Supporting Information

Evaluation of the relationship between momentum wakes behind moving vehicles and dispersion of vehicle emissions using near roadway measurements

Yu Ting Yu[†‡], Sheng Xiang[†‡], Kenneth E. Noll[*†]

† Department of Civil, Architecture and Environmental Engineering, Illinois Institute of Technology, Chicago, IL 60616, USA.

‡ School of Environment, State Key Joint Laboratory of Environment Simulation and Pollution Control, Tsinghua University, Beijing 100084, PR China

*Corresponding author contact information: noll@iit.edu

There are 3 table and 2 figures in total in the SI.

Number of pages: 11

S1. Instrumentation. Table S1. Summary of technical description of the instruments used in this study.

Equipment	Manufactory	Measured variables	Time Resolution	Accuracy
AQ-Expert	E-Instruments	CO_2,	10 s	±2% of reading ±10 ppm
P-trak 8525	TSI instrument	UFP number concentration	1 s	Not specified
3D-ultrasonic anemometer	Gill instrument	Wind speed	1 s	<1.5% @ 12m/s
		Wind direction	1 s	<2º @ 12m/s
Digital record	Sony	Vehicle flow rate, Vehicle speed	n.a.	n.a.

In this study, all the instruments were operated simultaneously with 5-min sampling interval. Each 5-min was a separate sampling period.

S2. Determination of initial $\boldsymbol{\sigma_z}$. Based on the relationship between vehicle height and wake length shown in Fig.5 (regression line), the "wake length" for this study can be determined as 11.9 m and 100.4 m for cars and trucks, respectively (H_{LDV}=1.4 m and H_{HDV}=4.1 m). By applying the derived "wake length" and vehicle height into the "wake area model" (Eq. (7)-(9)) (see section 2.1), the "effective wake area" can be determined as 16.7 m^2 for one LDV and 411.7 m^2 for one HDV. With vehicle density that measured from field measurements and "wake area model" (Eq. (7)-(9)), $\sigma_{z,init,WAKE}$ that related to vehicles can be determined (see Table S2).

$\sigma_{z,init}$ for CALINE4 and AERMOD were back calculated by running the dispersion models. The vehicle flow rate, emission factors and meteorological data were first entered into CALINE4 and AERMOD to calculated pollutant concentrations. When running the CALINE4 and AERMOD, surface roughness was set as 0.01 m for LSD (covered by sands and grass) and 0.05 m for DRE (covered by grass and low-rise building)[1].Then applied Eq. (10) to back calculate $\sigma_{z,init}$ for CALINE4 and AERMOD.

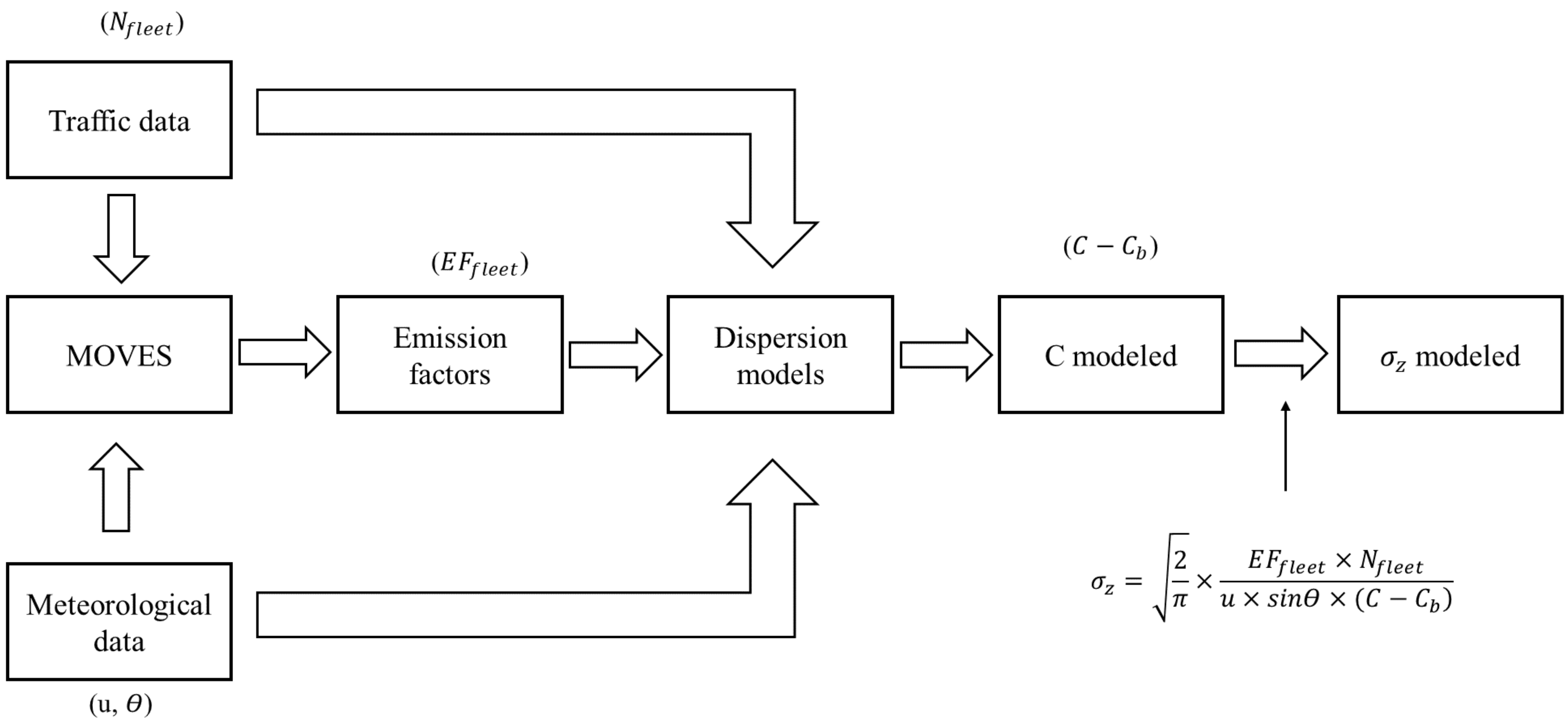

Figure S.1. Flow diagram of back calculated $\sigma_{z,init}$ for CALINE4 and AERMOD with critical factors identified.

S3. Calculation of Pollutant Concentrations. Pollutant concentrations were calculated by Eq. (S.1) that is a transformation of Eq. (4). The $C_{calculated}$ is the calculated pollutant concentrations contributed by vehicles (background subtracted). The definition of the components ($EF_{fleet}, N_{fleet}, u, \theta$) are provided in the main text (Section 1.2).

$$C_{calculated} = \sqrt{\frac{2}{\pi}} \times \frac{EF_{fleet} \times N_{fleet}}{u \times sin\theta \times \sigma_z} \qquad \text{Eq. (S.1)}$$

Pollutant concentrations were calculated by inputing $\sigma_{z,init,WAKE}$ parameterized in this study and directly running CALINE4 and AERMOD. Then the calculated pollutant concentrations were compared to our previous concentration measurements[2,3].

Table S1 presents the data from previous measurements and results of example calculation in this study near LSD and DRE. The data interval of 7,000 veh h^{-1} ≤ total flow rate < 8,000 veh h^{-1} near LSD and DRE is selected because it has the largest amount of retrievable data (n=173). The main calculation process of simulated pollutant concentrations is shown in Fig. S2.

Table S2. Example calculations using Eq. (S.1) for 7,000 veh $h^{-1} \leqslant$ total flow rate < 8,000 veh h^{-1} data interval near LSD and DRE from previous measurements[1, 2]. Only average data for each 1,000 data intervals under free flow driving conditions were used for comparison.

Roadway		LSD	DRE	DRE	DRE	DRE	DRE
Perpendicular wind speed (m)		1.3	1.7	1.7	1.7	1.7	1.7
CO_2 fleet EF ($\times 10^{-3}$ g $m^{-1} veh^{-1}$)		175	192	213	236	262	300
UFP fleet EF (pt $m^{-1} veh^{-1}$)		1.2×10^{10}	2.5×10^{10}	4.4×10^{10}	5.7×10^{10}	6.4×10^{10}	7.5×10^{10}
HDV flow rate (veh h^{-1})		0	141	304	481	663	810
LDV flow rate (veh h^{-1})		7350	7418	7097	7026	6890	6870
HDV density ($\times 10^{-3}$ m^{-1})		n.a	1.5	3.3	5.4	7.6	9.7
LDV density ($\times 10^{-3}$ m^{-1})		74	66	63	64	62	68
$\sigma_{z,init}$ (m)	Parameterized[#]	2.94	3.52	4.21	5.09	5.96	6.93
	CALINE4	2.81	2.40	2.40	2.40	2.40	2.40
	AERMOD	3.51	4.28	4.30	4.31	4.33	4.34
$C_{calculated}$ (g m^{-3})	Parameterized[$]	0.075	0.054	0.049	0.046	0.043	0.044
	CALINE4	0.078	0.079	0.086	0.097	0.108	0.126
	AERMOD	0.062	0.043	0.046	0.050	0.054	0.062
CO_2 measured[%] (g m^{-3})		0.080	0.052	0.051	0.052	0.054	0.052
$C_{calculated}$ (pt cm^{-3})	Parameterized[$]	5136	7018	10109	10985	10594	10866
	CALINE4	5370	10305	17758	23335	26361	31411
	AERMOD	4284	5645	9478	12113	13289	15494
UFP measured (pt cm^{-3})		5616	7385	11371	11882	13839	13411

[#]: Parameterized $\sigma_{z,init,WAKE}$ is calculated based on Eq. (7)-(9) with vehicle density from field measurements and σ_z that not related to vehicles (1.7m for LSD, 1.8m for DRE).

[$]: Calculated concentrations are calculated by applying parameterized $\sigma_{z,init,WAKE}$ to Eq. (S.1) with vehicle emission data (EF_{fleet}, N_{fleet}) and meteorology data (u, Θ) from field measurements.

[%]: Measured concentrations are background subtracted.

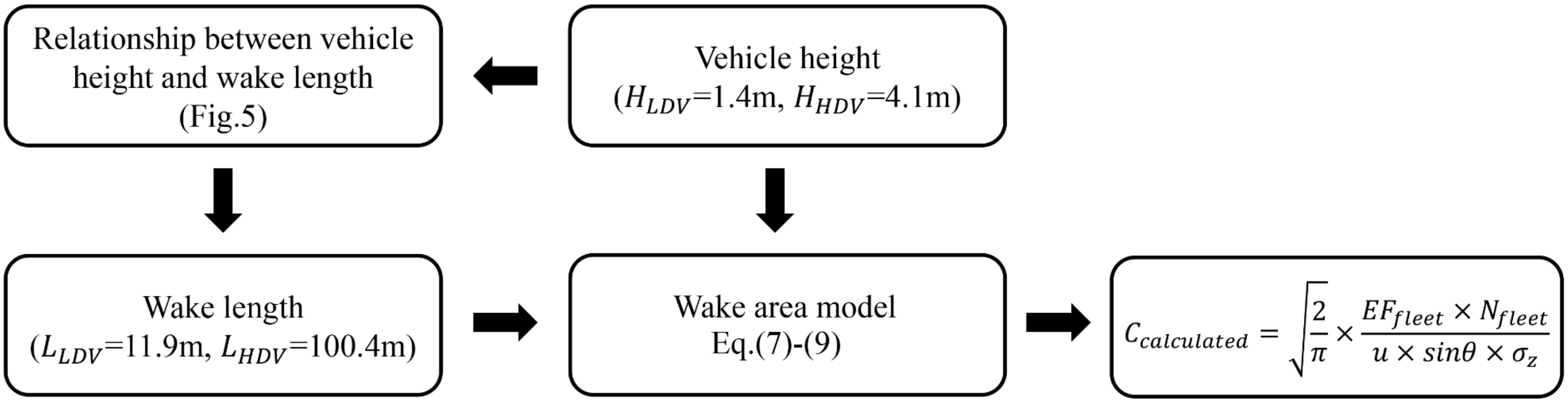

Fig. S2. The flow chart for calculation of pollutant concentrations with parameterized $\sigma_{z,init,WAKE}$.

S4. Model performance measures. Calculated concentrations from the models were compared to field measurements and model performance were quantified using common statistical parameters[4]. These parameters include Pearson correlation (R), normalized mean square error (NMSE), fractional bias (FB), and the fraction of data within a factor of two (FAC2). These parameters are defined using notation of C_o and C_p for observed and predicted concentrations, respectively. For our case, C_o is the field measured concentrations, C_p is the model calculations.

The parameters are defined as following:

$$R = \frac{\overline{(C_o - \overline{C_o})(C_p - \overline{C_P})}}{\sigma c_o \sigma_{cp}}$$ Eq. (S.2)

$$NNSE = \frac{\overline{(C_o - C_p)^2}}{\overline{C_o C_p}}$$ Eq. (S.3)

$$FB = \frac{(\overline{C_o} - \overline{C_p})}{0.5(\overline{C_o} + \overline{C_p})}$$ Eq. (S.4)

By definition, FAC2 is the fraction of model prediction within the range:

$$0.5 \leq \frac{C_p}{C_o} \leq 2$$ Eq. (S.5)

R is a measure of how well the linear relationship between the measure and model results, with a range of -1 to 1. A value of 0 means no correlation, and ±1 means a perfect correlation. NMSE is a measure evaluating the deviation of model results on point to point basis, and 0 value indicates ideal model performance. FB is a measure showing correlation between measure and model results on average basis, with a range of -2 to 2.

Where the negative and positive values indicate over- and underpredict of measurements, respectively. A value of 0 indicates ideal model performance. FAC2 is a measure showing the fraction of data falling into the "factor of two envelope", a value of 1 indicate ideal model performance.

Table S3. Overall performance of calculated pollutant concentrations against pollutant concentrations field measurement that derived from Fig. 7.

Model	R[a]		NMSE[b]		FB[c]		FAC2[d]	
	CO_2	UFP	CO_2	UFP	CO_2	UFP	CO_2	UFP
Wake Area Model	0.92	0.92	0.02	0.11	0.01	0.14	1	1
CALINE4	0.50	0.47	0.21	0.26	-0.23	-0.25	0.80	0.83
AERMOD	0.63	0.82	0.16	0.17	0.13	0.24	0.90	0.93

In Table S3, compared to CALINE4 and AERMOD, the "wake area model" showed larger value of R and lower NMSE. This indicated that there is stronger linear relationship and smaller differences between the calculated concentrations from "wake area model" and the measured concentrations. The "wake are model" showed FB value close to 0 and FAC2 with value of 1. This indicated that the calculated average concentrations from the "wake area model" were closer to the field measurements compared to CALINE4 and AERMOD with no extreme under- and overpredictions.

Reference

1. Arya, P.S., **2001**. Introduction to micrometeorology. Elsevier.
2. Xiang, S.; Yu, Y. T.; Hu, Z.; Noll, K. E., Characterization of dispersion and ultrafine-particle emission factors based on near-roadway monitoring Part II: Heavy duty vehicles. *Aerosol and Air Quality Research* **2019,** *19*, 2421-2431.
3. Xiang, S.; Yu, Y. T.; Hu, Z.; Noll, K. E., Characterization of Dispersion and Ultrafine-particle Emission Factors Based on Near-roadway Monitoring Part I: Light Duty Vehicles. *Aerosol and Air Quality Research* **2019,** *19*, 2410-2420.
4. Chang, J.C. and Hanna, S.R., **2004**. Air quality model performance evaluation. Meteorology and Atmospheric Physics, 87(1-3), pp.167-196.