp. Res. Part C Emerg. Technol.* **2017**, *75*, 45–83. [CrossRef]
19. Zegeye, S.; De Schutter, B.; Hellendoorn, J.; Breunesse, E.; Hegyi, A. Integrated macroscopic traffic flow, emission, and fuel consumption model for control purposes. *Transp. Res. Part C Emerg. Technol.* **2013**, *31*, 158–171. [CrossRef]
20. Gori, S.; La Spada, S.; Mannini, L.; Nigro, M. A dynamic mesoscopic emission model for signalized intersections. In Proceedings of the 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013), The Hague, The Netherlands, 6–9 October 2013; pp. 2212–2217.
21. Knoop, V.L.; Van Lint, H.; Hoogendoorn, S.P. Traffic dynamics: Its impact on the macroscopic fundamental diagram. *Phys. A Stat. Mech. Its Appl.* **2015**, *438*, 236–250. [CrossRef]
22. Ramezani, M.; Haddad, J.; Geroliminis, N. Dynamics of heterogeneity in urban networks: Aggregated traffic modeling and hierarchical control. *Transp. Res. Part B Methodol.* **2015**, *74*, 1–19. [CrossRef]
23. Daganzo, C.F.; Geroliminis, N. An analytical approximation for the macroscopic fundamental diagram of urban traffic. *Transp. Res. Part B Methodol.* **2008**, *42*, 771–781. [CrossRef]
24. Geroliminis, N.; Daganzo, C.F. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transp. Res. Part B Methodol.* **2008**, *42*, 759–770. [CrossRef]
25. Aimsun. Aimsun Next 22 Users Manual. 2022. Available online: https://docs.aimsun.com/next/22.0.1/ (accessed on 27 July 2022).
26. Gayah, V.V.; Daganzo, C.F. Clockwise hysteresis loops in the macroscopic fundamental diagram: An effect of network instability. *Transp. Res. Part B Methodol.* **2011**, *45*, 643–655. [CrossRef]
27. Halakoo, M.; Yang, H. Evaluation of Macroscopic Fundamental Diagram Transition in the Era of Connected and Autonomous Vehicles. In Proceedings of the 2021 IEEE Intelligent Vehicles Symposium (IV), Nagoya, Japan, 11–17 July 2021; pp. 1188–1193.
28. Ortigosa, J.; Menendez, M.; Gayah, V.V. Analysis of network exit functions for various urban grid network configurations. *Transp. Res. Rec.* **2015**, *2491*, 12–21. [CrossRef]
29. Filom, S.; Amiri, A.M.; Razavi, S. Applications of machine learning methods in port operations—A systematic literature review. *Transp. Res. Part E Logist. Transp. Rev.* **2022**, *161*, 102722. [CrossRef]
30. Knoop, V.L.; Hoogendoorn, S.P. Empirics of a generalized macroscopic fundamental diagram for urban freeways. *Transp. Res. Rec.* **2013**, *2391*, 133–141. [CrossRef]
31. U.S. Census Bureau. US Census Data. 2020. Available online: https://www.census.gov/2020results (accessed on 27 July 2022).
32. Tsubota, T.; Bhaskar, A.; Chung, E. Macroscopic fundamental diagram for Brisbane, Australia: Empirical findings on network partitioning and incident detection. *Transp. Res. Rec.* **2014**, *2421*, 12–21. [CrossRef]
33. Saeedmanesh, M.; Geroliminis, N. Clustering of heterogeneous networks with directional flows based on "Snake" similarities. *Transp. Res. Part B Methodol.* **2016**, *91*, 250–269. [CrossRef]
34. Yildirimoglu, M.; Sirmatel, I.I.; Geroliminis, N. Hierarchical control of heterogeneous large-scale urban road networks via path assignment and regional route guidance. *Transp. Res. Part B Methodol.* **2018**, *118*, 106–123. [CrossRef]

35. Rodriguez-Rey, D.; Guevara, M.; Linares, M.P.; Casanovas, J.; Salmerón, J.; Soret, A.; Jorba, O.; Tena, C.; García-Pando, C.P. A coupled macroscopic traffic and pollutant emission modelling system for Barcelona. *Transp. Res. Part D Transp. Environ.* **2021**, *92*, 102725. [CrossRef]
36. Saffari, E.; Yildirimoglu, M.; Hickman, M. A methodology for identifying critical links and estimating macroscopic fundamental diagram in large-scale urban networks. *Transp. Res. Part C Emerg. Technol.* **2020**, *119*, 102743. [CrossRef]