MOVES-Matrix 3.0: On-Road Energy and Emission Modeling with High-Performance Supercomputing

October 2024

A research report from the National Center for Sustainable Transportation

Hongyu Lu, Georgia Institute of Technology
Michael O. Rodgers, Ph.D., Georgia Institute of Technology
Randall Guensler, Ph.D., Georgia Institute of Technology

National Center for Sustainable Transportation
A USDOT University Transportation Center

Georgia Tech School of Civil and Environmental Engineering
College of Engineering

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NCST-GT-RR-24-29	N/A	N/A

4. Title and Subtitle
MOVES-Matrix 3.0: On-Road Energy and Emission Modeling with High-Performance Supercomputing

5. Report Date
October 2024

6. Performing Organization Code
N/A

7. Author(s)
Hongyu Lu, https://orcid.org/0000-0002-0170-7169
Michael O. Rodgers, Ph.D., https://orcid.org/0000-0001-6608-9333
Randall Guensler, Ph.D., https://orcid.org/0000-0003-2204-7427

8. Performing Organization Report No.
N/A

9. Performing Organization Name and Address
School of Civil and Environmental Engineering
Georgia Institute of Technology
790 Atlantic Drive, Atlanta, GA 30332-0355

10. Work Unit No.
N/A

11. Contract or Grant No.
USDOT Grant 69A3551747114

12. Sponsoring Agency Name and Address
U.S. Department of Transportation
Office of the Assistant Secretary for Research and Technology
1200 New Jersey Avenue, SE, Washington, DC 20590

13. Type of Report and Period Covered
Final Research Report
(March 2021 – October 2023)

14. Sponsoring Agency Code
USDOT OST-R

15. Supplementary Notes
DOI: https://doi.org/10.7922/G2H130CC
Dataset DOI: https://doi.org/10.5281/zenodo.13961059

16. Abstract
The Georgia Tech research team developed MOVES-Matrix 3.0 based on the EPA's MOVES3 (version 3.1.0) energy use and emission rate model by running MOVES3 thousands of times on the PACE supercomputing cluster across all combinations of input variables and storing the output as lookup tables. MOVES-Matrix 3.0 allows on-road energy consumption and emissions modeling to be conducted more than 800 times faster than running MOVES, while it generates the exact same results, as verified in this report. MOVES-Matrix 3.0 was designed similarly to its predecessor, MOVES-Matrix 2014, but required extensive code modifications to accommodate changes in the MOVES3 environment (including a shift from MySQL to MariaDB and incorporation of new vehicle source sub-types and operating parameters). The review of the fuel and I/M scenarios indicated that MOVES3 now defines 122 modeling regions, as compared with 109 regions in MOVES 2014b (different matrices need to be developed each modeling region). The development of matrices for each modeling region takes approximately 15-20 days on the PACE supercomputing cluster given our assigned resources (compared with only 5-7 days to develop matrices for MOVES 2014). A case study of 3,000 roadway links using Atlanta's matrices confirmed that MOVES-Matrix 3.0 produces the exact same energy consumption and emissions results as MOVES3, but execution modules operate 800 times faster using MOVES-Matrix lookups than running MOVES for any single run.

17. Key Words
Emission Rate Modeling; Energy Use Modeling; MOVES; MOVES-Matrix

18. Distribution Statement
No restrictions.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	27	N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

About the National Center for Sustainable Transportation

The National Center for Sustainable Transportation is a consortium of leading universities committed to advancing an environmentally sustainable transportation system through cutting-edge research, direct policy engagement, and education of our future leaders. Consortium members include: the University of California, Davis; California State University, Long Beach; Georgia Institute of Technology; Texas Southern University; the University of California, Riverside; the University of Southern California; and the University of Vermont. More information can be found at: ncst.ucdavis.edu.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

The U.S. Department of Transportation requires that all University Transportation Center reports be published publicly. To fulfill this requirement, the National Center for Sustainable Transportation publishes reports on the University of California open access publication repository, eScholarship. The authors may copyright any books, publications, or other copyrightable materials developed in the course of, or under, or as a result of the funding grant; however, the U.S. Department of Transportation reserves a royalty-free, nonexclusive and irrevocable license to reproduce, publish, or otherwise use and to authorize others to use the work for government purposes.

Acknowledgments

This study was funded, partially or entirely, by a grant from the National Center for Sustainable Transportation (NCST), supported by the U.S. Department of Transportation (USDOT) through the University Transportation Centers program. The authors would like to thank the NCST and the USDOT for their support of university-based research in transportation, and especially for the funding provided in support of this project. The authors would also like to thank staff from the City of Atlanta and Atlanta Regional commission for providing network data and assistance.

NCST

MOVES-Matrix 3.0: On-Road Energy and Emission Modeling with High-Performance Supercomputing

A National Center for Sustainable Transportation Research Report

October 2024

Hongyu Lu, Graduate Research Assistant

Michael O. Rodgers, Ph.D., Regents Researcher and Adjunct Professor

Randall Guensler, Ph.D., Professor

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

[page intentionally left blank]

NCST

TABLE OF CONTENTS

List of Tables

List of Figures

MOVES-Matrix 3.0: On-Road Energy and Emission Modeling with High-Performance Supercomputing

EXECUTIVE SUMMARY

This report summarizes the development of MOVES-Matrix 3.0, an energy and emission rate lookup system, based upon EPA's latest energy use and emission rate model known as MOVES3 (version 3.1.0). The MOVES-Matrix system developed by the Georgia Tech team can be used to conduct on-road energy consumption and emissions modeling more than 800-times faster than using MOVES directly. The concept of using iterative model processing and matrix generation was first applied with the MOBILE model nearly 30 years ago (Guensler et al. 2004, 2000; Guensler and Leonard II 1995). MOVES-Matrix follows the same iterative approach, where a multidimensional array is generated from MOVES model run outputs. MOVES-Matrix 1.0 for Atlanta, GA was originally created by running the MOVES 2014 model 146,853 times (across all combinations of input variables). The iterative runs populate matrices that can be queried in any energy and emissions analyses using Python® scripts and/or other applications, for any set of fleet compositions, model year distributions, environmental conditions, and on-road operating conditions. The MOVES-matrix lookup system provides the exact same emission results as using MOVES 2014 directly for each analysis of an individual transportation road link, ensuring the validity of using the matrix outputs and facilitating their use in regulatory analysis. The resulting emission rate matrices allow users to connect link-by-link vehicle activity data for large projects with applicable emission rates, and to evaluate changes in emissions for dynamic transportation systems in near-real-time, without having to launch MOVES (because all possible MOVES runs for the modeling region have already been pre-processed).

In developing MOVES-Matrix 3.0, the research team applied the same conceptual design used in MOVES-Matrix 2014; however, all of the programs employed in generating matrices on the distributed computing servers had to be updated to account for changes in the updated MOVES3 environment (e.g., internal use of MariaDB instead of MySQL, changes in internal data structures, and addition of vehicle source sub-types). The development of MOVES-Matrix 3.0 required significant code modifications to: 1) integrate the installation of MariaDB on the supercomputing cluster (without root access), 2) account for new source type physics parameters (vehicle specific power/scaled-tractive power configurations for various source types by model year), 3) address changes in input parameters, and 4) process significantly greater numbers of mandatory iterations within each individual model run.

MOVES3 also employs more unique combinations of fuel supply regions and inspection and maintenance (I/M) programs than the team faced in developing the original MOVES-Matrix. MOVES3 now includes 40 fuel scenarios (vs. 22 fuel scenarios in MOVES 2014b) and 87 I/M scenarios in MOVES3 (vs. 84 I/M scenarios in MOVES 2014b). The increase in fuel scenarios represents changes to fuel formulation regions and one-to-many relationships between counties and fuel formulation regions by year. MOVES-Matrix is processed for each unique combination of fuel scenario and I/M scenario, which MOVES-matrix defines as a modeling

region. The number of MOVES-Matrix modeling regions has increased from 109 regions in MOVES 2014b to 122 regions in MOVES3 (representing a 12% increase in resources required to develop all matrices for the nation).

Each iterative process to develop matrices for a modeling region takes about 15 to 20 days to process on the PACE supercomputing cluster given the considerable resources assigned to the research team by PACE (a Georgia Tech supercomputer Co-op) to the modeling team. The case study presented in this report for Atlanta's matrices demonstrate that MOVES-Matrix 3.0 generates the exact same emission and energy consumption results with MOVES3, while running the MOVES-Matrix lookup module is about 800 times faster than running a MOVES3 scenario directly for analytical work. This provides tremendous advantages in regional modeling assessments across a variety of fleet compositions and on-road operating conditions.

1. Introduction

The MOtor Vehicle Emission Simulator (MOVES) model (U.S. EPA, 2023a) provides significantly improved emission rates compared to the older MOBILE series of models (U.S. EPA, 2016), primarily because MOVES emission rates are more modal in nature, better representing emissions as a function of instantaneous (1Hz) speed and acceleration. Coupling high-resolution on-road vehicle activity data with appropriate MOVES emission rates further advances research efforts designed to assess the environmental impacts of transportation design and operation strategies (U.S. EPA 2023a). Hot spot analysis and near-road dispersion modeling for environmental impact assessment also benefit from using more accurate vehicle activity data in time and space, and the application of high-resolution emission rates to actual on-road driving conditions (Guensler et al. 2021; Liu et al., 2019b). MOVES3 has been updated and improved from the previous version of MOVES 2014 by: 1) incorporating the latest data on vehicle populations, emission rates, and regional fuels; 2) better accounting for engine starts, truck hoteling (extended idling); 3) incorporating the forecasted impacts of new truck and passenger vehicle standards (U.S. EPA, 2022b). U.S. EPA published the latest version of MOVES4 in August 2023 (U.S. EPA, 2023a, 2023b); the research team is currently working on developing MOVES-Matrix 4.0 based on the same MOVES-Matrix 3.0 design.

Because emissions are a complex function of many locally-dependent variables, 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 user inputs to properly characterize any specific emission scenario. Extensive labor is required to prepare MOVES input files. In addition, running MOVES is time consuming, because emission calculations begin with base emission rates, which are internally adjusted by various correction factors such as temperature, humidity, fuel property, etc. This also makes MOVES difficult to use for large-scale transportation networks that experience dynamic changes in on-road fleet composition and operating conditions that affect corrections factors during the day (CRC, 2017).

The latest Atlanta Regional Commission (ARC) Travel Demand Model network now includes 148,000 roadway segment links (nearly doubling the number of links compared to the 2019 model with 74,500 links). It is nearly impossible to perform emissions modeling for a dynamic network of this size using individual MOVES emission rates for each link, when fleet composition and on-road operating conditions change dynamically over the course of a day. On a typical personal computer (PC), depending on the pollutant types to be modeled, MOVES requires around 10-30 seconds to process emissions for one link for a unique fleet and operating condition. To obtain the composite emission rates for 1,000 roadway links in Atlanta, where the fleet composition and operating conditions vary every hour on every road segment, and where temperatures and humidity values vary by hour of day and month, and for the three Atlanta fuels (summer, winter, and transition), would require nearly 32 million individual MOVES runs, which would take ten years to run on a typical PC. This is based on 1,000 road segments, with operations of each hour across 24 hours, in 21 temperature bins scenarios (10-110 °F in 5 °F bins), 21 humidity bins scenarios (0%-100% in 5% bins), and 3 fuels supply scenarios (summer, winter, and transition fuel supply), which yields 31,752,000 individual

MOVES runs. Shortcuts can be taken to reduce the number of MOVES runs required (e.g., many runs yield the exact same emission output across a certain temperature/humidity ranges as they are insensitive to several pollutant types, and there is no need to run the model for low temperatures during summer months). However, modeling across all such operating conditions on PCs is impractical. High-performance modeling approaches are needed to assess large-scale dynamic networks.

Regulations require the latest approved regulatory model be used in all transportation and air quality planning and assessment work (at the time of this research, MOVES3 was the latest; however, MOVES4 is now the most recent model) (U.S. EPA, 2020a, 2023a, 2023b, 2023c). To improve modeling efficiency, but at the same time ensure that regulatory requirements for use of MOVES (2014 at that time) are met, the research team developed MOVES-Matrix (Liu et al. 2019a). The MOVES model was run more than a hundred thousand times to generate a multidimensional energy use and emission rate matrices for all combinations of MOVES input variables (i.e., MOVES runspecs and user csv supplied data, see U.S. EPA 2023).

Modelers can query MOVES-Matrix emission rates generated through the iterative modeling process, and apply these emission rates in any analytical activity that can be performed with MOVES outputs, without ever having to launch MOVES or transfer MOVES modeling output files into the analyses. The scenario runs demonstrate that MOVES-Matrix can finish the emissions computation tasks more than 200 times faster than using the MOVES batch mode and the results generated are exactly the same (Liu et al. 2019a).

MOVES-Matrix can be used to model the emissions from individual vehicles and eco-driving applications (Xu, et al., 2018; Guensler, et al., 2017), and can be easily coupled with vehicle activity analysis (Li et al. 2017, 2016; Xu et al. 2017; Liu et al. 2019b) by importing second-by-second vehicle operations. The matrix structure also facilitates sensitivity analysis and uncertainty assessment of MOVES algorithms (Lu et al., 2021) and can be integrated with large-scale transit operation data for energy use modeling (Fan et al. 2022). MOVES-Matrix can also be applied to different transportation models, such as travel demand models (Xu, et al. 2018), dispersion models (Kim, 2020; Kim et al., 2020; Liu et al., 2019b, 2017; Lu et al., 2022), and microscopic traffic simulation models (Xu et al. 2016). For example, assessment of the emissions and energy consumptions of the ARC network (149,000 roadway segment links as described above) based on output from activity-based model (ABM) only takes a few hours to run using MOVES-Matrix (Zhao, 2020). In 2017/2018, the team also developed MOVES-Matrix 2.0, which also supported the rapid analysis of engine starts, truck hoteling, evaporative sources, brake/tire wear (Xu et al. 2018).

Georgia Tech continues to collaborate with an expanding set of MOVES-Matrix users. The research team integrated MOVES-Matrix in various EPA meetings and presentations, and provided FHWA with model set and the matrices of Washington D.C. The team provided the matrices of Taxes to researchers at Texas A&M Transportation Institute, and the modeling tool has been integrated into their statewide live online roadway performance assessment tools. Georgia Tech has collaborated with CARTEEH to implement MOVES-Matrix online repository.

The team are also working with University of Tennessee (matrices of Tennessee), University of Wyoming (matrices of Wyoming and Utah), Pennsylvania State University (matrices of Pennsylvania), the University of Vermont (matrices of Vermont), Rensselaer Polytechnic Institute (matrices of New York), Beijing Jiaotong University in Beijing (matrices of Atlanta, GA), Tongji University in Shanghai (matrices of California), MIT Senseable City Lab (matrices of New York City, NY) and Inha University in Incheon, Korea (matrices of Atlanta, GA) by providing the matrices that they need for their research. Consulting firms such as AECOM and Rybinski Engineering (matrices of Delaware) are also using MOVES-Matrix for their projects. The continued advancements in the state-of-art techniques to distribute MOVES in Georgia Tech's high-performance parallel computing system (PACE) system was highlighted in the PACE Newsletter in Spring 2020. At Georgia Tech, MOVES-Matrix is utilized in undergraduate courses of CEE4620: Environmental Monitoring and Impact Assessment, CEE4670: Transportation and Health, and the Undergraduate Capstone projects. MOVES-Matrix is also incorporated in graduate courses CEE6625: Transportation, Energy, and Air Quality, and CEE6701: Urban Transportation Planning.

Because MOVES is the regulatory model required for use by EPA, the research team must make sure that any modeling approach that employs MOVES3 outputs in a lookup matrix process yield exactly the same emission rates as when applying any individual MOVES run output directly to the same transportation operating conditions.

In this project, the team identified all of the major updates within MOVES3 by reviewing the EPA documentation (U.S. EPA, 2020a, 2020b, 2020c), and assessed changes in the MOVES3 default database (movesdb20220802). The MOVES3 model was configured and installed on the PACE supercomputing cluster (Georgia Tech, 2023), and thousands of model runs were launched to iterate all possible combinations of input variables. The output emission rates were stored as look-up tables, and a Python®-based module was developed to query and assemble the emission rates, to ensure that the runs provide the exact same outputs as MOVES3. A case study of 3,000 links were performed to demonstrate the efficiency of the running module, and the results were compared against MOVES3 (based on the exact same input configurations).

2. Methodology

MOVES-Matrix is composed of the outputs from a huge number of MOVES model runs. The basic process is to run MOVES across all combinations of variables that affect output emission rates, where each MOVES run yields a pollutant emission rate for single vehicle source type, model year, vehicle fuel type, on-road operating condition (average speed and road type, or a single on-road VSP/STP operating mode bin), calendar year, temperature, humidity condition, and other applicable regional regulatory parameters (fuel properties, I/M program characteristics). After conducting more than 100,000 MOVES model runs, the resulting MOVES emission rate matrix (MOVES-Matrix) can be queried to obtain the exact same emission rates that would be obtained for any MOVES model run, without ever having to launch MOVES again, or transfer MOVES outputs into the analyses.

Figure 1 provides an overview of MOVES-Matrix application process (Liu et al. 2019). Users first identify the MOVES-Matrix subset required, by specifying calendar year, fuel month, and meteorology data. The user can then access each cell that contains an emission rate for a specific vehicle class and model year from MOVES-Matrix and weight each emission rate by on-road activity to reassemble the fleet emission rate. Figure 2 shows the emission rate assembly process for MOVES-Matrix. The weighting process is exactly the same as is used inside MOVES to generate a fleet composite emission rate for a link (i.e., no approximations are employed in the process). Hence, MOVES-Matrix yields the exact same emission rates as a direct MOVES run, but in a fraction of the time, because all potential MOVES runs are pre-processed and the user is simply looking up applicable results in the matrix. Each MOVES run has already performed the complex emission rate calculations and adjustments for temperature, humidity, fuel composition, I/M program, etc., and MOVES-Matrix contains the resulting emission rates. Hence, the user can apply MOVES-Matrix results very efficiently, and the system is efficient enough to enable large-scale and real-time emission estimation.

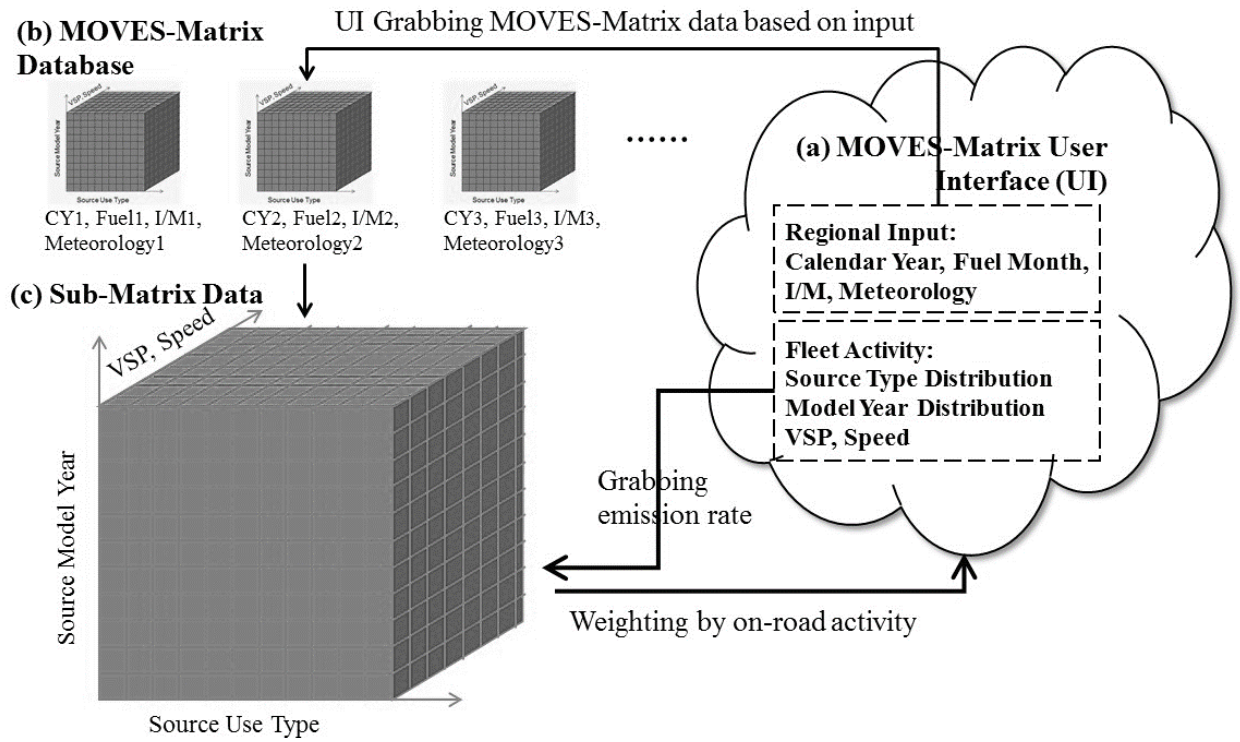

Figure 1. MOVES-Matrix Conceptual Flow (Liu et al. 2019a)

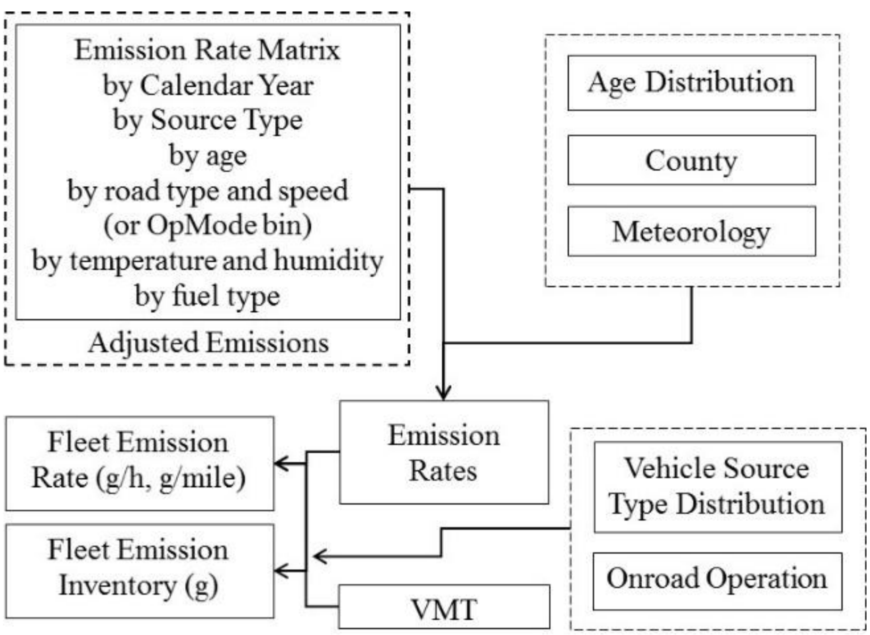

Figure 2. MOVES-Matrix Data Processing Overview (Liu et al. 2019a)

The team performed a thorough review of the configurations of fuel and I/M programs in MOVES3, and compared these settings vs. MOVES 2014. The team updated the MOVES-Matrix region directory accordingly (added new regions due to any finer break of fuel and I/M combinations, and removed redundant regions as needed to reduce run time), and developed a comprehensive estimate of run time required to complete the runs for all modeling regions in the nation.

The research team has priority access to the Partnership for an Advanced Computing Environment (PACE) high performance computing (HPC) clusters of Phoenix (Georgia Tech, 2023). PACE is a collaboration between Georgia Tech faculty and the Office of Information Technology, and was established for the primary purpose of providing an environment for distributed high performance computing. The configuration of MOVES3 (version 3.1.0) on PACE required that internal MOVES model updates be accommodated, which includes support of Java 17 and the operating system update to Red Hat Enterprise Linux (RHEL) 7.9 from RHEL 7.6 (a major update in February 2021 that changed the Linux Kernel). The team also made significant efforts to accommodate the PACE migration to Slurm in early 2023, which slightly reduced the running efficiency of the MOVES-Matrix generation.

A major update from MOVES2014 to MOVES3 was the introduction of MariaDB to replace MySQL. Although PACE did have MySQL installed, MariaDB was not on the supported software list, and the latest PACE policy prohibits software installation requests that involve deployment of database servers. The research team managed to configure the computing clusters to launch MariaDB without root access to PACE, by installing MariaDB binary tarballs.

When a MOVES job is launched on a cluster machine, the scripts first install a MariaDB server (MySQL for MOVES 2014) without root authority on the supercomputing cluster (by unzipping the binary tarball version of MariaDB and migrating the my.cnf file), and then start the SQL server on an available port. The script then proceeds to install MOVES3 on the machine by unzipping the MOVES 3.1.0 source files on the disk. The installations of both MariaDB and MOVES3 have to be configured to the temporary storage directory on each cluster, which is automatically deleted after each job (to avoid clogging the cluster). Therefore, the MariaDB and MOVES installations need to be launched every time a job is submitted.

MOVES command line Java processes are then launched to create input and output database files respectively. The output files are zipped and stored on PACE persistent storage. Given PACE's limits on the number of jobs that can be launched simultaneously (500 jobs according to PACE policy) and on the number of computational nodes that can be used at the same time (3,000 nodes per research group), the team further optimized the scripts to better utilize the computational resources by allocating two nodes to each matrix job (i.e., one node to run the thread of SQL, and one to host the thread of MOVES).

For each modeling region (defined by a unique combination of fuel and I/M configuration for which a suite of MOVES iteration runs is performed), PACE runs MOVES across a total of 21 calendar years × 3 fuel months × 23 temperature bins (0-110 degrees with 5-degree interval) × 21 humidity bins (0-100% with 5% interval) = 30,429 MOVES runs. For each of these runs, the input variables are iterated to include all 13 source types, all 31 model year groups, all facility types, all opMode bins, and all pollutants.

The team completed the model runs for the Atlanta, GA (Fulton County), and performed a case study to make sure that MOVES-Matrix 3.0 generated the exact same results as MOVES3. The speed performance associated with running MOVES-Matrix versus launching MOVES individually was also reassessed to compare with the prior MOVES-Matrix model version. A total of 3,000 roadway links were coded with randomized input variables. Traffic volume ranged from 500 to 1,000 vehicles (at an interval of one vehicle), and average speeds were set between 1.0 mph and 80.0 mph (at an interval of 0.1 mph). The facility types included all of the MOVES types: urban restricted highway, urban unrestricted highway, rural restricted highway, and rural unrestricted highway. The source type and vehicle age distributions were also randomized, making sure that each link included all 13 source types and 31 vehicle age groups.

The testing environment was based on these specifications to compare the running time of MOVES vs. MOVES-Matrix. CPU: Intel I7-8700K @ 3.70GHz; RAM: 32 GB GDDR4 (2×16GB modules); Operating System: Microsoft® Windows 10 Pro.

The emission results of total gaseous hydrocarbons (pollutant ID #1), carbon monoxide (CO) (pollutant ID #2), oxides of nitrogen (NO_x) (pollutant ID #3), Volatile Organic Compounds (VOC) (pollutant ID #87), Atmospheric CO_2 (pollutant ID #90), total energy consumption (pollutant ID #91), primary exhaust PM_{10} (pollutant ID #100), and primary exhaust $PM_{2.5}$ (pollutant ID #110) were compared between MOVES vs. MOVES-Matrix.

3. Results and Discussion

The team evaluated MOVES-Matrix 3.0 performance by comparing the runs across numerous scenarios (running module queries of emission rates) against the results taken from MOVES3 directly, to ensure MOVES-Matrix 3.0 generates the exact same results with MOVES3. The team assessed the matrix generation time per region, as presented in Section 3.1, and that for the nation, by reviewing the regional definitions in MOVES3 based on fuel and I/M configurations and comparing them to MOVES 2014, as detailed in Section 3.2. The comparison of run outputs for MOVES-Matrix 3.0 vs. MOVES3, and runtime efficiency (computer time to perform the applications with MOVES-Matrix lookups vs. running MOVES3 directly to obtain the outputs), can be found in Section 3.3.

3.1 Generation of Matrices by Region

One of the major updates in terms of emissions and energy use modeling is the new source type physics parameters for vehicle-specific-power/scaled-tractive-power configurations (U.S. EPA, 2020d, 2023) for various source types based on model year (listed in the "sourceusetypephysics" table within the MOVES3 database). The updated MOVES3 database introduced 185 unique combinations of VSP/STP parameters (out of 206 potential iterative combinations), while MOVES 2014 only employed 22 unique combinations. The team updated the corresponding code for both the generation of the matrices and the MOVES-Matrix running module to address the changes in input parameters. These parameter updates led to a significant increase in iteration run time to generate matrices for MOVES-Matrix 3.0 (for MOVES3) compared to MOVES-Matrix (for MOVES 2014) on the supercomputing cluster, due to the larger number of iterations. Empirical tests showed approximately 30% more time to finish a MOVES3 run than a MOVES 2014 run on the supercomputing clusters, given the same amount of runtime resources assigned by PACE.

Empirical tests showed that setting up MariaDB and MOVES3 typically takes around 30 minutes, while completing a single MOVES run ranges from 15 to 45 minutes (to iterate through all source types, all vehicle model year groups, all facility types, all pollutants, and all opMode bins, for a single combination of temperature and humidity). Both MariaDB and MOVES3 must be installed in a temporary storage directory on the computing clusters (i.e., the system variable of {TMPDIR}), which is created for each job. Once a job concludes, the PACE scheduling system clears this directory, removing the installed environment.

To minimize installation time, the goal is to execute as many MOVES runs as possible within a single installation. However, increasing the number of MOVES model runs per job extends the total run time (referred to as 'wall time' on PACE), which in turn lengthens the queue time. Extended wall time also make the jobs vulnerable of disruptions from system updates, hardware failures, and other interruptions.

Given these considerations, the team opted for a 24-hour wall time for each job. Empirical tests suggested that between 50 and 80 MOVES runs can be completed within this timeframe.

PACE clusters work on a scheduling algorithm that dynamically assigns node computation power based on the requested resources, and intense requests from other users might delay the launch of the MOVES jobs (longer queue time).

Generating matrices for a region takes approximately 3 to 7 days for MOVES 2014, and take around 15 to 20 days for MOVES3 due to the increased number of sub-source types (VSP/STP configurations) that increased the running time, and the additional burden to install MariaDB.

3.2 Fuel Specifications and Inspection/Maintenance Programs

MOVES 2014b includes 3,229 counties. MOVES3 introduced the Kusilvak Census Area, Alaska to the county list (but users can still use its former name of Wade Hampton Census Area). Otherwise, the lists of modeled counties are identical for the two versions of MOVES. The Kusilvak Census Area and Wade Hampton Census Areas are marked with different county IDs in MOVES, but their fuel supply and I/M programs are exactly the same in the MOVES3 database (not surprising, given they are essentially referring to the same administrative division).

The fuel supply updates in MOVES3 yield 40 fuel supply scenarios, given the increased number of fuel formulation regions and the updated one-to-many county-to-region relationships (there are only 22 fuel supply scenarios in MOVES 2014b). The review of the fuel supply updates ("fuelsupply" table in the default MOVES databases) identified 23 fuel formulation regions in MOVES3 (fuel formulation region #700000000 that covers various Boroughs and Census Areas in Alaska) rather than 22 regions in MOVES 2014b. The resulting change in these MOVES3 parameters yields a large increase in one-to-many corresponding relationships between counties vs. the fuel formulation regions. In MOVES 2014b, every County corresponds to only one fuel region (across all the years), while in MOVES3, one County can correspond to various fuel regions (one fuel region per year). An example is provided below, where Gwinnett County, GA (County ID #13135) corresponds to three fuel formulation regions (#100000000, #170000000, and #178000000) in MOVES3 while it only corresponds to one (#170000000) in MOVES 2014b. The fuel scenarios for the state of Georgia are shown in Figure 3, where the updated fuel supply settings can be observed for the metro Atlanta area (two scenarios in MOVES 2014b and three in MOVES3).

Table 1. Fuel Formulation Regions for Gwinnett County, GA in MOVES 2014b vs. MOVES3

Database	Fuel Formulation Region # (1990 to 2001)	Fuel Formulation Region # (2002 to 2015)	Fuel Formulation Region # (2016 to 2019)	Fuel Formulation Region # (2020 to 2060)
MOVES 2014b	170000000	170000000	170000000	170000000
MOVES3	100000000	170000000	178000000	100000000

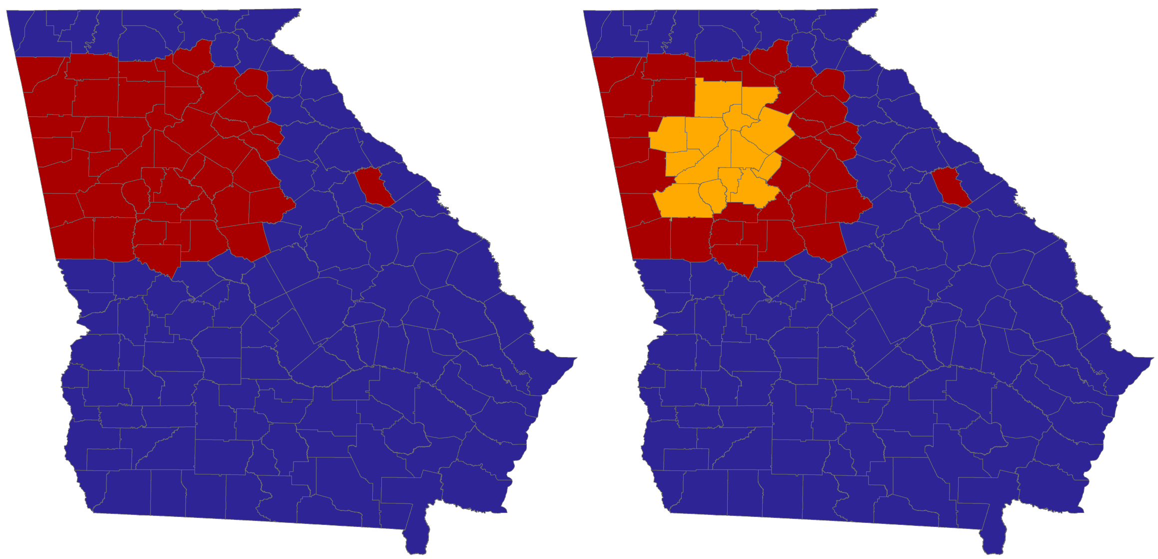

Figure 3. Fuel Supply Scenarios in MOVES 2014b (Left) vs. MOVES3 (Right) for Georgia

The team updated the examining algorithm to check the unique I/M programs, and found 87 I/M scenarios in MOVES3 (version 3.1.0), whereas MOVES 2014b has 84 scenarios. Major updates on the I/M programs include a finer division of New Jersey, allocating I/M programs to various counties in North Carolina and Louisiana (no default I/M programs in MOVES 2014b for these counties), etc.

The review of the default fuel supply configurations and I/M programs were reviewed by examining the MOVES3 database (the tables of "fuelsupply" and "imcoverage") vs. MOVES 2014b indicated that MOVES3 introduced finer resolution in terms of defining the regions, and it was found that MOVES 2014b defined a total of 109 regions (unique combinations of fuel vs. IM programs), and that MOVES3 (version 3.1.0) defined 122 regions. An increase of 10.9% more modeling regions are found in MOVES3, which increases the run time by about the same amount to run nationwide matrices.

3.2 Emission and Energy Consumption Results

The more refined definitions of VSP/STP parameters in MOVES3 also led to an increase in processing time to prepare sub-matrices for model applications, due to the assembly of emission rates across a larger number of combinations (and a slightly increased time for input/output). Instead of hard coding the VSP/STP parameters for MOVES-Matrix 2014, the MOVES-Matrix 3.0 running module integrates a separate input file of the VSP/STP parameters, while the rest of the algorithm remains the same with the previous running module, where the emission rates are weighted by source type, model year, and opMode bin activity (by facility type), as shown in Equation (1).

$$ER_{fleet} = \sum_{ST} \sum_{MY} \sum_{OpMode(FT)} ST\% \times MY\%_{ST} \times ER_{ST,MY} \times ER_{OpMode} \quad (1)$$

For the MOVES-Matrix 3.0 application, the increase in run time is minimal, only about 1% to 2%, when processing a small number of tasks (specifically when the I/O of the VSP/STP table is a factor). Pulling emission rates from MOVES3 matrices remains highly efficient, compared with running MOVES3 (which takes significantly longer to run). When processing a large number of tasks, such as over 100 roadway links, the increase in run time for MOVES3-Matrix is negligible.

The running module finished the 3,000-link case study within approximately 14 seconds, while the MOVES3 application took approximately 3 hours and 20 minutes to finish the same emission rate processes. This was consistent with MOVES-Matrix 2014 vs. MOVES2014, where MOVES-Matrix 2014 was more than 200 times faster than running MOVES. The results of MOVES-Matrix 3.0 and MOVES3 are shown in Figure 4 and Figure 5, where the results are exactly the same. The case study indicates that the algorithm for developing MOVES-Matrix 2014 (that pre-running MOVES across all possible combination of input variables and store the output as lookup tables) also applies to MOVES3. MOVES-Matrix 3.0 generates the exact same results as MOVES3, but now runs more than 800 times faster than MOVES, given the net increase in MOVES3 run time.

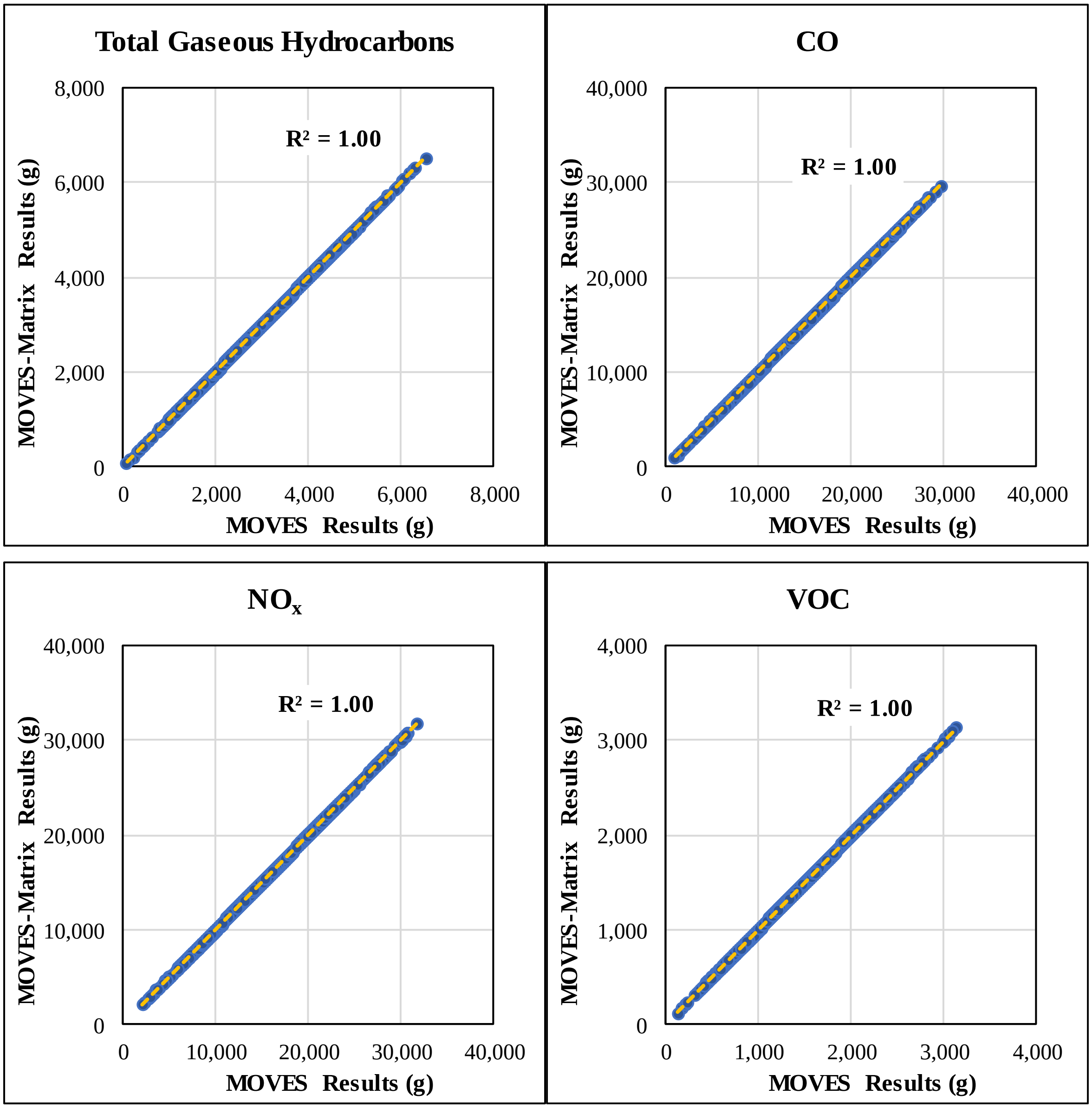

Figure 4. Results Comparison between MOVES-Matrix 3.0 vs. MOVES3: Total Gaseous Hydrocarbons, CO, NO_x and VOC

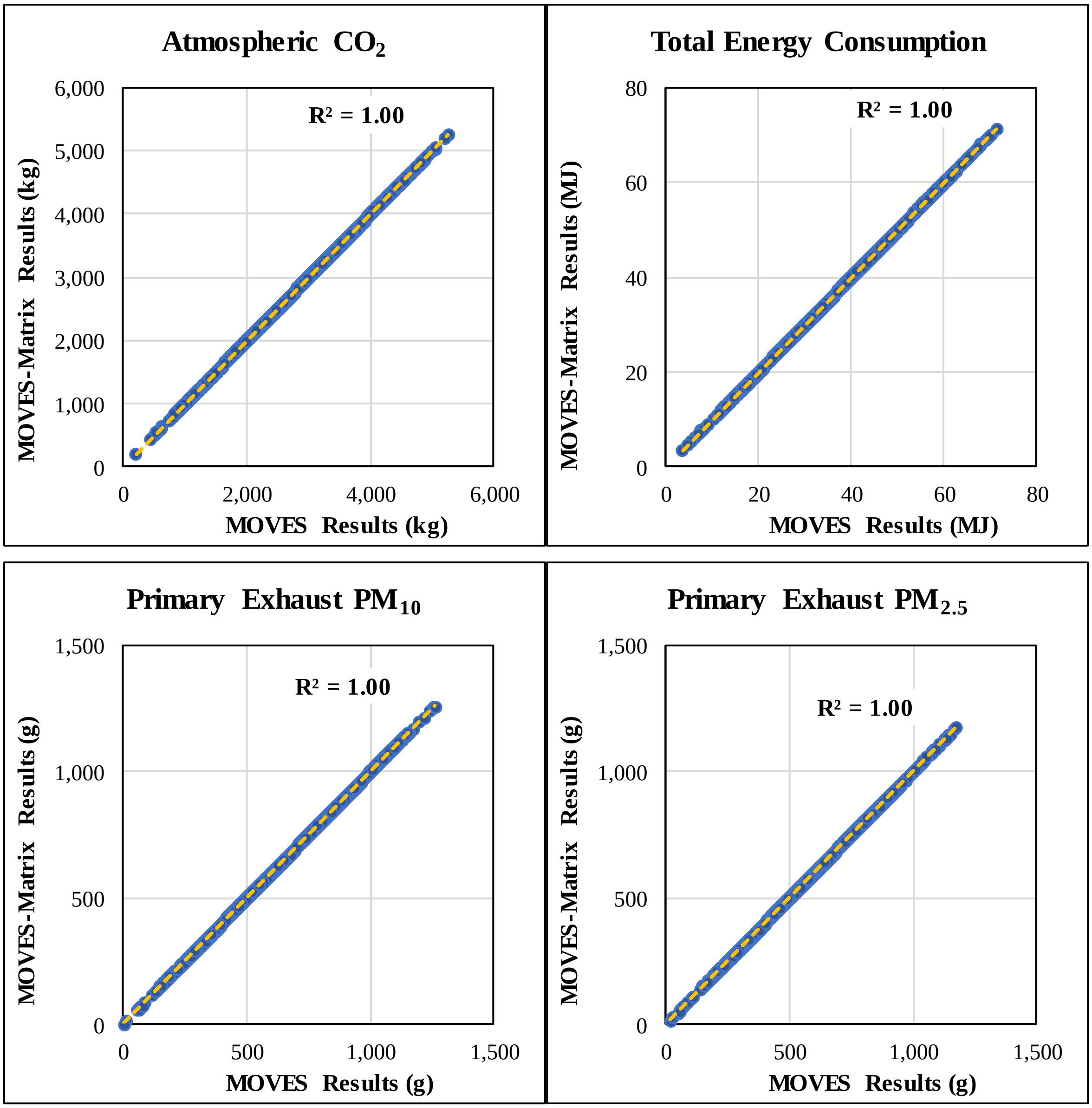

Figure 5. Results Comparison between MOVES-Matrix 3.0 vs. MOVES3: Atmospheric CO_2, Total Energy Consumption, Primary Exhaust PM_{10} and Primary Exhaust $PM_{2.5}$

4. Conclusions and Ongoing Work

This report documents the development of MOVES-Matrix 3.0, based on the latest version of EPA's regulatory model of MOVES3. MOVES3 was configurated and launched on the PACE supercomputing cluster for a total of 21 calendar years × 3 fuel months × 23 temperature bins (0-110 degrees with 5-degree interval) × 21 humidity bins (0-100% with 5% interval) = 30,429 runs. For each run, the input variables were iterated to include all 13 source types, all 31 model year groups, all facility types, all opMode bins, and all pollutants. The output was converted to lookup tables, and a run module was developed to help users assemble the emission rates to fleet results.

The internal model review of program software updates and default fuel supply configurations and I/M programs in MOVES3 vs. MOVES 2014b indicated that a total of 122 regions are defined in MOVES3 compared with 109 regions in MOVES 2014b. The case study of 3,000 roadway links modeled in both MOVES3 and MOVES-Matrix 3.0 indicated that MOVES-Matrix 3.0 generates the exact same results with MOVES3 model, but runs 800 times faster.

The team finished the model runs for Atlanta, GA, and is working on applying the same algorithm to the latest version of MOVES4. Given that there is no change in terms of the scripted environment (and model algorithm) from MOVES3 to MOVES4, the team anticipates that the research design is still valid and that MOVES-Matrix 4.0 can be generated using the same algorithm.

References

CRC. 2017. "Scoping Study for Rewrite of MOVES for Efficiency." CRC Report No. A-110. http://crcsite.wpengine.com/wp-content/uploads/2019/05/CRC-A110-Final-Report_31Aug17.pdf.

Fan, Huiying, Hongyu Lu, Ziyi Dai, Reid Passmore, Angshuman Guin, Kari Watkins, and Randall Guensler. 2022. "Combined Effect of Changes in Transit Service and Changes in Occupancy on Energy Consumption per Passenger-Mile." *Transportation Research Record: Journal of the Transportation Research Board*. August 2022. https://doi.org/10.1177/03611981221111116.

Georgia Tech, Georgia Institute of Technology (2023). Partnership for an Advanced Computing Environment. https://pace.gatech.edu. Accessed October 1, 2023.

Guensler, Randall, Hongyu Lu, William Reichard, Ziyi Dai, Tian Xia, Angshuman Guin, and Michael O. Rodgers (2021). "AERMOD, RLINE, and RLINEXT Case Study Analyses in Atlanta, GA." Prepared for the Federal Highway Administration and Georgia Department of Transportation. Final Report (FHWA-GA-21-2024). https://trid.trb.org/view/1896759. November, 2021.

Guensler, Randall, and John D. Leonard II. 1995. "A Monte Carlo Technique for Assessing Motor Vehicle Emission Model Uncertainty." In *Transportation Congress, Volume 2; American Society of Civil Engineers*, 2098–2115. New York, NY, U.S.

Guensler, Randall, Karen K. Dixon, Vetri Venthan Elango, and Seungju Yoon. 2004. "MOBILE Matrix: Application of Georgia Statewide Multimodal Transportation Planning Tool for Rural Areas." *Transportation Research Record: Journal of the Transportation Research Board* 1880 (1): 83–89. https://doi.org/10.3141/1880-10.

Guensler, Randall, Haobing Liu, Yanzhi "Ann" Xu, Alper Akanser, Daejin Kim, Michael P. Hunter, and Michael O. Rodgers. 2017. "Energy Consumption and Emissions Modeling of Individual Vehicles." *Transportation Research Record: Journal of the Transportation Research Board* 2627 (January): 93–102. https://doi.org/10.3141/2627-11.

Guensler, Randall, Michael O. Rodgers, John D. Leonard II, and William Bachman. 2000. "A Large Scale Gridded Application of the CALINE4 Dispersion Model." In *Transportation Planning and Air Quality IV: Persistent Problems and Promising Solutions*, edited by Ed Arun Chatterjee. New York, NY: American Society of Civil Engineers.

Kim, Daejin. 2020. "Large-Scale, Dynamic, Microscopic Simulation for Region-Wide Line Source Dispersion Modeling." Ph.D Dissertation, Georgia Institute of Technology.

Kim, Daejin, Haobing Liu, Michael O. Rodgers, and Randall Guensler. 2020. "Development of Roadway Link Screening Model for Regional-Level Near-Road Air Quality Analysis: A Case Study for Particulate Matter." *Atmospheric Environment* 237 (September): 117677. https://doi.org/10.1016/j.atmosenv.2020.117677.

Li, Hanyan, Haobing Liu, Xiaodan Xu, Yanzhi “Ann” Xu, Michael O Rodgers, and Randall Guensler. 2016. “Emissions Benefits from Reducing Local Transit Service Deadheading: An Atlanta Case Study.” In *95th Annual Meeting of the Transportation Research Board*. Washington, D.C., U.S.

Li, Hanyan, Yuanbo Wang, Xiaodan Xu, Haobing Liu, Angshuman Guin, Michael O Rodgers, Michael P Hunter, Jorge A Laval, and Randall Guensler. 2017. “Assessing the Time, Monetary, and Energy Costs of Alternative Modes.” In *97th Annual Meeting of Transportation Research Board*. Washington, D.C., U.S.

Liu, Haobing, Randall Guensler, Hongyu Lu, Yanzhi Xu, Xiaodan Xu, and Michael O. Rodgers. 2019a. “MOVES-Matrix for High-Performance On-Road Energy and Running Emission Rate Modeling Applications.” *Journal of the Air & Waste Management Association*, July, 10962247.2019.1640806. https://doi.org/10.1080/10962247.2019.1640806.

Liu, Haobing, Michael O. Rodgers, and Randall Guensler. 2019b. “Impact of Road Grade on Vehicle Speed-Acceleration Distribution, Emissions and Dispersion Modeling on Freeways.” *Transportation Research Part D: Transport and Environment* 69: 107–22. https://doi.org/10.1016/j.trd.2019.01.028.

Liu, Haobing, Xiaodan Xu, Michael O. Rodgers, Yanzhi “Ann” Xu, and Randall Guensler. 2017. “MOVES-Matrix and Distributed Computing for Microscale Line Source Dispersion Analysis.” *Journal of the Air & Waste Management Association* 67 (7): 763–75. https://doi.org/10.1080/10962247.2017.1287788.

Liu, Haobing, Yanzhi Xu, Michael O Rodgers, Alper Akanser, and Randall Guensler. 2016. “Improved Energy and Emissions Modeling for Project Evaluation (MOVES-Matrix).” https://ncst.ucdavis.edu/wp-content/uploads/2014/09/10-10-2016-NCST_MOVESMatrix_FinalReport_version-4.pdf.

Lu, Hongyu, Haobing Liu, Daejin Kim, Tian Xia, Ziyi Dai, William Reichard, Yingping Zhao, Angshuman Guin, Michael O. Rodgers, and Randall Guensler. 2022. “Comparative Analysis of AERMOD-Predicted Pollutant Concentrations by Input Source Type in Atlanta, GA.” *Transportation Research Record: Journal of the Transportation Research Board*. October 2022. https://doi.org/10.1177/03611981221123806.

Lu, Hongyu, Haobing Liu, Tian Xia, Angshuman Guin, Michael O. Rodgers, and Randall Guensler. 2021. “Investigation of MOVES Project-Level Uncertainty: Impact of Temporal- and Spatial- Aggregation of Onroad Operating Conditions on Emission Rate Estimates”. In *100th Annual Meeting of the Transportation Research Board*. Washington, D.C., U.S.

U.S. EPA. 2016. “Description and History of the MOBILE Highway Vehicle Emission Factor Model.” U.S. Environmental Protection Agency. 2016. https://www.epa.gov/moves/description-and-history-mobile-highway-vehicle-emission-factor-model.

U.S. EPA. 2020a. “Emission Adjustments for Temperature, Humidity, Air Conditioning and Inspection and Maintenance for Onroad Vehicles in MOVES3.” https://www.epa.gov/sites/production/files/2020-12/documents/420r20013.pdf.

U.S. EPA. 2020b. “Exhaust Emission Rates for Heavy-Duty Onroad Vehicles in MOVES3.” https://www.epa.gov/sites/production/files/2020-11/documents/420r20018.pdf.

U.S. EPA. 2020c. “Exhaust Emission Rates for Light-Duty Onroad Vehicles in MOVES3.” https://www.epa.gov/sites/production/files/2020-12/documents/420r20019.pdf.

U.S. EPA. 2020d. “Fuel Supply Defaults: Regional Fuels and the Fuel Wizard in MOVES3.” https://www.epa.gov/sites/production/files/2020-11/documents/420r20017.pdf.

U.S. EPA, United State Environmental Protection Agency (2023a). MOVES and Other Mobile Source Emissions Models. https://www.epa.gov/moves. Last Upated on September 22, 2023.

U.S. EPA, United State Environmental Protection Agency (2023b). Latest Version of MOtor Vehicle Emission Simulator (MOVES) - MOVES3: Latest Version of Motor Vehicle Emission Simulator. https://www.epa.gov/moves/latest-version-motor-vehicle-emission-simulator-moves. Last Upated on September 22, 2023.

U.S. EPA, United State Environmental Protection Agency (2023c). MOVES Versions in Limited Current Use. https://www.epa.gov/moves/moves-versions-limited-current-use. Last Upated on September 22, 2023.

Xu, Xiaodan, Hanyan Li, Haobing Liu, Michael O. Rodgers and R. Guensler. 2018. “Evaluation of Transit Ecodriving in Rural, Suburban, and Urban Environments.” *Transportation Research Record: Journal of the Transportation Research Board* 2672 (8): 152-164. https://doi.org/10.1177/0361198118797778.

Xu, Xiaodan, Haobing Liu, J.M. Anderson, Yanzhi “Ann” Xu, Michael Hunter, M.O. Rodgers. 2016. “Estimating Project-Level Vehicle Emissions with VISSIM and MOVES-Matrix.” In *95th Annual Meeting of the Transportation Research Board*. Washington, D.C., U.S.

Xu, Yanzhi, Hanyan Li, Haobing Liu, Michael O. Rodgers, and Randall L. Guensler. 2017. “Eco-Driving for Transit: An Effective Strategy to Conserve Fuel and Emissions.” Applied Energy, 194: 784–97. https://doi.org/10.1016/j.apenergy.2016.09.101.

Zhao, Yingping. 2020. “Distributional Impact Assessment of Transportation Policy Using Activity-Based Models with Path Retention.” Ph.D Dissertation, Georgia Institute of Technology.

Data Summary

Products of Research

As described in this report, the team developed the emission and energy use matrices based on iterative runs of EPA's MOVES model, and the results of all possible combinations of input variables were generated on the PACE supercomputing cluster at Georgia Tech.

Data Format and Content

The energy and emission rate results are directly pulled from MOVES outputs, and can be used to generate the exact same results as regulatory analyses. The matrices are public domain and can be found at https://zenodo.org/records/13961059.

Data Access and Sharing

The format and content of the MOVES-Matrix data sets are documented in the NCST MOVES-Matrix overview and training documents https://github.com/gti-gatech/moves_training.

Reuse and Redistribution

The MOVES-Matrix data are open source can be downloaded, used, and freely redistributed using the link provided above.

The energy and emission rate results matrices should be cited as follows:

Lu, H. (2024). MOVES-Matrix 3.0: On-Road Energy and Emission Modeling with High-Performance Supercomputing [Data set]. Zenodo. https://doi.org/10.5281/zenodo.13961059