JOURNAL OF THE AIR & WASTE MANAGEMENT ASSOCIATION
2017, VOL. 67, NO. 7, 763–775
http://dx.doi.org/10.1080/10962247.2017.1287788

TECHNICAL PAPER

MOVES-Matrix and distributed computing for microscale line source dispersion analysis

Haobing Liu, Xiaodan Xu, Michael O. Rodgers, Yanzhi "Ann" Xu, and Randall L. Guensler

School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, USA

ABSTRACT
MOVES and AERMOD are the U.S. Environmental Protection Agency's recommended models for use in project-level transportation conformity and hot-spot analysis. However, the structure and algorithms involved in running MOVES make analyses cumbersome and time-consuming. Likewise, the modeling setup process, including extensive data requirements and required input formats, in AERMOD lead to a high potential for analysis error in dispersion modeling.

This study presents a distributed computing method for line source dispersion modeling that integrates MOVES-Matrix, a high-performance emission modeling tool, with the microscale dispersion models CALINE4 and AERMOD. MOVES-Matrix was prepared by iteratively running MOVES across all possible iterations of vehicle source-type, fuel, operating conditions, and environmental parameters to create a huge multi-dimensional emission rate lookup matrix. AERMOD and CALINE4 are connected with MOVES-Matrix in a distributed computing cluster using a series of Python scripts. This streamlined system built on MOVES-Matrix generates exactly the same emission rates and concentration results as using MOVES with AERMOD and CALINE4, but the approach is more than 200 times faster than using the MOVES graphical user interface. Because AERMOD requires detailed meteorological input, which is difficult to obtain, this study also recommends using CALINE4 as a screening tool for identifying the potential area that may exceed air quality standards before using AERMOD (and identifying areas that are exceedingly unlikely to exceed air quality standards). CALINE4 worst case method yields consistently higher concentration results than AERMOD for all comparisons in this paper, as expected given the nature of the meteorological data employed.

Implications: The paper demonstrates a distributed computing method for line source dispersion modeling that integrates MOVES-Matrix with the CALINE4 and AERMOD. This streamlined system generates exactly the same emission rates and concentration results as traditional way to use MOVES with AERMOD and CALINE4, which are regulatory models approved by the U.S. EPA for conformity analysis, but the approach is more than 200 times faster than implementing the MOVES model. We highlighted the potentially significant benefit of using CALINE4 as screening tool for identifying potential area that may exceeds air quality standards before using AERMOD, which requires much more meteorology input than CALINE4.

PAPER HISTORY
Received September 15, 2016
Revised January 4, 2017
Accepted January 24, 2017

Introduction

Transportation conformity is required by the Clean Air Act section 176(c) (42 U.S.C. 7506(c)) to ensure that federal funding and approval are given to highway and transit projects that are consistent with ("conform to") the air quality goals established in each state implementation plan (SIP). Conformity to the purpose of the SIP means that transportation activities will not cause new air quality violations, worsen existing violations, or delay attainment of the National Ambient Air Quality Standards (NAAQS) (EPA, 2016a).

Hot-spot analysis is defined in 40 CFR 93.101 as an estimation and comparison of likely future localized pollutant concentration with current pollutant concentration and NAAQS. Hot-spot analysis requires detailed modeling of the impacts of transportation project emission sources on the surrounding environment using microscale dispersion analysis. The U.S. EPA has published transportation conformity guidance for hot-spot analysis in particulate matter (PM) nonattainment and maintenance area (EPA, 2015a), as well as project-level carbon monoxide (CO) hot-spot analysis guidance (EPA, 2015b). According to the conformity guidance, MOVES (Motor Vehicle Emission Simulator) is designated as the official mobile emission model (EPA, 2015c) for air quality analysis. AERMOD and CAL3QHCR are the recommended air dispersion models (EPA, 2016b). Guidance for California is slightly different, due to its stricter emissions rules, requiring

CONTACT Haobing Liu haobing.liu@gatech.edu School of Civil and Environmental Engineering, Georgia Institute of Technology, 790 Atlantic Drive, Atlanta, GA 30332-0002, USA.

EMFAC as the emission model (Caltrans, 2016a). CALINE4 is only accepted by the U.S. EPA for CO dispersion analysis in California (Caltrans, 2016b), unless separate approval is granted by the U.S. EPA or interagency modeling work group.

For the remainder of the United States, MOVES is the approved regulatory model released by the U.S. EPA for estimating emissions from the vehicle fleet (EPA, 2015c). In MOVES, a "binning" approach is applied where inside the MOVES model vehicle activity is distributed to vehicle-specific power operating mode bin fractions and bin-specific emission rates are assigned to these bins to estimate emissions. MOVES includes an emission database with base emission rates for each pollutant within each operating mode bin, vehicle regulatory class, model year from 1960 and project to 2050, and at each age level. Through internal calculations, emission rates are weighted by operating mode distribution, and adjusted by air conditioning, fuel properties, inspection and maintenance (I/M) program elements, and meteorology factors, and then aggregated to the on-road fleet (using fleet composition and vehicle miles traveled [VMT] data) to obtain fleet emission rates and an emission inventory.

MOVES also allows users to specify local vehicle operation in the form of local driving cycles and operating mode distributions. Because MOVES facilitates emission modeling as a function of speed and acceleration, the model can make full use of a variety of new fleet activity data, such as streaming machine vision data, smartphone location global positioning system (GPS) data, and traffic simulation. Hot-spot analysis and near-road air quality modeling can benefit from the use of more accurate vehicle activity data and the application of high-resolution emission rates for observed on-road vehicle operation conditions. Other refined input is also required for MOVES modeling, including meteorology, calendar year, fuel specifications, I/M program elements, traffic volumes, fleet age distributions and vehicle type distributions.

Because emissions are a complex function of many locally dependent variables (i.e., local fleet, operations, fuel supply, and I/M strategy), and because MOVES integrates a number of aggregation functions for use in emission estimation at state and county levels, the interface is complex and requires numerous inputs to properly characterize any specific emission scenario modeled by a user. Significant labor is required to prepare MOVES input files. In addition, running MOVES is time-consuming, because emission calculations always begin with base emission rates that are internally adjusted by various correction factors such as temperature, humidity, fuel property, and so on. Hence, MOVES is difficult to use for large-scale transportation networks that experience dynamic changes in on-road fleet composition and operating conditions that affect internal model correction factors and predicted on-road emission rates.

Similarly, complex setup procedures for dispersion modeling also have a high potential for introducing analytical error. Vallamsundar and Lin (2012) showcased the use of MOVES and AERMOD for transportation conformity analysis in accordance with EPA's guidance as an example for others to better understand the proposed process of conformity analysis and model setup. Wu and Niemeier (2016) highlighted potential issues associated with excluding zones and improperly setting up receptors.

The dynamics and fluctuation of traffic flow and meteorology suggested a need for a smoother connection between traffic, emission, and dispersion models. Hence a systematic and automatic process for micro-scale line source dispersion modeling for air quality seems a reasonable goal. An automated method should also help to identify modeling uncertainties that arise from the dynamic nature of traffic and the near-roadway atmospheric environment. Currently, both AERMOD and CALINE4 support only one single meteorology and traffic scenario at a time, and one pollutant type in each run, limiting the efficacy of large-scale modeling using these tools.

Motivated by these challenges, the research team developed an approach that employs MOVES-Matrix, a high-performance emission rate lookup system, with two dispersion models. The goals of this development were to increase modeling speed and to ensure that the same emission and concentration results are obtained as when using MOVES2014a, the latest version of MOVES model directly as outlined in the regulatory approach recommended by U.S. EPA. The MOVES-Matrix modeling framework runs more than 200 times more quickly than running MOVES directly because the MOVES2014a emission rates have already been preprocessed into a matrix for use in applications such as simulation modeling, regional emissions modeling, and now dispersion modeling.

In this paper, the research team reports on the result of implementing MOVES-Matrix directly with CALINE4 and AERMOD, on a distributed computing platform, using Python scripts to automate the connections between the dynamic traffic and meteorological data, MOVES2014a emission rates embedded in MOVES-Matrix, and the two dispersion models. The process obtains the same emission rates and concentrations for a corridor, in a fraction of the time it currently takes to perform such analyses using traditional methods.

Methodology

MOVES-Matrix

This study utilized MOVES-Matrix to generate vehicle emission rates for use in subsequent line source modeling. MOVES-Matrix is essentially a multidimensional array containing emission rate outputs from a huge number of MOVES model runs. The matrix for Atlanta was generated from 146,853 individual MOVES model runs (Liu et al., 2016). The basic process is to run MOVES2014a across all variables that affect output emission rates, where each iteration yields a pollutant emission rate for a uniform vehicle source type (all vehicles represented in the run are a specific type of vehicle), a uniform model year (age group), a specific vehicle fuel type (gasoline, diesel, compressed natural gas, etc.), a specific on-road operating condition (average speed and road type, or a single on-road vehicle specific power [VSP] operating mode bin), a single calendar year, other applicable regional regulatory parameters (fuel properties, I/M program characteristics), and a specific temperature and humidity condition. After conducting the MOVES runs, the resulting MOVES emission rate matrix (MOVES-Matrix) can be queried to obtain the exact same emission rates that are obtained from any individual MOVES model run, without ever having to launch MOVES again, or transfer MOVES outputs into the analyses.

To develop the MOVES-Matrix emission rate database for each region of interest, in total 146,853 MOVES runs were prepared, including the combination of 21 calendar years (2010–2025, 2030, 2035, 2040, 2045, 2050), three fuel supply scenarios (winter, summer, and transition fuels), 111 temperature bins (0°F–110°F temperatures in 1°F intervals), and 21 humidity bins (0–100% relative humidity in 5% intervals), which yields $21 \times 3 \times 111 \times 21 = 146{,}853$ iterations. MOVES is configured in a distributed computer cluster, and it takes 18 days to finish all these iterative runs. The MOVES-Matrix emission rates can then be queried from any analytical platform without ever having to launch MOVES or transfer MOVES modeling output files into the analyses.

MOVES-Matrix emission rates were grouped into individual submatrices, with each submatrix storing emission rates for all source types, all source model years, all on-road operations (speed bins or operating mode bins), for one specific calendar year, one month, one temperature, one relative humidity, one fuel supply (by year and month), and one I/M strategy (by year). In this way, a small subset of emission rates can be extracted from the matrix based on the user's year, month, and meteorology inputs. This structure helps support emission control strategy analysis, given that users tended to assume a single temperature, humidity, and fuel, when exploring the impacts of strategies on traffic activity and emissions. Using a submatrix is significantly faster than extracting data from the full 90-billion-cell emission rate matrix.

MOVES starts with a set of baseline emission rates, and these baseline emission rates are adjusted during each run before they are connected to activity data. MOVES-Matrix stores these adjusted emission rates for all scenarios, including the scenario of interest that can be extracted for use in subsequent analyses. As there are no code modifications, no correction factors, and no approximations involved, the outputs from MOVES-Matrix are exactly the same as obtained from running the MOVES model directly. However, since the results are preprocessed, extracting emissions factors for a particular scenario from MOVES-Matrix is more than 200 times faster than running MOVES directly. Python, Java, Perl, or any other similar scripting program can be used to link MOVES-Matrix emission rates with travel demand models, traffic simulation, monitored data, and dispersion models. MOVES-Matrix is open source and collaborative. More information on setup, implementation, and application of MOVES-Matrix can be found in Liu et al. (2016).

MOVES-Matrix and dispersion model connection

The mechanism for connecting MOVES-Matrix with the dispersion models is shown in Figure 1. Input data can be divided into dynamic and static input. Hourly dynamic input data (e.g., traffic volume, on-road operating speeds, meteorology) are stored in a traffic activity and meteorology database. Static input parameters include link geometry, geographic data, and receptor coordinates and normally do not change within any single analysis.

At the beginning of each model run, the system extracts a submatrix containing emission rate and energy consumption rates applicable to the scenario of interest. This extraction from MOVES-Matrix is based on the calendar year and month of the analysis, and the temperature and humidity range of the analysis (multiple matrices are extracted as temperature and humidity conditions change over the course of an analysis).

Emissions processing is the same as used by MOVES in project-level modeling. Corresponding traffic volumes, fleet composition, and on-road operating conditions are extracted from the fleet database for each link and are used to weight the individual emission rates generated the composite emission rate. This weighting combines on-road vehicle activity, as defined by source type and model year distribution, and the amount of on-road activity by operating mode bin or speed bin to calculate a composite emission rate for each link. The emission rate weighting function is:

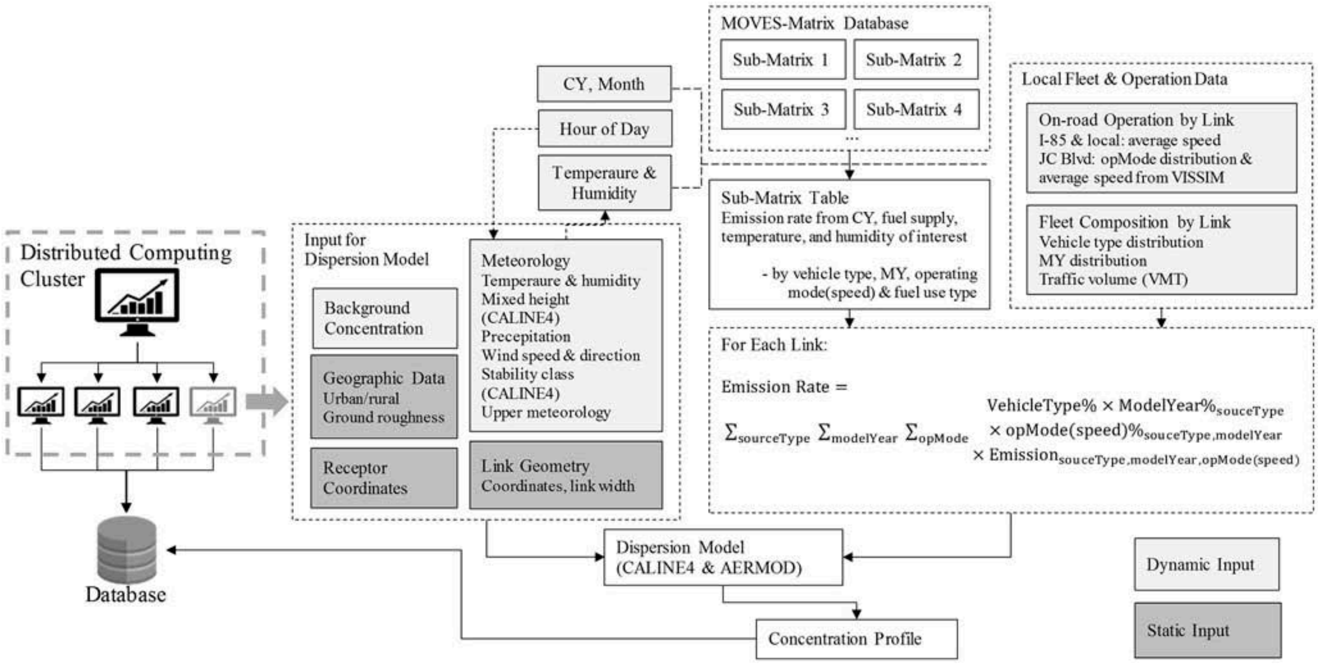

Figure 1. MOVES-Matrix connection with dispersion model for conformity analysis.

$$\text{Emission Rate Bylink} = \sum_{\text{SourceType}} \sum_{\text{modelYear}} \sum_{\text{opMode}} \text{SourceType\%} \times \text{modelYear\%}_{\text{sourcesType}} \times \text{opMode(or SpeedBin)\%}_{\text{sourceType,modelYear}} \times \text{emissionRate}_{\text{sourceType,modelYear,opMode}}$$

$$\text{Link Total Emissions} = \text{VMT} \times \text{EmissionRate} \quad (2)$$

The fleet average emission rate, calculated in the preceding, serves as the emission rate input for CALINE4 and AERMOD use. In this process, the traffic volume, on-road operating conditions, and meteorology data are uploaded for each hour of the year, or a total of 8,760 groups of input combinations (24 hours × 365 days). All of the model connection processes are automated through Python scripts. The input formats required by CALINE4 and AERMOD are stored as templates, with static input information stored in advance. For each iterative dispersion model run, the templates are used to feed dynamic input data, including meteorology, and emissions data from MOVES-Matrix. This automated updating process based on verified templates significantly reduces the chance of processing error.

Using a distributed computing cluster in dispersion modeling

In total, 8,760 hour scenarios (24 hours × 365 days) need to be modeled for a site, which takes a very long time on a typical desktop computer. For this study, the research team obtained priority access to the Partnership for an Advanced Computing Environment (PACE) high-performance computing (HPC) cluster at Georgia Tech. Similar to other distributed computing clusters, PACE was established for the primary purpose of providing an environment for distributed high-performance computing (PACE, 2016). Participating researchers can benefit from the large-scale computing and storage infrastructure, which is organized in the forms of shared queues and distributed computational runs. PACE also provides a variety of software options, and supports almost all the programming languages that are normally used.

The research team takes advantage of the PACE system by submitting dispersion model job iterations to the PACE cluster. The model result, including emission inventory results from MOVES-Matrix, and concentration results from CALINE4 and AERMOD, are assembled to the specified output directory for post processing. In previous research, parallel computing has been utilized to run MOVES and generate vehicle activity and emission for air quality modeling with SMOKE (Faler et al., 2012), which is a system to support emission controls decision making at regional-level. In preparation of on-road vehicle module at SMOKE, the vehicle activity data is needed from MOVES, including VMT, vehicle type/road class distribution and, average speed distribution. Different from SMOKE that is for regional-level analysis, the scope of this research is in microscale. MOVES-Matrix enables users to input roadway average speed (which is essentially applying default driving cycles embedded in MOVES model), as well as apply second-by-second local vehicle speed-acceleration, which better represent on-road vehicle engine load in response of traffic conditions at freeway or intersections. The local operating input is very important in hot-spot analysis in particulate matter (PM) nonattainment and maintenance area (U.S. EPA, 2015a), as well as project-level carbon monoxide (CO) hot-spot analysis (U.S. EPA, 2015b). While SMOKE-MOVES was developed for MOVES and SMOKE connection (Zubrow and Baek, 2011), there was few such system to connect MOVES output with AERMOD or CALINE4 for transportation conformity and hot-spot analysis at microscale.

MOVES-Matrix implementation in dispersion modeling

This section introduces the procedure of implementing MOVES-Matrix with CALINE4 and AERMOD for dispersion modeling in an Atlanta, GA, case study for calendar year 2011.

Case study

This research team selected a 2.5 × 2.5 km (1.55 × 1.55 mile) area in Gwinnett County in the Atlanta metropolitan area as the case study. This study area includes a 3.5-km (2.2-mile) section of I-85 (two-way, six lanes in each direction, 70 mph speed limit), including two entrance ramps, two exit ramps, a 2.8-km (1.74-mile) connecting two-way arterial segment, and associated local roads. This area was selected because of its high traffic volumes, representative of a typical major arterial interchange in suburban Atlanta. There are also detailed data available for analysis in this area. The roadway and layout and receptor locations for the case study are shown in Figure 2.

Local fleet and operation data

For emissions modeling, local fleet and on-road vehicle operating condition data were prepared, including vehicle class distribution, model year distribution, traffic volumes, and operating speeds. The on-road vehicle class distribution and light-duty vehicle model year distributions in the I-85 corridor are derived from video data collected during three days in January 2011 at the Jimmy Carter Boulevard (JCB) site. General vehicle classifications of passenger cars, light-duty trucks, buses, and heavy-duty trucks were manually counted from 2.5 hr of video observations during an a.m. peak period. On-road license-plate data were also transcribed from the video and matched to the motor vehicle registration database to return vehicle characteristics, including model year, fuel type, and body style. Such information can be used to derive detailed MOVES light-duty vehicle class and model year distributions, when combined with the Highway Performance Monitoring System (HPMS) vehicle class distribution (Federal Highway Administration [FHWA], 2013), U.S. EPA certification data (EPA, 2016c), and the class definitions in the MOVES model (Liu et al., 2015). The national 2011 vehicle subfleet composition was used for heavy-duty vehicles, because the majority of heavy-duty vehicles are not registered in Georgia. Details related to vehicle classification process could be found in Liu et al. (2015). Table 1 summarizes the fleet composition of I-85 highway and local roads used in this analysis.

Traffic volumes and operating speeds on I-85 are obtained from the Georgia Department of Transportation (DOT) NaviGAtor intelligent transportation system (ITS). The system recorded traffic volumes and spot speed data for each lane at 20-sec resolution based on the video detection machine vision system. The speed and volume data used in this analysis are aggregated for each hour over the course of a month (24 bins of 1 hr each). Figure 3 shows the monthly averaged 1-hr traffic volume and operating speed of I-85 in each direction from January to December 2011. The peak period is 7:00–9:00 a.m. for southbound, and 5:00–7:00 p.m. for northbound. The variation in hourly volume and speed is large, which will yield similar variability in emissions and near-road concentrations (although not necessarily proportional, because fleet composition and on-road operating conditions affecting emission rates are also variable).

Table 2 shows the data source of fleet and operation data. Traffic volume and fleet composition on JCB and other local roads were obtained from manual processing of in-field video collection during a p.m. weekday peak period (4:00–6:00 p.m.). Since the hourly volume and speed of JCB and local roads in other time period were unavailable, the peak-hour volume was scaled to obtain hourly data in 24 hr according to the volume proportion on I-85. Through the manual process, five vehicle types were classified: motorcycles, light-duty vehicles, buses, single-unit heavy-duty trucks, and combination heavy-duty trucks. The five classes were then subclassified into the 13 MOVES vehicle source types (e.g., dividing light-duty vehicles into passenger cars, passenger trucks, and light commercial trucks) using county-level fleet distributions. Vehicle operations on JCB were estimated using a VISSIM-based simulation model (Xu et al., 2016), one of the most popular microscale traffic simulation software systems. VISSIM simulates individual vehicle movements and interactions, and is typically used to predict on-road

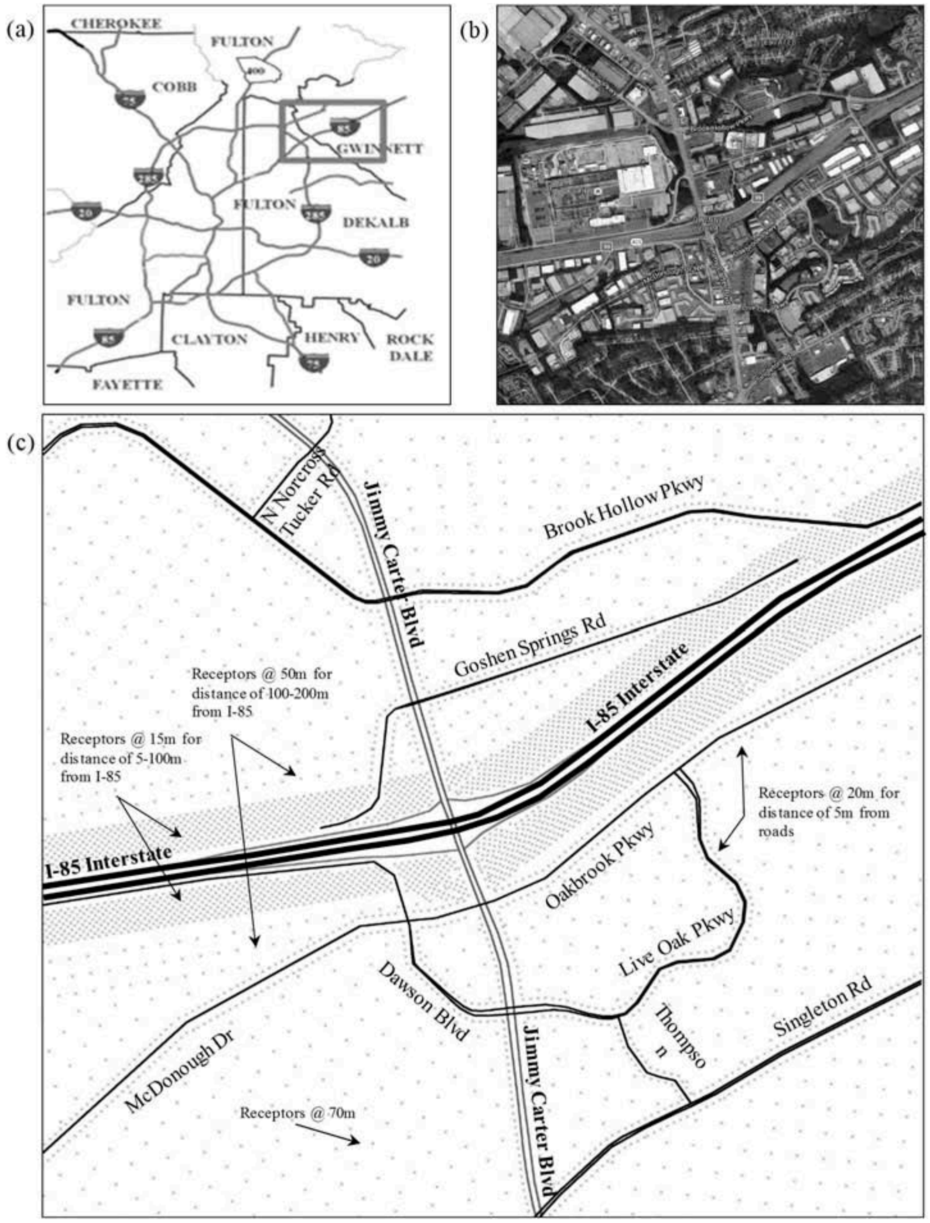

Figure 2. Line source dispersion modeling case study. (a) Case-study map. (b) Case-study satellite. (c) Line source and receptors setup.

vehicle operations. Through the VISSIM model, second-by-second position and speed information of each vehicle was simulated. The VSP of each second were calculated and assigned with MOVES operating mode bin based on vehicle type, instantaneous speed, and acceleration. The speeds of other local roads were assumed to be 25 mph. The vehicle age distribution from the county-level registration database was applied to both JCB and the local roads.

Dispersion model setup

CALINE4 "worst-case" analysis was utilized to simulate the impacts at each receptor for the model-estimated worst-case wind direction, given the model input data for traffic volumes, emission rates, and meteorology. For example, the CO and PM concentration of a point could reach the worst case when the mixing height is extremely low (e.g., 50 m) and the location is immediately downstream from a large emission source. The atmospheric stability class in CALINE4 is set as class E, representing a stable condition. Since CALINE4 worst case always assumes wind direction for a low wind speed and preselected stability, surface roughness, and so on, little meteorological information is required in conducting this process.

AERMOD requires much more refined input for atmospheric conditions than CALINE4, some of which may be unavailable in typical application cases. The AERMOD modeling setup is also more

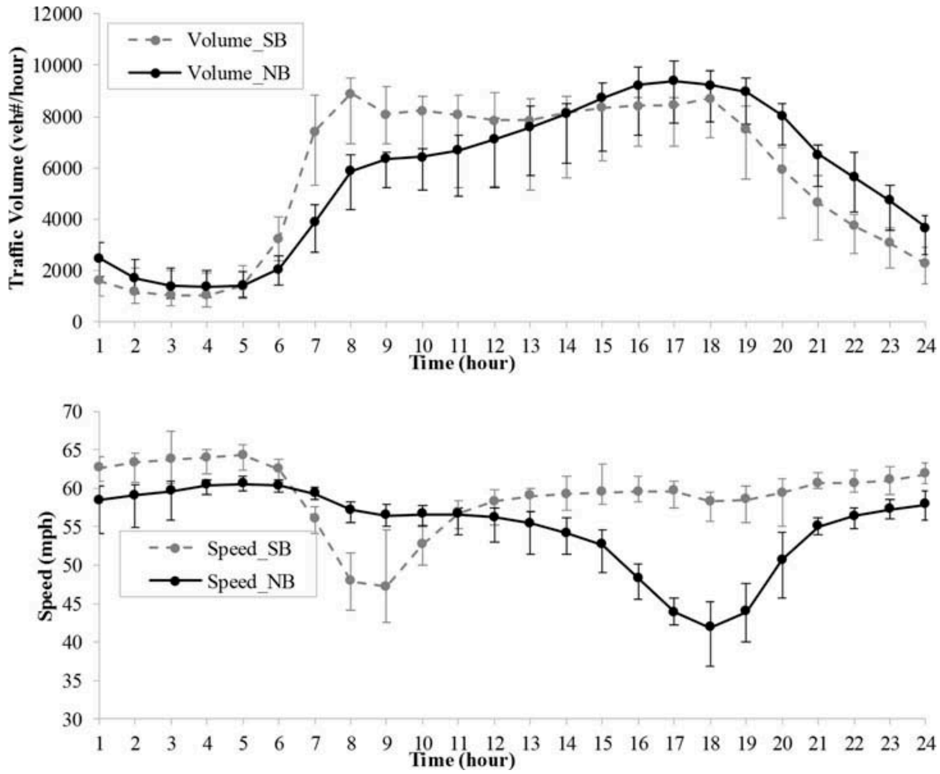

Figure 3. 2011 Monthly average per hour traffic volume and speed on I-85.

Table 1. Fleet composition of I-85 and local roads.

Fleet composition	I-85	Local roads
Light-duty vehicle (%)	93.9	95.7
Bus (%)	0.3	0.4
Heavy-duty truck (%)	5.8	1.9
Light-duty vehicle 0–9 years (%)	72.8	71.0

Table 2. Overview of fleet and operation data source.

Input	Data source	Collection period
Vehicle type distribution	Field video manual count	2.5 hr of video observation 2011, winter: morning peak periods
Age distribution	Field-captured license plates	106,676 plate observations 2011, January: 3 weekdays
Traffic volume and operating speed	NaviGAtor machine vision system: traffic volumes and speed data in 5-min averaging bins	2011, January–December

complicated than for CALINE4, requiring much more effort to develop the model scripting process. Fortunately, the Atlanta Regional Commission provided the team with Atlanta 2011 hourly AERMET (the meteorology data preprocessor for AERMOD) meteorology data from site 33.3825°N, 84.4269°W, along with the resulting wind rose diagram (Figure 4). Since the upper air data in this area are not available in this study, upper air data from Texas are downloaded from Texas Commission of Environmental Quality, to ensure that AERMOD can be launched, which shouldn't significantly affect near-road dispersion results, as the surface concentration is of concern, and the highway-induced concentration significantly decayed with distance. Wind speed, wind direction, wind direction variations, temperature, and humidity from the AERMET file were converted for CALINE4 modeling.

The layout of modeled roadway link sources is shown in Figure 2c. Both AERMOD and CALINE4 used a Cartesian coordination system with a self-defined origin point. In CALINE4, the coordinates of the endpoints for the link centerlines and mixing zone width (link width plus 5 m on each side) were specified to identify the location and geometry of the link segments. The same geometric approximation method is also applied in the AERMOD model with the coordinates of the four ends of the "rectangle" determined separately. The "AREA" method was utilized to simulate the road source in AERMOD (Wu and Niemeier, 2016).

Eight signalized intersections along Jimmy Carter Boulevard were included in the case. Air quality close to intersections is likely to be influenced by vehicle

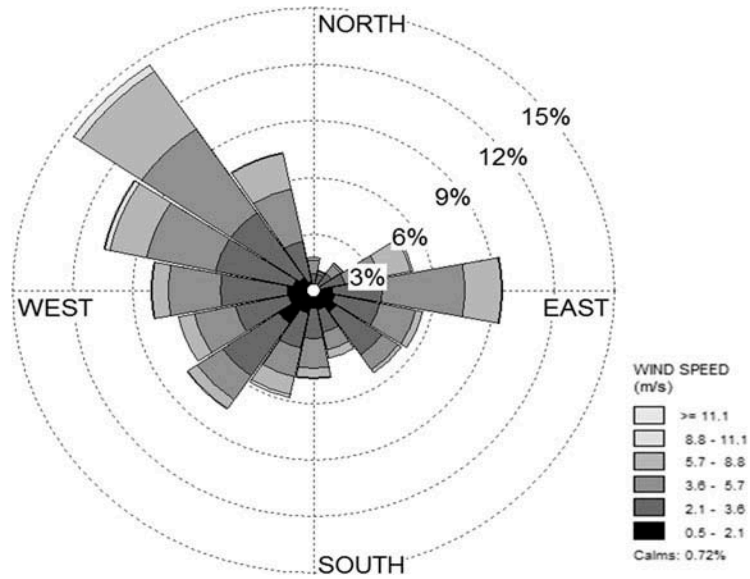

Figure 4. Wind rose diagram in the year 2011.

stop-and-go operations. To refine the influence of the operating characteristics in the intersections, the JCB road source was divided into three types of refined "modal link," based on the hot-spot analysis guide (EPA, 2015a): (1) a "queue link," located upstream of intersection with the length determined by average queue length from VISSIM; (2) an "acceleration link," with speed increasing from 0 mph until $a < 0.1$ m/sec^2; and (3) a "cruising link," links other than "queued" and "acceleration" modes. Second-by-second operating mode bin and speed were assigned into each "modal link" to generate MOVES operating mode distribution and average speed, which can be directly used as input for emission modeling in MOVES or MOVES-Matrix. In CALINE4, considering that only straight lines are used as road sources, curved roads would need to be approximated with multiple straight segments. The road networks with 20 links were then further divided into 228 sublinks specified as rectangles, with longer edges as along the road directions, and shorter edges as road width.

Results from previous research have suggested that insufficient resolution of receptors may yield inaccurate concentration contours (Wu and Niemeier, 2016). From the EPA hot-spot analysis guidance (EPA, 2015a), receptors should be sited as close as 5 m from the road source. Also, the guidance recommends that receptors be placed with a finer spacing (e.g., 25 m) closer to the source, and with wider spacing (e.g., 100 m) farther away from the source. Because the wind direction will influence where maximum impacts are likely to occur, and is likely to change, receptors need also be placed in all directions surrounding the source. In this study, high-resolution receptors were placed at an interval of 15 m from 5 to 100 m from I-85, every 50 m from 100 to 200 m from I-85, and every 70 m farther away. Receptors were also set for a distance of 5 m from other roads at an interval of 20 m. The receptors setup can be seen in Figure 2c.

Table 3 summarizes the dispersion modeling tasks in this study. CO 1-hr concentration, PM_{10} 24-hr concentration, and $PM_{2.5}$ 24-hr and annual concentration are estimated. An emission inventory for each hour is calculated in MOVES-Matrix using hourly traffic data and meteorology data. The research team modified the array definitions in the original CALINE4 source code (which limited modeling to 20 receptors and 20 links per run) to allow 500 receptors and 250 links per run. To run our case study with 4,974 receptors and 228 links, 10 CALINE4 runs were needed for each scenario

Table 3. Dispersion modeling scope in the year of 2011.

Pollutant	Scope	NAAQS	MOVES-Matrix runs	CALINE4 runs	AERMOD runs
CO	1-hr	35 ppm, not to be exceeded more than once per year	8,760 runs	87,600 runs	8,760 runs
	8-hr	9 ppm, not to be exceeded more than once per year			
$PM_{2.5}$	24-hr	35 μg/m^3, 98th percentile, averaged over 3 years	8,760 runs	87,600 runs	365 runs
	Annual	12 μg/m^3			
PM_{10}	24-hr	150 μg/m^3, not to be exceeded more than once	8,760 runs	87,600 runs	365 runs

hour. Thus, 87,600 total CALINE4 runs were required to complete all hourly runs for all receptors through the year 2011. For the CALINE4 model, after the runs are finished, the hourly concentrations values can be easily aggregated and averaged in 24-hr and annual periods to obtain the concentration metric of interest. In AERMOD modeling, the actual meteorology file is imported into the analysis, and hourly emission rates are imported through the AERMOD "HOUREMIS" module. The AERMOD model itself can aggregate the results to averaged concentration index of interest, which corresponds to NAAQS metrics.

According to the NAAQS, the CO 1-hr concentration limit is 35 ppm, the 8-hr limit is 9 ppm, and the PM_{10} 24-hr limit is 150 μg/m^3. These standards are not allowed to be exceeded more than once per year. In addition, the $PM_{2.5}$ 24-hr limit is 35 μg/m^3, for which the 98th percentile of the concentration is not allowed to exceed over 3 years. The $PM_{2.5}$ annual average limit is 12 μg/m^3. The modeled results from AERMOD and CALINE4 are compared with the NAAQS.

Results

Operating performance

A model performance comparison for result outputs and run time was conducted between using MOVES GUI and CALINE4/AERMOD (MOVES GUI method), and distributed computation of MOVES-Matrix and CALINE4/AERMOD in PACE (MOVES-Matrix method). The results from these two methods are exactly the same, with R^2 equal to 1 for the x–y scatterplot. Only the intercept is as small as 10^{-11} due to the rounding errors from calculation. Figure 5 shows the example of CO emissions for 228 links, and receptor concentrations modeled for August 19, 2011, at 8:00–9:00 a.m. Given that the two approaches achieve the same exact emissions and concentration results it could be argued that the MOVES-Matrix method should be considered able to meet EPA's standards for use in regulatory analysis.

Model run time was also compared between using the MOVES GUI method and the MOVES-Matrix method (Table 4). The run time for the using MOVES GUI is based on an average speed of 24 runs in the computer with configuration of Intel Xeon CPU W3550 at 3.07 GHz, Windows 7 64-bit, RAM 6 GB. The running time based on MOVES-Matrix and PACE was recorded. Using the MOVES GUI, the emission inventory for 8,760 hr across 228 links takes 193 hr to complete, plus additional time to run the dispersion models. The MOVES-Matrix approach operating on the PACE computing cluster system completed all of emissions calculations and all of the dispersion modeling runs within 3 hr. In general, the dispersion modeling based on MOVES-Matrix and distributed computing system is more than 200 times faster than the normal procedures based on MOVES GUI, CALINE4, and AERMOD. The fast calculation speed of MOVES-Matrix can provide a platform that can be employed with newer and bigger data sets (e.g., INRIX GPS data, traffic simulations, smartphone data, etc.) and supports dynamic, real-time emission modeling and dispersion modeling, as well as short-term air quality warnings if unusual weather and/or heavy traffic congestion is predicted. The MOVES-Matrix application can also be used to enhance dispersion modeling sensitivity analysis in response to surrounding environment and traffic conditions.

Concentration results

The concentration results of all receptors, including highest 1-hr CO, highest 8-hr CO, 98th percentile 24-hr $PM_{2.5}$, annual $PM_{2.5}$, and highest PM_{10} concentration from AERMOD and CALINE4 worst case, were also compared and are shown in Figure 6. The comparison is conducted not for evaluation of the prediction accuracy, but to show the relationships of the results between two models: For all indexes and pollutant types modeled, CALINE4 worst-case results are consistently higher than AERMOD results, by a ratio of 1.3 to 4.0. This is not surprising because the CALINE modeling employed the worst-case wind angle approach (to identify the maximum concentration that will result at a receptor if the wind comes from the worst-case direction) and AERMOD is run with actual meteorology input for the local area of interest (wind speed, direction, and changes in both are accounted for over time, but it is not a worst-case analysis). Some other field studies have shown that AERMOD tends to underestimate PM concentrations (Chen et al., 2008; Kesarkar et al., 2007; Zhang et al., 2008). The MOVES-Matrix emission modeling approach will facilitate direct comparisons of numerous dispersion models, which should help researchers identify the conditions under which the dispersion model predictions diverge from each other (perhaps helping researchers to identify problems with internal dispersion algorithms).

It is interesting to note that the CO, PM_{10} and $PM_{2.5}$ 24-hr results from the CALINE4 worst case method all meet the NAAQS (see Figure 6). The comprehensive CALINE4 run results indicate that researchers can safely conclude that the project will not violate ambient air quality standards, and that it is not necessary to run the more complex AERMOD approach with respect to project-level CO (will not cause or contribute to a NAAQS violation). However, for the $PM_{2.5}$ annual

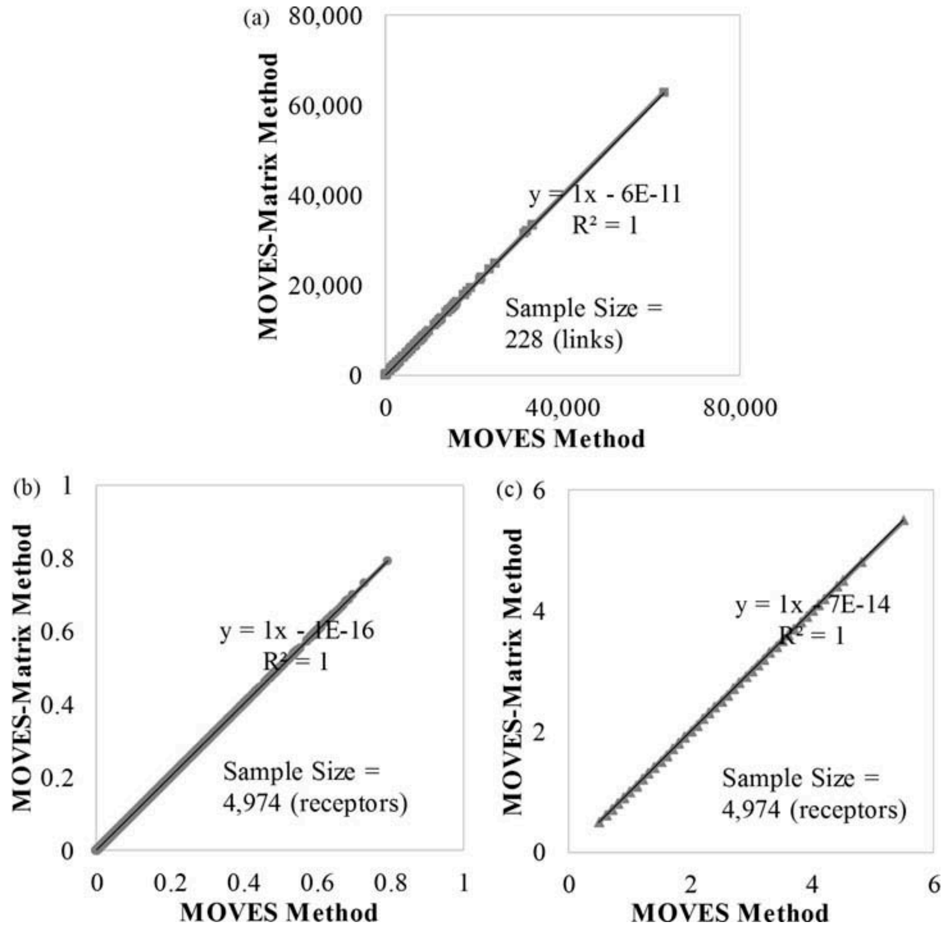

Figure 5. CO emissions and concentrations results by MOVES GUI and Matrix. (a) CO emission results (g/hr in 1-mile segment). (b) CO concentration from AERMOD (ppm). (c) CO concentration from CALINE4 (ppm).

Table 4. Model run time comparison.

Modeling scenario (over the year of 2011)	MOVES GUI method			
	Emissions modeling	Dispersion modeling	Total	MOVES-Matrix method, total
CALINE4				
CO 1-hr and 8-hr	193 hr	49 hr	242 hr	2 hr
$PM_{2.5}$ 24-hr and annual	193 hr	49 hr	242 hr	2 hr
PM_{10} 24-hr	193 hr	49 hr	242 hr	2 hr
AERMOD				
CO 1-hr and 8-hr	193 hr	438 hr	631 hr	2.9 hr
$PM_{2.5}$ 24-hr and annual	193 hr	30 hr	223 hr	2 hr
PM_{10} 24-hr	193 hr	24 hr	217 hr	2 hr

average concentration, CALINE4 results show that the area within around 35–50 m of the I-85 highway could possibly exceeded the NAAQS under persistent worst-case conditions (Figure 7b), whereas the AERMOD results indicated all the area meet the standard. Unlike the 1-hr analyses, where one might expect that worst-case conditions will be encountered at some point during the year, analyses that aggregate continuous worst-case conditions over a whole year are representing conditions that will simply never occur. Given these findings, it seems worthwhile to continue undertaking these kinds of detailed sensitivity and comparative analyses for the purposes of developing a set of rules that will allow CALINE4 to serve as an approvable screening tool.

Conclusion

This study introduced the application of MOVES-Matrix, a high-performance emission modeling system, in

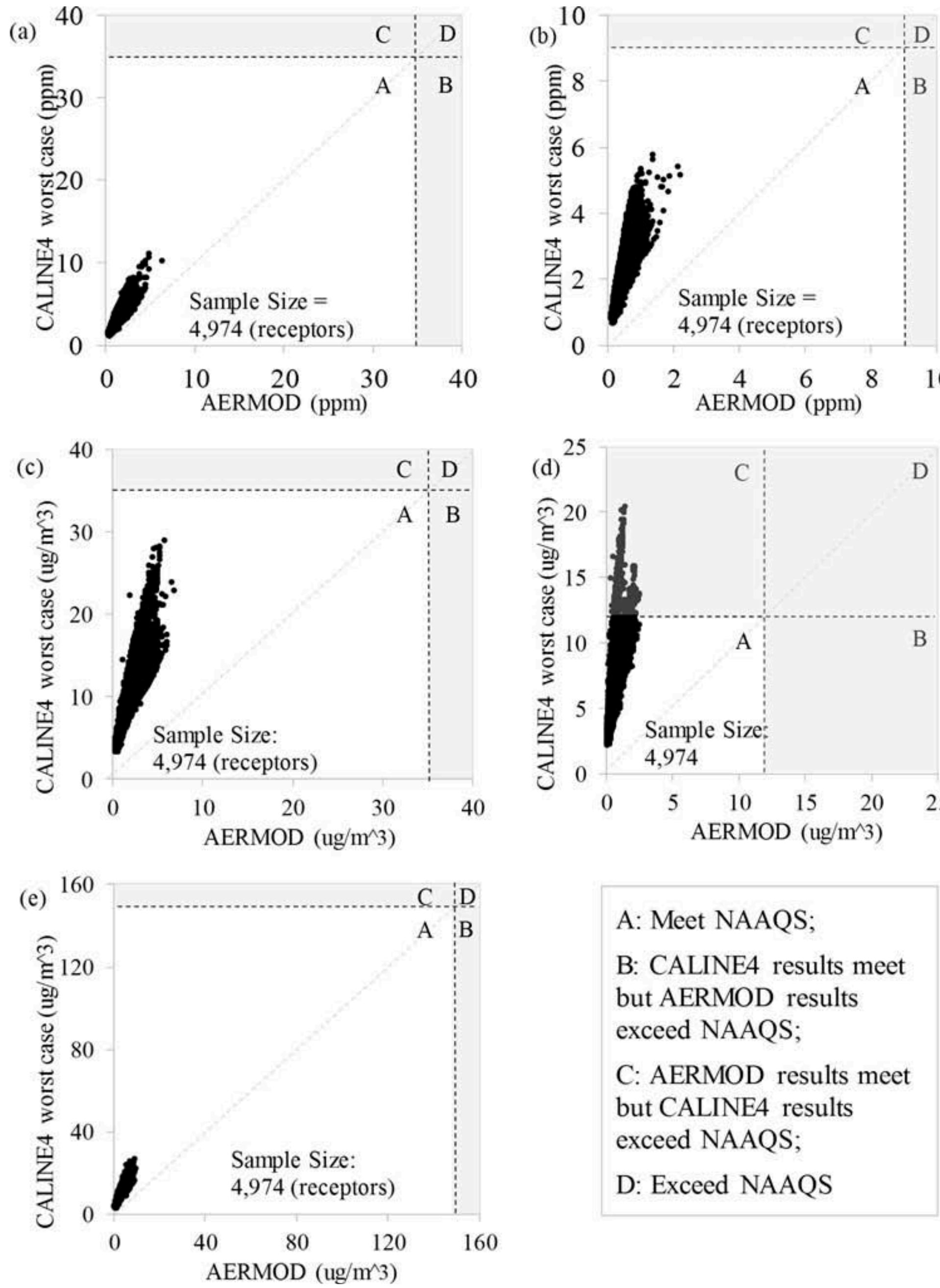

Figure 6. Modeling concentration from CALINE4 and AERMOD. (a) CO Highest hourly concentration. (b) CO highest 8-hr concentration. (c) $PM_{2.5}$ 98th percentile 24-hr concentration. (d) $PM_{2.5}$ annual concentration. (e) PM_{10} Highest 24-hr concentration.

dispersion modeling. The MOVES-Matrix connections with AERMOD and CALINE4 models are automated through Python scripts. Hour-by-hour traffic operation data and meteorological data are assigned to the system for distributed computing. Because the MOVES2014a emission rates outputs are contained in MOVES-Matrix, and no approximations or corrections are employed, the emission results from the MOVES-Matrix are exactly the same as using MOVES GUI. This means that the MOVES-Matrix model obtains the same results as the standard regulatory dispersion analysis. The research team believes that the MOVES-Matrix modeling approach is ready for regulatory review and approval for dispersion modeling.

MOVES-Matrix stores adjusted emission rates as a function of temperature, humidity, I/M strategy, fuel, and so on. Unlike the MOVES model that starts its calculation from the base emission rate for each run, there are no code modifications or use of correction factors for the emission rates in MOVES-Matrix, and only filtering and query are needed to obtain the

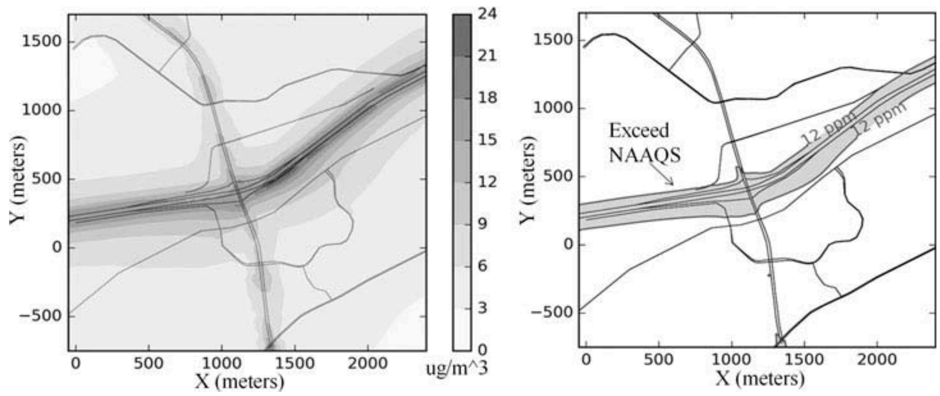

Figure 7. $PM_{2.5}$ annual concentration and areas exceeded NAAQS from CALINE4: (a) Annual concentration. (b) Area exceeded NAAQS.

adjusted emission rates. MOVES-Matrix can finish the emissions computation tasks 200 times more rapidly than using the MOVES GUI. Rapid calculation speeds on the distributed computing cluster allows MOVES-Matrix to couple with various sources of big data for vehicle activity. Hot-spot analysis and near-road air quality modeling benefit from the use of more accurate vehicle fleet specification and on-road vehicle operating condition data (high-resolution emission rates for on-road driving conditions). MOVES-Matrix also has the potential for use in real-time emission and dispersion modeling, as well for use in as short-term air quality warning if unusual weather and/or heavy traffic congestion is predicted. MOVES-Matrix is an open source system that anyone can use. It can be operationalized in Java, Python, Perl, or any similar scripting program to link MOVES emission rates with other software such as dispersion models. This automated linkage will smooth the connection between emission model and dispersion model, and thus minimize data-processing errors.

Through the results comparison between using AERMOD and using CALINE4 worst case, the research team also found that the modeled concentration from CALINE4 worst case are consistently higher than for AERMOD for all pollutant indexes evaluated. Compared with AERMOD, which requires refined meteorological input, one of the promising advantages of using CALINE4 worst case is it requires very little meteorology input data, which saves a lot of data collection work. The research team believes that before venturing into a refined and complicated AERMOD analysis, the user-friendly CALINE4 line source model with worst-case method could be used as a screening tool to identify the potentially high concentration areas and areas where violations are not likely to occur. If NAAQS or the specified air quality requirement is met even with CALINE4 worst case, it is very likely that the same conclusion is get by using AERMOD, since it shows consistently lower concentration results. More detailed research is needed to verify this conclusion and to explore the specific conditions that either CALINE4 worst case can or cannot perform well as this a screening tool.

Funding

This work was funded by the National Center for Sustainable Transportation.

About the authors

Haobing Liu is a graduate research assistant at the School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA.

Xiaodan Xu is a graduate research assistant at the School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA.

Michael O. Rodgers, Ph.D., is a principal research scientist and adjunct professor at the School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA.

Yanzhi “Ann” Xu, Ph.D., is as research engineer II at the School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA.

Randall L. Guensler, Ph.D., is a professor at the School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA.

References

California Department of Transportation. 2016a. EMFAC. http://www.dot.ca.gov/hq/env/air/pages/emfac.htm (accessed July 31, 2016).

California Department of Transportation. 2016b. CALINE4 software. http://www.dot.ca.gov/hq/env/air/software/caline4/calinesw.htm (accessed July 31, 2016).

Chen, H., S. Bai, D. Eisinger, D. Niemeier, and M. Claggett. 2008. Modeling uncertainties and near-road PM2.5: A comparison of CALINE4, CAL3QHC and AERMOD. University of California–Davis (UCD) Caltrans Air Quality Project. University of California–Davis, Davis, CA.

Faler, W., H. Michaels, and W. Aikman. 2012. Using cloud computing to do large numbers of MOVES runs. 20th International Emission Inventory Conference. Tampa, FL, August 13–16. https://www3.epa.gov/ttn/chief/conference/ei20/session8/wfaler.pdf (accessed November 2016).

Federal Highway Administration. 2013. *Highway performance monitoring system: Field manual*. http://www.fhwa.dot.gov/ohim/hpmsmanl/chapt3.cfm (accessed July 31, 2016).

Kesarkar, A.P., M. Dalvi, A. Kaginalkar, and A. Ojha. 2007. Coupling of the weather research and forecasting model with AERMOD for pollutant dispersion modeling. A case study for PM_{10} dispersion over Pune, India. *Atmos. Environ.* 41(9):1976–88. doi:10.1016/j.atmosenv.2006.10.042

Liu, H., Y. Xu, M. Rodgers, and R. Guensler. 2015. Developing vehicle classification inputs for project-level MOVES analysis. *Transportation Research Record* 2503:81–90. doi:10.3141/2503-09

Liu, H., Y. Xu, M. Rodgers, A. Akanser, and R. Guensler. 2016. Improved energy and emissions modeling for project evaluation (MOVES-Matrix). National Center for Sustainable Transportation. http://ncst.ucdavis.edu/wp-content/uploads/2014/09/10-10-2016-NCST_MOVESMatrix_FinalReport_version-4.pdf (accessed November 31, 2016).

PACE. 2016. Partnership for an Advanced Computing Environment. http://www.pace.gatech.edu (accessed July 31, 2016)

U.S. Environmental Protection Agency. 2015a. Transportation conformity guidance for quantitative hot-spot analyses in $PM_{2.5}$ and PM_{10} nonattainment and maintenance areas. EPA Report EPA-420-B-15-084. https://www.epa.gov/state-and-local-transportation/project-level-conformity-and-hot-spot-analyses#pmguidance (accessed March 6, 2017).

U.S. Environmental Protection Agency. 2015b. Using MOVES2014 in project-level carbon monoxide analyses. EPA Report EPA-420-B-15-028. https://www.epa.gov/state-and-local-transportation/project-level-conformity-and-hot-spot-analyses#coguidance (accessed March 6, 2017).

U.S. Environmental Protection Agency. 2015c. MOVES2014a. https://www3.epa.gov/otaq/models/moves (accessed July 31, 2016).

U.S. Environmental Protection Agency. 2016a. National ambient air quality standards (NAAQS). https://www.epa.gov/criteria-air-pollutants/naaqs-table (accessed July 31, 2016).

U.S. Environmental Protection Agency. 2016b. Preferred/recommended models. https://www3.epa.gov/scram001/dispersion_prefrec.htm (accessed July 31, 2016).

U.S. Environmental Protection Agency. 2016c. Light-duty vehicle and engine emission certification. http://www.epa.gov/otaq/cert.htm (accessed July 31, 2016).

Vallamsundar, S., and J. Lin. 2012. MOVES and AERMOD used for PM2.5 conformity hot spot air quality modeling. *Transportation Research Record* 2270:39–48. doi:10.3141/2270-06

Wu, Y., and D. Niemeier. 2016. Strategy of AERMOD configuration for transportation conformity hot-spot analysis. Paper presented at the 95th Transportation Research Board Annual Meeting, Transportation Research Board of the National Academies, Washington, DC, January 10–14.

Xu, X. H. Liu, Y. Xu, M. Rodgers, and R. Guensler. 2016. Estimating project-level vehicle emissions with VISSIM and MOVES-Matrix. *Transportation Research Record* 2570:107–17. doi:10.3141/2570-12

Zhang, Q., Y. Wei, W. Tian, and K. Yang. 2008. GIS-based emission inventories of urban scale: A case study of Hangzhou, China. *Atmos. Environ.* 42(20):5150–65. doi:10.1016/j.atmosenv.2008.02.012

Zubrow, A., and B.H. Baek. 2011. SMOKE-MOVES. U.S. Environmental Protection Agency air quality seminar. https://www.epa.gov/moves/moves-workshops-and-presentations (accessed November 2016)).