A Modeling Framework for Near-Road Population Exposure to Traffic-Related $PM_{2.5}$ and Environmental Equity Analysis: A Case Study in Atlanta, Georgia

September 2024

A Research Report from the National Center for Sustainable Transportation

Hongyu Lu, Georgia Institute of Technology

Dr. Haobing Liu, Tongji University

Dr. Michael O. Rodgers, Georgia Institute of Technology

Dr. Randall Guensler, 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. NCST-GT-RR-24-27	**2. Government Accession No.** N/A	**3. Recipient's Catalog No.** N/A
4. Title and Subtitle A Modeling Framework for Near-Road Population Exposure to Traffic-Related $PM_{2.5}$ and Environmental Equity Analysis: A Case Study in Atlanta, Georgia		**5. Report Date** September 2024 **6. Performing Organization Code** N/A
7. Author(s) Hongyu Lu, GRA, https://orcid.org/0000-0002-0170-7169 Haobing Liu, Ph.D., https://orcid.org/0000-0002-3588-4593 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 Georgia Institute of Technology School of Civil and Environmental Engineering 790 Atlantic Dr, Atlanta, GA 30332		**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 (April 2019 – September 2024) **14. Sponsoring Agency Code** USDOT OST-R

15. Supplementary Notes
DOI: https://doi.org/10.7922/G2JS9NSS
Dataset DOI: https://doi.org/10.5281/zenodo.13878218

16. Abstract
In this study, a modeling framework for population exposure to traffic-related $PM_{2.5}$ with high spatiotemporal resolution is proposed and applied to the I-575/I-75 Northwest Corridor (NWC) in Atlanta, GA, for environmental equity analysis. The analyses retrieved trip data from the Atlanta Regional Commission's (ARC) Activity-Based Model 2020 (ABM2020), after implementing path retention algorithms (Zhao, et al., 2019) to generate individual travel paths for more than 20 million predicted vehicle trips. Emission rates for each link were retrieved from MOVES-Matrix given the ABM link speed and facility type, the ARC's county-level fleet composition data, and regional fuel properties and I&M program parameters. High-resolution downwind concentration profiles were predicted using EPA's AERMOD microscale dispersion model with AERMET meteorology profiles for a huge array of receptors. Trip-end locations were derived from the ABM trip data, and the on-road trajectories for each person-trip (vehicle trace data) were derived from the travel paths through network. ABM synthetic household and person data were used in demographic assessment, and linked to representative household latitude and longitude locations in the Epsilon 2019 household demographic dataset. Individual exposure to traffic-related $PM_{2.5}$ in time and space (average hourly concentration) was assessed by overlaying the second-by-second person location profiles (for 24 hours) against the hourly predicted $PM_{2.5}$ concentration profiles. The analyses summarize the results across 16 demographic groups and the aggregate population exposure are compared to assess potential impact differences across demographics. High-income households in the corridor were exposed to less traffic-related air pollution as they tended to live further from the freeways. The analyses did not reveal large disproportionate negative impacts on low income groups along this specific corridor, but lager disproportionate negative impacts are expected elsewhere in the metro area due to the spatial clustering of income groups along other corridors. Overall, the research demonstrates the applicability of the modeling framework and describes how the various elements (e.g., link screening, dispersion modeling, path tracing, etc.) are optimized on the supercomputing cluster.

17. Key Words Microscale air quality impact assessment, AERMOD, vehicle emissions, population exposure, environmental equity		**18. Distribution Statement** No restrictions.	
19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages** 52	**22. Price** 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: University of California, Davis; University of California, Riverside; University of Southern California; California State University, Long Beach; Georgia Institute of Technology; and University of Vermont. More information can be found at: ncst.ucdavis.edu.

Disclaimer

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

A Modeling Framework for Near-Road Population Exposure to Traffic-Related $PM_{2.5}$ and Environmental Equity Analysis: A Case Study in Atlanta, Georgia

A National Center for Sustainable Transportation Research Report

September 2024

Hongyu Lu, Graduate Research Assistant, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta

Dr. Haobing Liu, Professor, Tongji University, Shanghai, China

Dr. Michael O. Rodgers, Regents' Researcher and Adjunct Regents' Professor, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta

Dr. Randall Guensler, Guy J. Lookabaugh Professor, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta

NCST

[page intentionally left blank]

TABLE OF CONTENTS

List of Tables

List of Figures

A Modeling Framework for Near-Road Population Exposure to Traffic-Related $PM_{2.5}$ and Environmental Equity Analysis: A Case Study in Atlanta, Georgia

EXECUTIVE SUMMARY

Exposure to traffic-related air pollution, especially particulate matter (PM) is widely known to yield adverse health impacts. However, estimating daily population exposure to near-road pollutant concentrations is often limited by a lack of high spatial and temporal resolution of pollutant concentrations, high spatial and temporal variation in human activity, unquantified presence of mitigation factors that reduce pollutant exposure (e.g., HVAC system impacts on indoor concentrations), and confounding factors (such as other sources of indoor and outdoor air pollution). All these elements are of critical importance in quantifying individual and population exposure and for conducting environmental equity assessments across demographic groups. Existing models typically used to assess population exposure are mostly based on coarse input data (area-wide concentrations, for instance), and are not sensitive to variances in travel paths through the transportation system that may increase or decrease pollutant exposure. While previous studies have developed comprehensive modeling frameworks that integrate activity-based models, traffic emissions, dispersion modeling, and dynamic exposure assessment both on the move and stationary (e.g., Hatzopoulou and Miller, 2010; Beckx et al., 2009; Dias and Tchepel, 2018), our research advances this field by enhancing the spatial and temporal resolution of exposure assessments and optimizing the computational processes on a supercomputing cluster. Specifically, our framework leverages high-resolution individual travel paths and downwind concentration profiles to provide a more detailed analysis of population exposure and environmental equity. The research presented in this report implements an initial modeling framework designed to estimate the near-road population exposure to traffic-related fine particulate matter of 2.5µm or smaller ($PM_{2.5}$) at the individual and household level, while accounting concentration profiles, individual travel paths, trip-end activity, and demographic information. Although the analytical methods do not address mitigating and confounding factors affecting overall exposure, the tools provide a first step in conducting microscale impact assessment with person and vehicle trajectories through time and space.

The proposed modeling framework first retrieved the Atlanta Regional Commission's (ARC) Activity-Based Model 2020 (ABM2020-TIP-A1) output dataset, and implemented a path retention algorithm (Zhao, et al., 2019) to retain vehicle travel paths so that the team could generate second-by-second positions of each vehicle and person modeled by the ABM. High-resolution near-road pollutant concentration profiles were then predicted using AERMOD with variable receptor grids (Kim et al. 2019) and MOVES-Matrix (Liu et al. 2019) for the entire modeled subarea of I-575/I-75 Northwest Corridor (NWC) (Guensler et al. 2021). Finally, the exposure modeling is conducted by plotting the speed-time traces (second-by-second trajectories) of the modeled individuals in their vehicles through the transportation network, accounting for their time on the road and time at each destination. Licensed Epsilon

demographic data were integrated with ARC's household profiles to provide detailed demographic information including ethnicity, address, job, income, etc., at the household level. Individual-level exposure to traffic-related $PM_{2.5}$ was derived from detailed spatial and temporal attributions of travel activity and concentration input and was aggregated to 16 demographic groups that were defined using variables such as household size, number of workers, number of vehicles, household income groups, with and without children, etc. (the same demographic groups that were developed for a prior ARPA-E Department of Energy project). The hourly averages of population exposure were compared across the defined demographic groups by overlaying demographic characteristics over the concentration profiles.

The results indicated that population exposure of various income groups was correlated with distance to the modeled roadway from their households, given that travel activities outside the modeled subarea were not accounted for in this subarea analysis (i.e., the results were largely from the at-home exposure), and households with no or only one worker were found to be exposed to the highest amount of traffic-related $PM_{2.5}$. High-income groups that lived further from the NWC freeways were generally less exposed, while middle-income groups were found to have a slightly higher daily exposure. The overall exposure of low-income households and those with no vehicles were not among the highest exposure groups for this freeway corridor, but this may be more due to the relatively higher income of the households along this corridor and lower variability in household clustering by income than is seen along other metro area corridors.

While the patterns presented by the modeling results could change when the whole metro area is modeled, the research demonstrates the general applicability of the modeling framework and shows how such complex modeling can be conducted on a supercomputing cluster. The team is currently refining the synthetic household and synthetic fleet generation process, and expanding the analyses to the entire metro area.

Chapter 1. Introduction

Ambient air pollution has been widely associated with adverse impacts on human health. Exposure to traffic-related pollution is estimated to cause 15,000 premature deaths in the U.S., and between 184,000 and 242,000 premature deaths per year globally (Frey, 2018). Exposure to traffic-related particulate matter less than 2.5 micrometers in aerodynamic diameter ($PM_{2.5}$), is associated with aggravation of asthma, pulmonary dysfunction, lung cancer, heart disease, stroke, and other illnesses (Volk et al., 2013; Wang et al., 2014; WHO, 2018; Zhang et al., 2013).

In transportation planning and policy assessment, it is important to assess traffic-related equity with respect to benefits (e.g., mobility and accessibility) as well as burdens across demographic groups. Given that exposure to traffic-related air pollutants may lead to inequities in the distribution of health impacts (Clark et al., 2014; Luo et al., 2017; Wu, 2018), tools are needed to assess any disproportionate negative impacts across metro area populations. An equity analysis of transportation investments is required by federal and guidance issued under Title VI of the 1964 Civil Rights Act, and agencies who receive federal funding must make sure that People of Color and low-income populations are not denied the benefits of public investments (Marcantonio et al. 2017).

$PM_{2.5}$ is one of the six principal pollutants with National Ambient Air Quality Standard (NAAQS), established by the U.S. EPA under the Clean Air Act, along with ozone (O_3), carbon monoxide (CO), nitrogen dioxide (NO_2), and particulate matter less than 10 micrometers (PM_{10}). Transportation plans, transportation improvement programs (TIP), and projects must conform to the purpose of the State's air quality implementation plan (SIP) such that these activities will not "create or contribute to any new violations of the NAAQS, increase the frequency or severity of NAAQS violations or delay timely attainment of the NAAQS," as required by the Clean Air Act Section 176(c). Transportation actions in federal projects are subject to transportation conformity in terms of the funding and approval of Federal Highway Administration (FHWA) and/or Federal Transit Administration (FTA).

It is not always feasible to implement large-scale field measurement to collect travel data, quantify mobile source emissions, and measure hourly pollution concentration profiles, due to the technical and cost limitations. Given the dynamic nature of travel activities, variability of fleet composition and traffic flow, and complexity of the roadway networks, mathematical models are frequently used to make such assessments. It important to properly integrate these models into a comprehensive modeling framework, from travel activity to population exposure.

Equity analysis often requires that the modeling framework account for large-scale input both spatially (e.g., for a metropolitan area such as Atlanta, Georgia) and temporally (e.g., the modeling of one whole calendar year). The nature of exposure modeling (e.g., variability of trip patterns, and sensitivity of emissions and dispersion modeling) also requires high-resolution input data. The output of each modeling component serves as an input to the next stage of the modeling, so it is important to reconcile data resolution with model sensitivity. Input data that feed into every modeling component needs to be sufficiently detailed (disaggregated) in temporal and spatial attributes, and each model of the framework needs to be thoroughly reviewed and calibrated, to properly accommodate the fine-level input (appropriate sensitivity). Using coarse input data generated by a previous model in the chain can lead to larger (and sometimes unexpected) variability in the framework outcome; hence, it is important to investigate the uncertainty propagation through the chain of models to account for interacting sensitivity across the models (e.g., from vehicle emissions modeling to dispersion modeling).

It is also important that any proposed modeling framework accommodate spatial-temporal variability of the population activities, which requires detailed travel routes (e.g., individual or household-level) that comes with (or can be coupled with) traveler demographic profiles. Such modeling also needs to be compatible with any required regulatory analysis (and accepted by regulatory agencies). Hence, approved travel demand models (used by the local MPO), are generally coupled with the EPA's MOVES model for vehicle emissions modeling, and EPA's AERMOD model for dispersion modeling.

Development of these kinds of modeling frameworks typically comes with a large computation burden, and model calibration and model verification activities may be time-consuming due to sensitivity across the models that interact with each other. It is important to optimize the modeling process by filtering unnecessary workload for cost-effectiveness purposes, and to make sure the proposed optimization does not cause a significant impact on the overall predictions.

This research aims to develop a modeling framework for near-road population exposure to traffic-related $PM_{2.5}$ emissions. The research objectives can be described as follows:

1) Develop a refined spatial pollutant concentration exposure equity assessment framework with detailed demographic profiles - The household-level demographic profiles from the Epsilon dataset provide various attributes and may cover different households from the ABM household information. This work proposes to link the massive travel activity output from ABM2020 (more than 20 million vehicle trips per day) with the household information by integrating the ABM and Epsilon demographic profiles at the household level. Hourly pollutant exposure equity assessment will be conducted with the high-resolution spatial output from the exposure modeling framework. Marginalized or disadvantaged demographic groups who often bear a disproportionate burden of pollutant impacts will be identified and assessed.
2) Program a population exposure modeling framework with enhanced performance - The proposed population exposure modeling framework includes massive input data and significant amount of computation. This research proposes to accelerate modeling speed by reducing the calculation amount by reducing duplicate model runs (e.g., overlapped vehicle trajectories were removed in the path tracing component) or removing workload that does not significantly contribute to the model outcomes (link screening of source-receptor pairs).

The research presented in this report stems from ongoing dissertation research being conducted by Hongyu Lu in Georgia Tech's School of Civil and Environmental Engineering. This report presents the modeling framework methodology and data in Chapter 2, case study results and a discussion of the results for I-575/I-75 Northwest Corridor (NWC) in Chapter 3, and conclusions and recommendations in Chapter 4. The final dissertation work currently being conducted by Hongyu Lu will include the whole metro area with a refined synthetic fleet and household generation and the uncertainty assessment through the modeling chain.

Chapter 2. Data and Methodology

This chapter describes the proposed modeling framework components of path retention and traffic operation profiles, emissions modeling using MOVES-Matrix, dispersion modeling using AERMOD, and exposure modeling based on household pairing and demographic characteristics.

2.1 Overview of Modeling Framework

The proposed modeling framework consists of several modeling components (stages) as shown in Figure 1, started by retrieving the Atlanta Regional Commission's (ARC) Activity-Based Travel Demand Model 2020 (ABM2020-TIP-A1) output data set (more than 20 million trips), which are the same model outputs used to develop Atlanta's Transportation Improvement Program (TIP). A path-retention algorithm (Zhao, et al., 2019) was implemented for this model run to retain each new model-predicted vehicle travel path (and number of vehicles assigned to each new path), between each origin-destination pair, as derived in each step of the trip distribution process (Frank-Wolfe algorithms). The ABM-predicted trips between O-D pairs were then allocated to these routes in proportion to model predictions for randomized departure times within the modeled half-hour departure window. The second-by-second positions for each vehicle are estimated using the retained path (set of links traversed) and the speed for that time period on each traversed link. The model output also retains all basic information about the synthetic households used in trip generation and about the individuals assigned to each trip, facilitating further demographic analyses of the travel.

Near-road pollutant concentrations are predicted for the entire region by integrating vehicle activity with MOVES-Matrix (Guensler et al., 2017; Liu et al., 2019) and AERMOD (Kim et al., 2019) with a very large variably spaced receptor array in Georgia Tech's PACE distributed computing cluster. The PACE cluster allows the process of more than 500 dispersion modeling jobs simultaneously. The streamlined modeling system also incorporates a number of innovative features that provide huge reduction in computational resources and processing time for fine-resolution air quality impact analyses (Guensler, et al., 2021). To further accelerate modeling speed, a supervised link screening algorithm identifies and removes all irrelevant roadway sources-receptor pairs with high precision that barely affect the pollutant concentrations of each receptor (Kim, et al., 2022a).

Exposure modeling is conducted by tracing the traveled paths of modeled individuals through the transportation network to account for their time on the road and time at each destination. Over-the-road exposure and in-vehicle exposure were not directly assessed in this study. Receptors placed directly over the roadway were removed from the analysis, based on EPA guidance, and exposures were assessed using data from roadway-adjacent receptors. Modeling the differences between road-adjacent exposure and in-vehicle exposure (typically a lower concentration in the cabin) requires integration of additional processes that consider factors such as HVAC system performance, air exchange rates, and vehicle-specific dynamics. By aggregating the outputs into household/demographic group levels, the modeling framework supports equity analysis with respect to traffic-related pollutant emission/inhalation across the demographic groups.

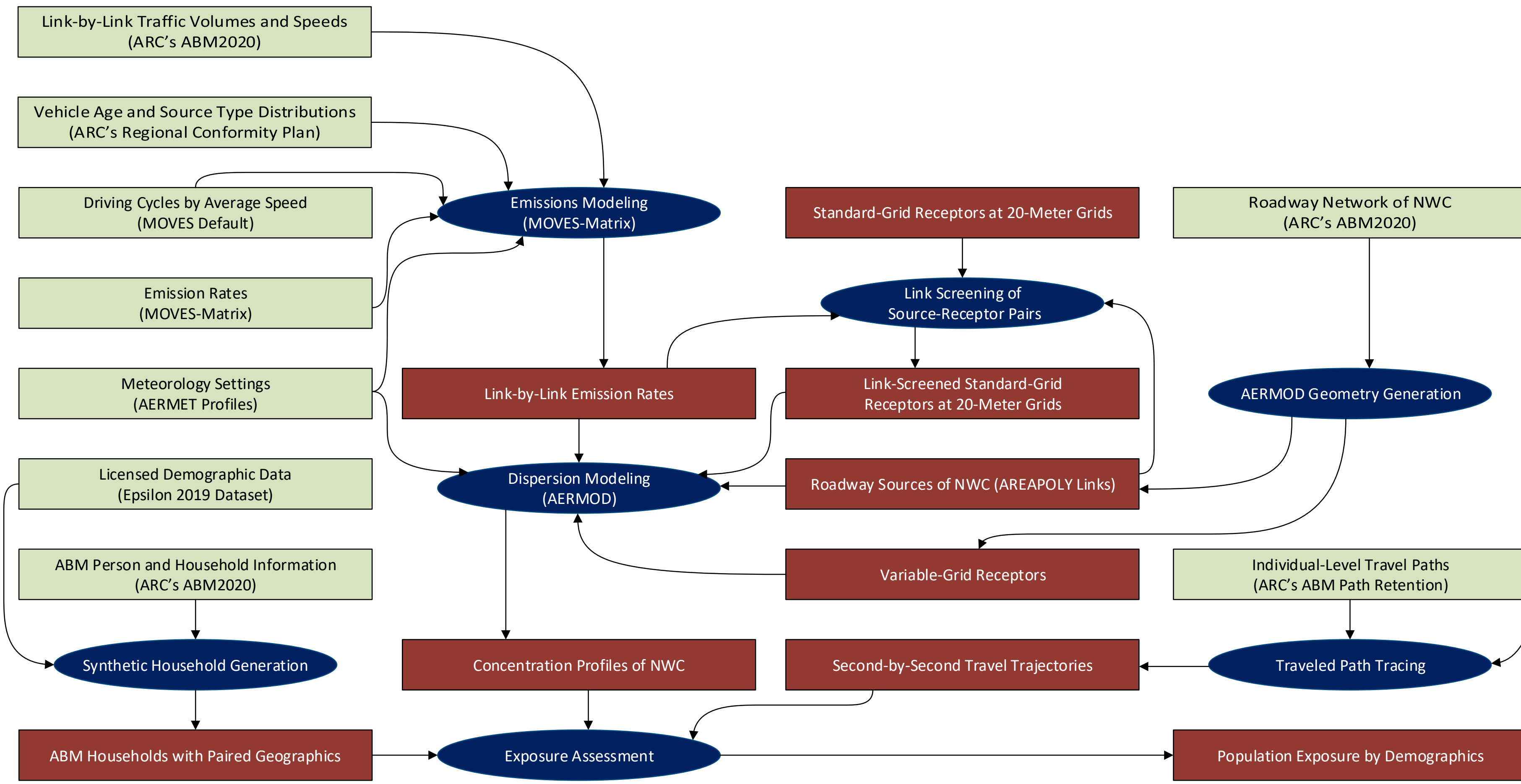

Figure 1. Flow Chart of the Proposed Modeling Framework of Population Exposure

2.2 Modeling Subarea: Northwest Corridor Description

The research team performed AERMOD microscale dispersion modeling for the entire Atlanta I-75/I-575 Northwest Corridor (NWC) subarea, including freeway corridors, managed lanes, connecting arterials, and intersections serving the NWC system (as shown in Figure 2). This corridor and subarea have been the subject of extensive emissions and dispersion modeling efforts by the Georgia Tech research team (Guensler, et al., 2021; Kim et al., 2020a) and encompasses a variety of projects of potential policy concern, including major intersections, managed lanes, and direct access ramps. The NWC also contains roadway sections with noise barriers, which can be modeled with the AERMOD RLINEXT source type.

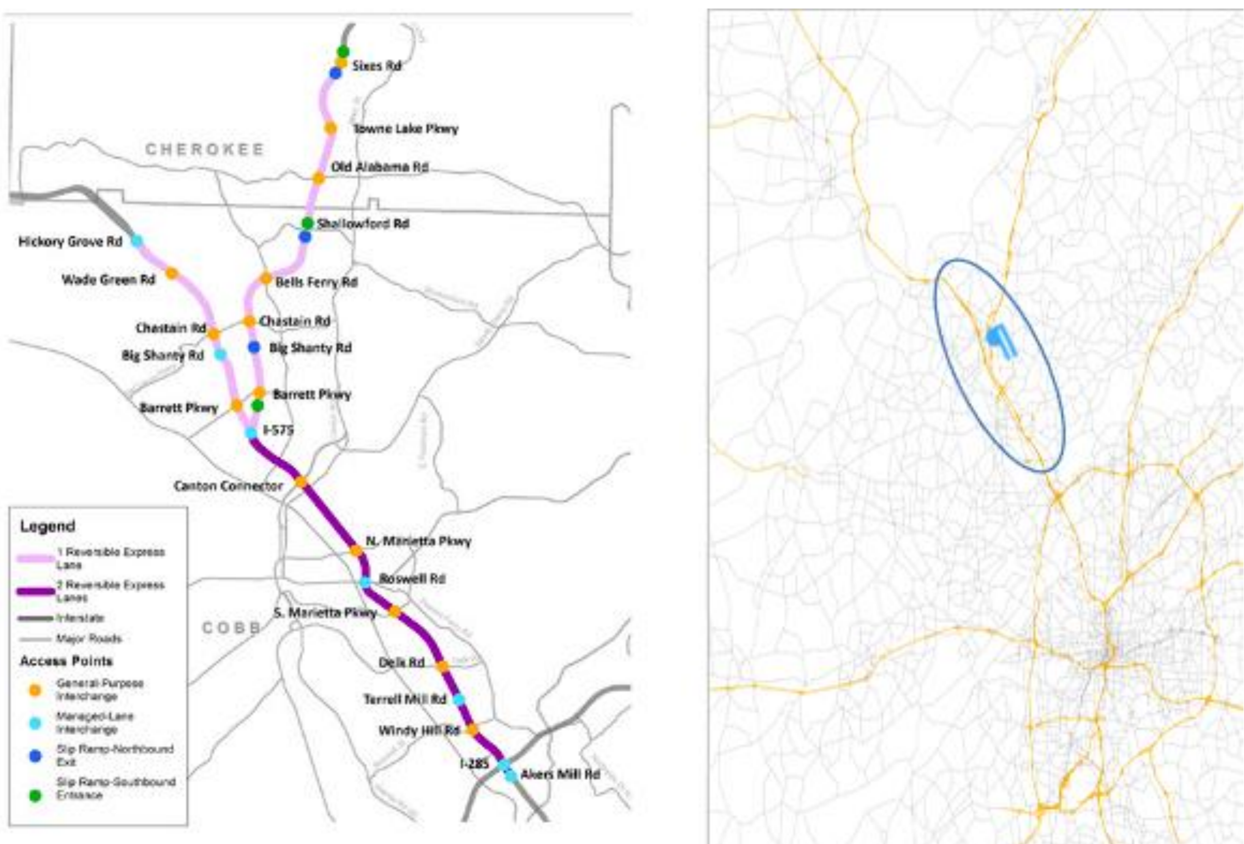

Figure 2. I-75/I-575 Managed Lanes Corridor and the Modeled Sub Area

2.3 Traffic Operation Profiles: Activity-Based Model

The roadway network adopted in this study is the Activity-Based Travel Demand Model (ABM) by Atlanta Regional Commission (ARC). The specific version employed is ABM2020-TIPA1-2020, which is same version run by ARC staff for the Transportation Improvement Program (TIP), for calendar year 2020, with the path retention method applied and modeled with the ARC planning assumptions for the transportation network, land use, and household demographics for calendar year 2020 (Zhao, et al., 2020; Zhao 2020). The path retention method allows modelers to retain the paths between origin and destination predicted by the ABM's internal

Franke-Wolfe algorithms. With path retention, model-predicted link-by-link vehicle traverses through the road network are available for analysis. The ABM models trips in half-hour bins, so a randomization algorithm was designed and implemented to generate the trip departure timestamps (described in Section 2.6.1 in this report). Therefore, both a full trip record with the departure time and arrival time and origin and destination information and its corresponding specific traces at each time and location are available for energy and emissions modeling. The ABM2020-TIPA1-2020 model data is available in geodatabase format and includes a links layer and a nodes layer available. In total, the model network includes 149,967 links and 66,418 unique nodes. The NWC freeway network (and connecting arterial segments) includes 1,570 of the regional roadway links.

For vehicle activity, the team used the Atlanta Regional Commission's ABM2020 model outputs with path retention, which provides modeled travel demand traffic volumes, link speeds, and individual predicted paths through the network (second-by-second positions are derived from departure time and average speed on each link traversed by each vehicle, as described in Section 2.6.2). A total of 11,399 model-predicted trips that pass through a screen line at Roswell Road and I-75 for one hour of the morning peak period. Because the ABM retains household demographic data and person assignment for each trip, analysts can apportion modeling results across any demographic cluster derived from ABM model cut points. The team can also generate fleet composition (make, model, and age distributions) for those trips touching freeway links using 2019 freeway license plate observations. In regional conformity modeling efforts, the research team typically generates fleet composition for arterial trips using random pulls from trip origin traffic analysis zone (TAZ) vehicle registration mix employed. However, for this project, it was important that the analyses across time and space employ the same fleet composition and model year distributions in all analyses for control purposes. To that end, the team used the fleet composition data employed by the MPO in preparing their regional conformity analyses (discussed later in the report). Python scripts have been used to generate AERMOD input files for subsets of sources (link representations) and receptors, to run AERMOD on the PACE cluster, and to retain and compile results for each receptor in space and time.

2.3.1 Link-by-Link Volumes and Speeds by Hour

The ABM2020-TIPA1-2020 model generates hourly traffic volume and speed, the values remain static within each of the five time periods: EA (early AM), AM (AM peak), MD (midday), PM (PM peak), EV (late evening/night). The EA period ranges from 3:00 AM to 6:00 AM, the AM period ranges from 6:00 AM to 10:00 AM, the MD period ranges from 10:00 AM to 3:00 PM, the PM period ranges from 3:00 PM to 7:00 PM and the EV period ranges from 7:00 PM to 3:00 AM. Specifically, within each of the five time-periods, the ABM processes trips in half-hour increments. More details regarding the ABM2020 model can be found in the online model documentation (ARC 2019).

2.3.2 Fleet Composition

The input fleet composition to emissions modeling were extracted from the previous research by the team (Xu, et al., 2018) to provide source type distributions and age distributions to MOVES-Matrix. In support of the regional conformity plan, the 20-county nonattainment area was divided into 13-county vs. 7-county areas, where separate fleet compositions were applied for each area in MOVES modeling (ARC, 2015). The distribution of counties is shown in Table 1. Cobb County and Cherokee County belong to the 13-county area, and Bartow County belongs to the 7-county area. Four sets of fleet composition were used in support of the regional conformity plan for the year of 2017, 2024, 2030, and 2040, respectively. The fleet composition for calendar year 2017 was used in this research. The source type distribution is shown in Table 2, and the corresponding fleet composition for each modeled ABM link was allocated based upon county group membership.

Table 1. County Division of Regional Conformity Plan (ARC, 2015)

County Name	Area
Fulton	13-county
DeKalb	13-county
Cobb	13-county
Gwinnett	13-county
Rockdale	13-county
Henry	13-county
Clayton	13-county
Fayette	13-county
Douglas	13-county
Cherokee	13-county
Coweta	13-county
Forsyth	13-county
Paulding	13-county
Bartow	7-county
Carroll	7-county
Spalding	7-county
Newton	7-county
Walton	7-county
Barrow	7-county
Hall	7-county

Table 2. Input Source Type Distributions

Source Type #	Distribution in 13-County Area	Distribution in 7-County Area
11	2.11%	2.84%
21	53.91%	47.22%
31	31.00%	35.32%
32	10.12%	11.55%
41	0.03%	0.01%
42	0.02%	0.01%
43	0.33%	0.32%
51	0.04%	0.03%
52	1.25%	1.23%
53	0.09%	0.09%
54	0.13%	0.15%
61	0.62%	0.53%
62	0.35%	0.70%

2.4 Emissions Modeling: MOVES-Matrix

MOVES-Matrix for MOVES 2014b is a lookup-based energy-use emission rate modeling system developed by the team by pre-running MOVES 2014b in an iterative fashion on the PACE supercomputing cluster across combinations of all model inputs that affect emission rate outputs (Liu, et al., 2019). To develop the set of emission rates for the Atlanta metro area, the team ran MOVES 146,853 times, iterating across 21 calendar years (2010-2025, 2030, 2035, 2040, 2045, and 2050) × 3 fuel specifications (summer, winter and transition) x 111 temperature bins × 21 humidity bins. Each iteration produces emission rates for energy use and all pollutants (including pollutant composition types), for every vehicle source type, model year, and on-road operating conditions (by MOVES VSP bin and by average speed and facility type) for each calendar year, fuel specification month, temperature bin, and relative humidity bin. Because MOVES model outputs are also dependent on the specifications of the regional inspection and maintenance (I/M) program strategy (by calendar year), the model runs described above apply to Atlanta and any other region that employs the same fuel specification and I/M programs. Hence, the team is in the process of preparing emissions rate matrices for each of the 117 unique combinations of MOVES 2014b fuel and I/M programs across the United States. As of July 2021, the research team had generated full MOVES-Matrix outputs for 30 of these 117 modeling regions, covering 2,902 out of 3,228 counties as shown in Figure 3 (California does not use the MOVES model).

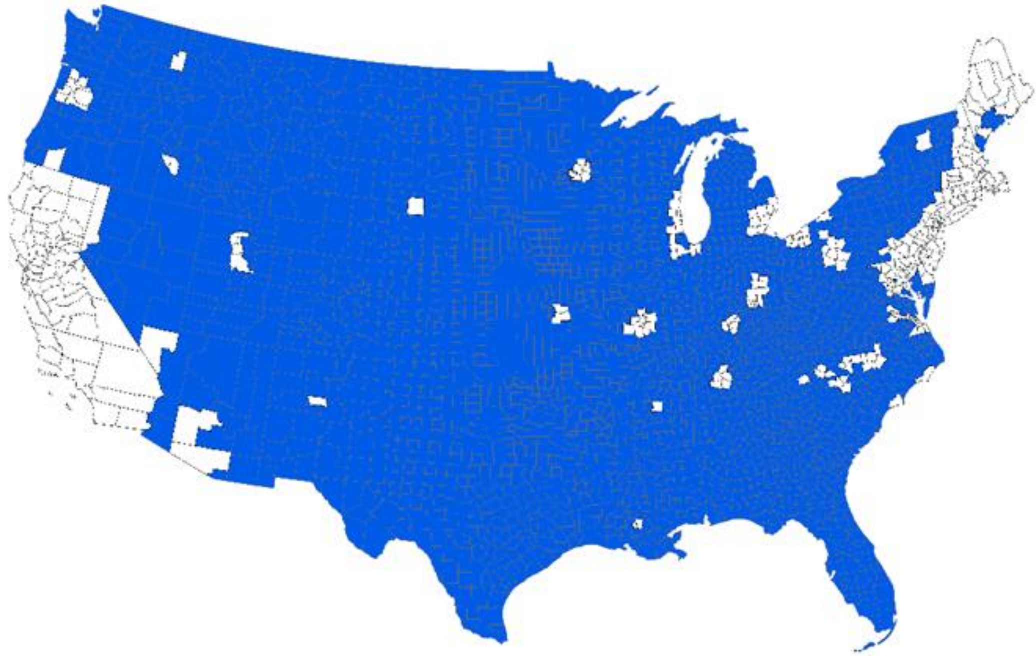

Figure 3. MOVES-Matrix Coverage Area

By pre-processing all possible combination of input variables, users can query the emission rates directly from MOVES-Matrix, rather than performing new MOVES modeling runs. Users need only identify the subset of the MOVES-Matrix needed for an analysis (calendar year, fuel month, and meteorology data) and run a set of queries to pull required emission rates (by vehicle type and model year and on-road activity) and the queries reassemble the fleet emission rate using the exact same weighting process that is used in MOVES. The queries run 200x faster than running MOVES and users never have to generate MOVES scenario input files, which improves analytical run time efficiency (Liu et al. 2019). The conceptual flow of MOVES-Matrix processing is illustrated in Figure 4.

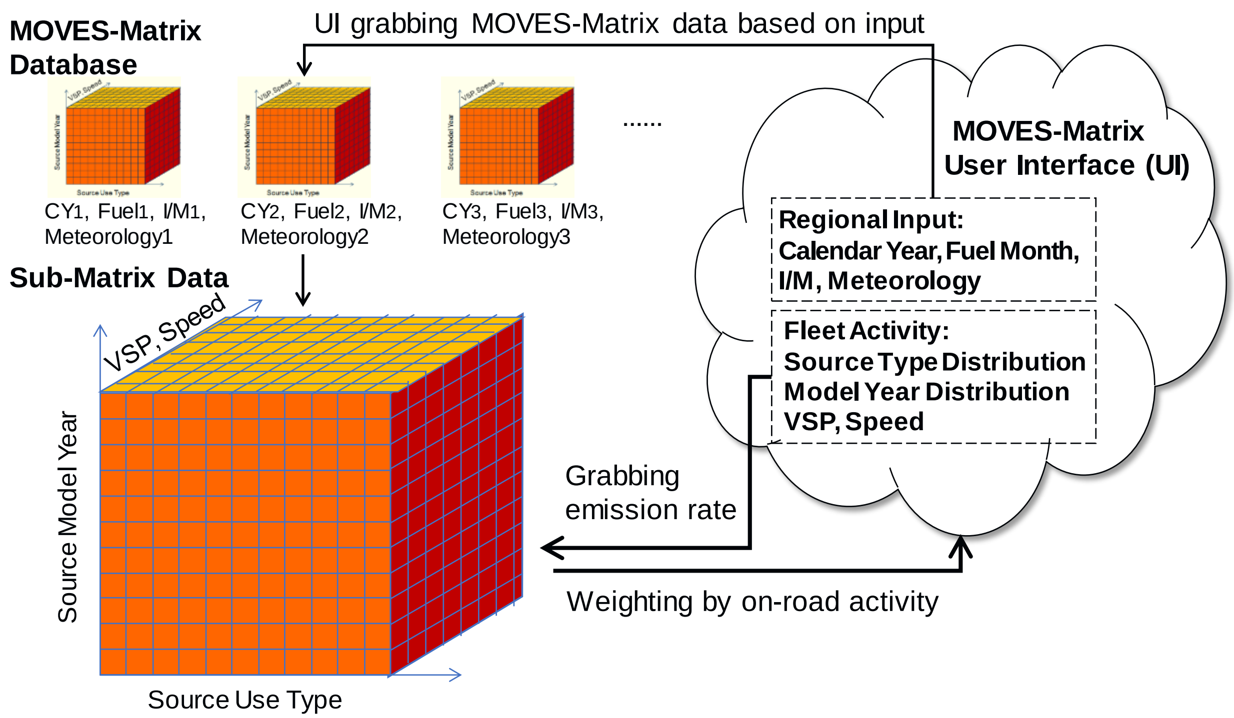

Figure 4. MOVES-Matrix Conceptual Flow (Liu et al., 2019)

MOVES-Matrix also enables rapid analyses of engine starts, truck hoteling, evaporative sources, brake/tire wear (Xu, et al., 2017), and can be used to model the emissions from individual vehicles (Guensler, et al., 2017). MOVES-Matrix can be easily coupled with vehicle activity analysis (Li, et al., 2017, 2016; Xu, et al., 2017) by importing second-by-second vehicle operations. MOVES-Matrix can also be applied to different transportation models, such as travel demand models (Xu, et al., 2018), and microscopic traffic simulation models (Xu, et al., 2016).

MOVES-Matrix is highly-desirable for regional-scale dispersion analysis (Kim et al., 2020b), with high-performance to deal with links from large-scale networks, variations in meteorology, and traffic operation input, and with its user-friendly nature to minimize potential human error in running MOVES (especially when it comes increased number of input links).

2.4.1 MOVES-Matrix Emission Rates

As discussed in the previous section, MOVES-Matrix provides a 90-billion cell look-up matrix for each modeling region. For the Atlanta metro area, MOVES-Matrix contains sub-matrices based on combinations of calendar year, season (Spring/Fall, Summer, Winter fuel season), temperature (0º-110º F with 1º F-bin interval, 111 bins in total for Atlanta), and relative humidity (0%-100% with 5%-bin interval, 21 bins in total for Atlanta). Meteorological data from AERMET are rounded to the appropriate temperature and humidity to link with appropriate sub-matrices for each MOVES-Matrix run.

For each hour in the year of 2019, corresponding sub-matrices within MOVES-Matrix were extracted using Python-based scripts for use in project-level emission rate calculations. Even though the hourly traffic volumes and average link speeds from ABM2020 are consistent across analysis days, and the distributions of vehicle source types and model year are uniform through the year and across analyses, the resulting emissions rates for each hour are different as a function of hourly environmental conditions.

2.4.2 Model Inputs and Outputs

Various sources of input data were used to pull hourly emission rate data from MOVES-Matrix for each analysis, as described in a previous section:

- Speed and volumes are derived from the ARC's ABM2020
- Driving cycles are embedded in MOVES for average speed and facility type
- Source type and vehicle age distributions by source type from regional conformity analyses provided by Georgia EPD
- AERMET meteorology data (hourly for the modeling year) from the regional air quality conformity analyses

Because the sub-matrices needed for the analyses are large (and require a lot of time to upload to PACE), and because MOVES-Matrix is so efficient, the emission rate modeling process was performed on local computers and results were used in PACE. It took less than 12 hours to model all ABM links (1,191 links) in MOVES-Matrix for NWC, for the 8,760 hours (24 hours × 365 days). The emission rate outputs were compiled and uploaded to PACE to provide the input emissions to dispersion modeling.

2.4.3 Meteorological Configuration: AERMET

A total of 8,760 hours of AERMET profiles (24 hours × 365 days) were obtained from Georgia EPD station #KATL/KFFC, which covers Cobb County, Cherokee County, and Bartow County (the counties in which the study area roads are located), for each hour of 2019, to provide the meteorology input (temperature, relative humidity, wind direction, wind speed, etc.) for emissions and for NWC dispersion modeling. The AERMET files (PFL file and SFC file) serve as an input to AERMOD and the hour-by-hour temperature and relative humidity profiles were used to extract and assign corresponding emission rates from the MOVES-Matrix (for MOVES 2014b) database to each individual hourly model run. The wind rose for the used AERMET profiles is shown in Figure 5 (prepared using OriginPro™ software). The temperature and humidity distributions are shown in Figure 6 and Figure 7, respectively.

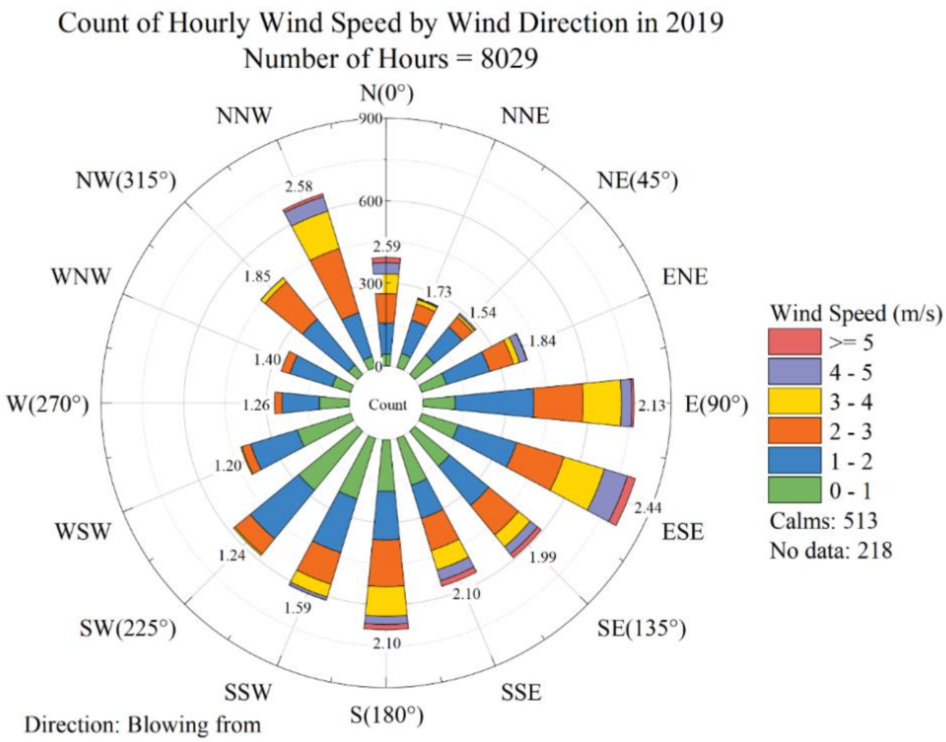

Figure 5. Wind Direction and Wind Speed Distributions

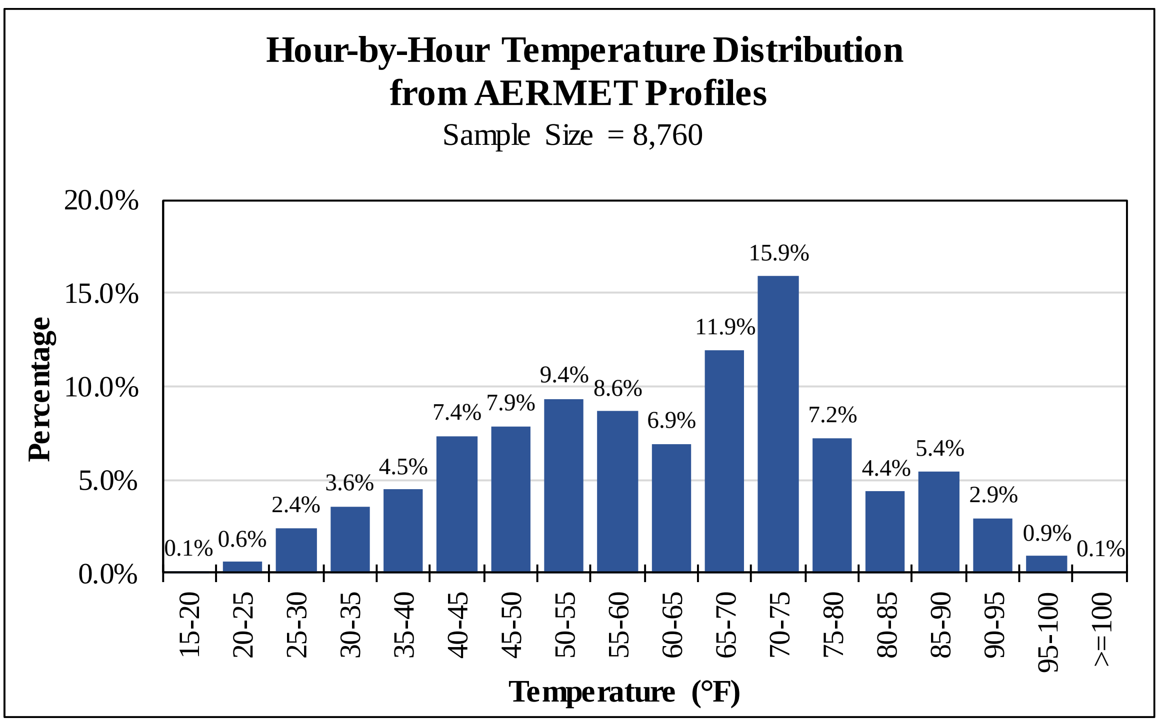

Figure 6. Temperature Distribution from AERMET Profiles

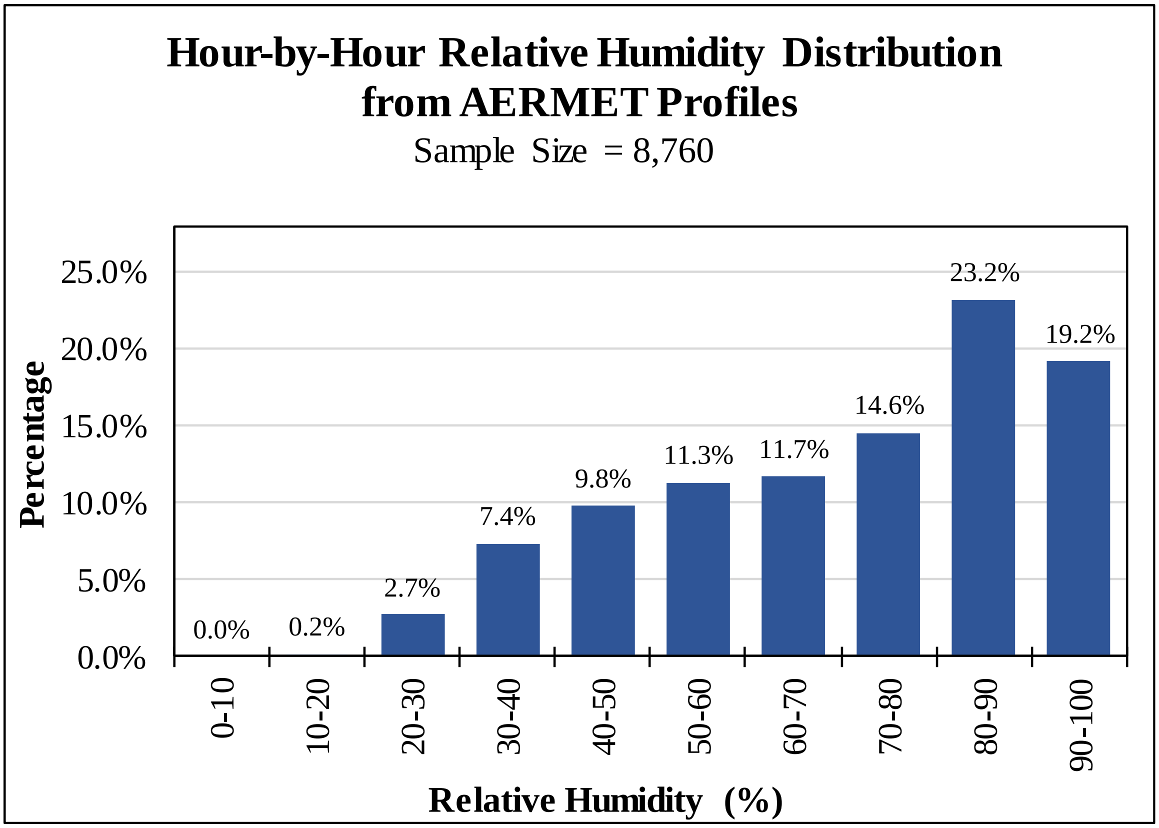

Figure 7. Relative Humidity Distribution from AERMET Profiles

2.5 Dispersion Modeling: AERMOD

EPA's regulatory microscale dispersion model of AERMOD (version v21112) was employed in this project to predict the hour-by-hour traffic-related $PM_{2.5}$ concentration profiles for the calendar year of 2019, based on the roadway network of the NWC subarea from ARC's ABM2020, traffic volumes and speeds predicted by the ABM in half-hour increments, the emission rates by link from MOVES-Matrix, and receptors of both standard grid of 20 meters (link-screened) and variable grid. More details on the manual and automatic creation of polygons, source validation, receptor placement and allocation of emission rates can be found in the FHWA report (Guensler et al., 2021).

2.5.1 Assignment of AERMOD Emissions Rates

The MOVES-Matrix output provided hour-by-hour (24 hours × 365 days) emissions rates in (grams/link/hour) for each ABM link and the same emissions rates by ABM link were used as the calculation starting point for emissions rate inputs to AERMOD across all AERMOD source types. These emissions rates were assigned to the corresponding AERMOD source types and were converted to the correct units for each AERMOD source type, as shown in Table 3. For LINE, AREAPOLY, and RLINE sources, if a link was divided into multiple sources, emissions per link were proportionally assigned to each source based on segment length (Guensler, et al., 2021). For this analysis the AERAPOLY source type was used.

Table 3. AERMOD Unit Conversion from MOVES-Matrix Output

Source Type	Required Unit	Calculation from grams/link/hour
LINE	grams/meter^2/second	Divided by source area and 3,600
AREAPOLY	grams/meter^2/second	Divided by source area and 3,600
VOLUME	grams/second	Divided by number of spheres and 3,600
RLINE	grams/meter^2/second	Divided by source area and 3,600
RLINEXT	grams/meter/second	Divided by source length and 3,600

2.5.2 Receptor Placement

AERMOD allows users to specify receptor locations. The receptors define the physical locations in x, y, z space for which pollutant concentrations will be predicted for every hour in the simulated year. Receptors allow users to assess pollutant concentration levels relative to nearby locations of concern (e.g., near schools or residential areas where individuals are likely to be exposed to pollutant concentrations for more than one hour) and to identify localized areas of high concentration (hot spots). Assessment of receptor concentrations allows modelers to identify regions that may exceed the NAAQS. The computing resources available for this project allowed the research team to assess as many receptors as desired, so a variety of receptor patterns were used in this study, including standard receptor grids and variable receptor grids.

Standard Grids

Standard grids with 200-meter spacing between receptors (Figure 8), 20-meter spacing between receptors, and 5-meter spacing between receptors were used in this study. Receptor grids provide a simple approach that requires minimal forethought, but they are computationally inefficient because many of the receptors are placed so far from the roadway that pollutant concentrations are not significant and contribute little information. Likewise, increased receptor density is desired near the roadway, where variation in pollutant concentration is likely to be the largest. It is difficult to strike a balance between sufficient receptor density near roadways and computational efficiency while using a standard grid. Hence, a variable grid approach is also useful as will be discussed in the next section.

Given the large size of the study area, receptors far from roadway emission sources yield very low receptor concentrations (e.g., less than 0.1 μg/m^3 across all input emissions rates and meteorology by hour). Removing such receptors from the analyses does not affect research outcomes. A link-screening tool is implemented to examine all link-receptor pairs and filter "low-impact" receptors to reduce required computational resources.

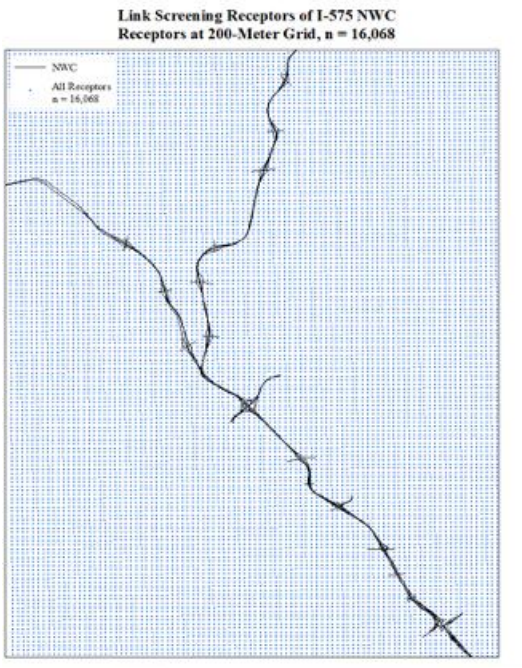

Figure 8. Placement of Receptors of 200-Meter Standard Grid

Variable Grids

A more advanced approach for placing the receptors is the dynamic grid method (Kim, 2020). A link screening process is applied before setting the variable grid (Kim 2020a), based upon source-receptor centration statistical significance assessed via machine learning. Specifically, a stepwise process is conducted to identify optimal receptor locations, with a forward search used to find the receptor that best fits the $PM_{2.5}$ concentration profile, and a backward search used to remove the receptor that is the worst fit for the $PM_{2.5}$ concentration profile. The iteration process continues, until the marginal change in mean squared error (MSE) is less than some pre-set critical threshold. The dynamic grid-receptor model identifies optimal receptor locations, removing receptors from the grid that contribute no significant additional information to the spatial concentration distributions. The variable grids that evolve from the process tend to have higher receptor density near roadway sources (where small differences in distance between receptors yield high differences in concentrations) and low receptor density further from roadway sources (where additional receptors yield very similar low concentrations). The more refined receptor grid minimizes the number of receptors used in modeling and speeds up the distributed modeling process.

Link Screening

The link screening tool previously developed by the team (Kim et al., 2020a) was implemented to eliminate roadway source-receptor pairs that do not contribute significantly to predicted downwind concentrations at a receptor. When no link-receptor pairs were significant for a receptor, the receptor was eliminated from all scenario analyses (reducing the number of receptors required to represent downwind concentrations). The machine learning-based screening tool was developed with the support of PACE (Kim et al., 2020b) to examine all source-receptor pairs across all input combinations (emissions rates, wind direction, wind speed, temperature, relative humidity, etc.) to filter receptors with summed predicted concentration (summed from all sources) of less than 0.1 µg/m^3 under all circumstances. These receptors can then be safely excluded from the analyses without any significant impact on the results.

An iterative filtering process is implemented to reduce the number of source-receptor pairs that need to be screened, further improving the efficiency of the link-screening tool. In the first iteration, receptors on a 200-meter grid (coarse grid to reduce computation burden) enter the screen, and “sensitive” receptors (with results larger than 0.1 µg/m^3 under any circumstance) are selected (i.e., these receptors are used in the dispersion modeling). The significant receptors naturally form a buffer area in which all receptors on a finer grid (20-meter and 5-meter grids) are selected. In the second iteration, the 20-meter grid receptors immediately outside the border of this area (that is, between a selected receptor and the adjacent screened receptor 200 meters further out) are examined. The 20-meter grid receptors are integrated in the next iteration, to expand the selected area (as a result, 5-meter receptors within the expansion are also selected). In the third iteration, the 20-meter grid receptors immediately outside the expanded border are screened, and the final selection is provided as a union of the selected receptors from all three iterations (Figure 9).

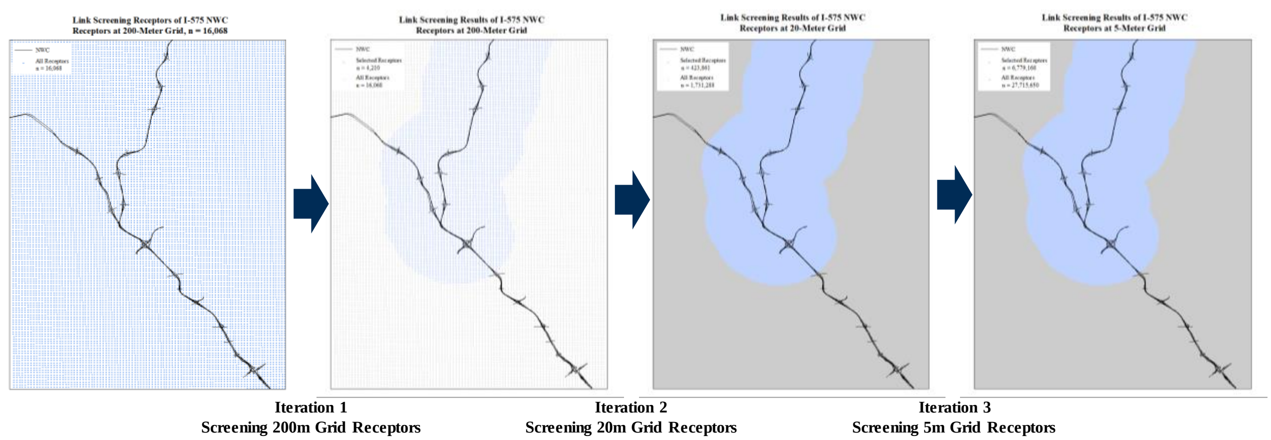

Figure 9. Iterative Implementation of Link-Screening Tool for I-575 NWC

The iterative implementation of the link-screening tool filtered approximately 75% of the receptors as being low-impact receptors and having little influence on the model output, and the total run time for screening was less than two hours (supercomputer not required), as shown in Table 4. The team recommends the use of the link-screening tool in future research projects that involve large-scale networks and fine-grid receptors for dispersion modeling.

Table 4. Performance of Link-Screening Tool

Network	Grid Resolution (Meter)	Input Number of Receptors	Selected Number of Receptors	Percentage of Receptors Excluded (Workload Reduction)	Screening Running Time (Minute)
I-575 NWC	5	27,715,651	6,779,168	75.5%	80
I-575 NWC	20	1,731,288	423,861	75.5%	5

2.6 Travel Paths: Path Retention of ABM

It is important to feed high-resolution travel information (second-by-second trajectory data in this assessment) into population exposure assessment. The trip output from ARC's ABM was allocated to departure periods of half-an-hour intervals (48 intervals for a day). The research team designed and implemented a randomization algorithm to generate the departure time (minute level) by trip from the cubic-splined departure time distributions by trip purpose, with the trip chains accounted for by making sure early trips finish before following trips. The travel paths were converted to second-by-second trajectories by traversing roadway links for the model-predicted link-by-link average speed. Travel trajectories by individual vehicle were then overlaid with the predicted concentration profiles for each hour to assess the population exposure during the year of 2019.

2.6.1 Trip Departure Timestamp Generation

Each trip predicted by ABM within a half-hour bin was assigned a departure time on a one-minute basis. The ABM-predicted trip information includes the trip departure time by period (at half-an-hour interval), as well as travel time (in minute), travel distance (in miles), origin terminal time (the time it takes until the individual enters the network, in minute, e.g., from the apartment to the roadway), destination terminal time (the time it takes for the individual to leave the roadway network until the individual arrives at the destination, in minute, e.g., from the roadway to a restaurant), and trip purpose (e.g., commute trips, recreational trips, etc.). The trip purposes are listed in Table 5 below.

Table 5. Trip Purposes from ARC's ABM

Abbreviation Code in ABM	Trip Purpose
atwork_business	Business trips at work
atwork_eat	Trips for eating at work
atwork_maint	Maintenance trips at work
eatout	Trips to eat out
escort_kids	Escort trips with kid(s)
escort_no kids	Escort trips with no kids
othdiscr	Other discretionary trips
othmaint	Other maintenance trips
school_drive	School drive
school_predrive	Pre-school drive
shopping	Shopping trips
social	Trips for social
university	University trips
work_bluecollar	Commute trips (blue collar jobs)
work_health	Commute trips (health-related jobs)
work_retailandfood	Commute trips (retail and food)
work_services	Commute trips (services)
work_whitecollar	Commute trips (white collar)

For a fine-level modeling framework that involves various models, it is important that each modeling process utilizes high-resolution input and output. Hence, the minute-by-minute trip departure timestamp is needed before the travel paths can be converted to second-by-second trajectories (trip travel time is already at minute-interval). Intuitively, the departure times follow a distribution across time of day due to patterned travel demand (e.g., morning peak hour for commute trips), and the departure time distribution needs to be modeled separately by trip purpose (e.g., commuting to work is most likely to be distributed in the morning peak hour, while commuting from work to home occurs mostly in the evening peak). The team analyzed the probability density functions of trip departure times by trip purpose by adopting natural cubic spline interpolations, and the team designed and implemented a randomization algorithm to generate trip departure timestamp at minute level.

Trips are generated and allocated to trip chains (tours) in the ABM output, where the departure of next trip cannot precede the original trip, and the trip chain constraints were accounted by making sure that trips within the same chains occur in correct sequences. Every trip in a trip chain is essentially constrained in two ways (forward and backward): 1) a new trip cannot start until the last trip is finished (a trip is finished at the timestamp that equals to the departure timestamp plus its travel time), which sets the early bound of the departure timestamp (forward checking), and 2) a trip must finish before a particular timestamp so that the next trip can start at its allocated period (half-an-hour slot), which sets the late bound of the departure timestamp (backward checking). All trip chains were examined to generate the early bounds and late bounds for each trip, and every trip was confined to departure only in the legitimate period between the early and late bounds. More detailed descriptions of the iterative process are in Appendix A, with the sample of legitimate periods of a dummy chain of seven trips.

The legitimate departure periods by trip were aggregated by trip purpose to generate the raw probability distributions, and a natural cubic spline interpolation was performed to generate the probability density functions by trip purpose, as shown in Figure 10.

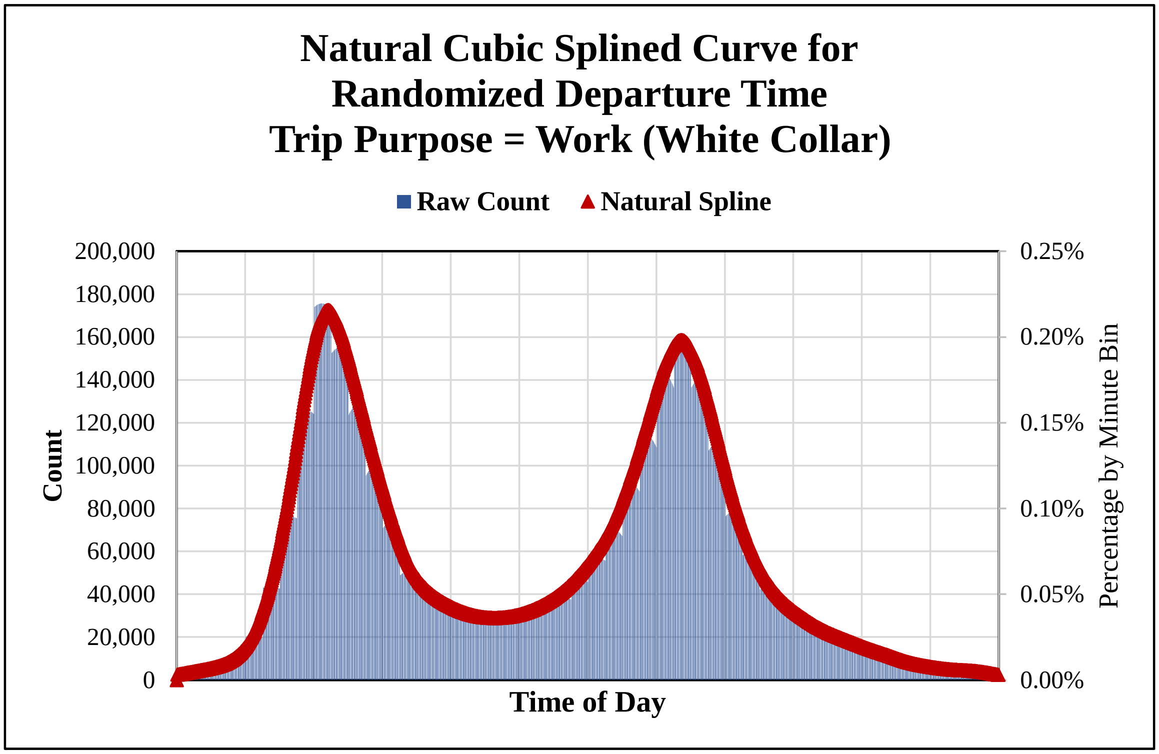

Figure 10. Splined Probability Density Function Generation

The legitimate departure period only ensures that the next trip can start and finish within the assigned half-an-hour slot (enough time left) if a reasonable departure time was assigned to the next trip. This does not guarantee that the next trip starts after the original trip finishes (legitimate periods could overlap with each other); hence, a sequence check is still needed to make sure all trips follow the correct order of the chain. For every trip chain, the randomized departure timestamps for all trips were generated at one time following the splined probability density functions, and if the generated trips do not follow the abovementioned sequence rules, another draw was performed with all timestamps re-generated until the constraints were met. The generated distributions of departure timestamps were compared with the splined probability density functions as a verification of the randomization, as shown in Figure 11 through Figure 14, and the generated distributions fit well with the splined distributions.

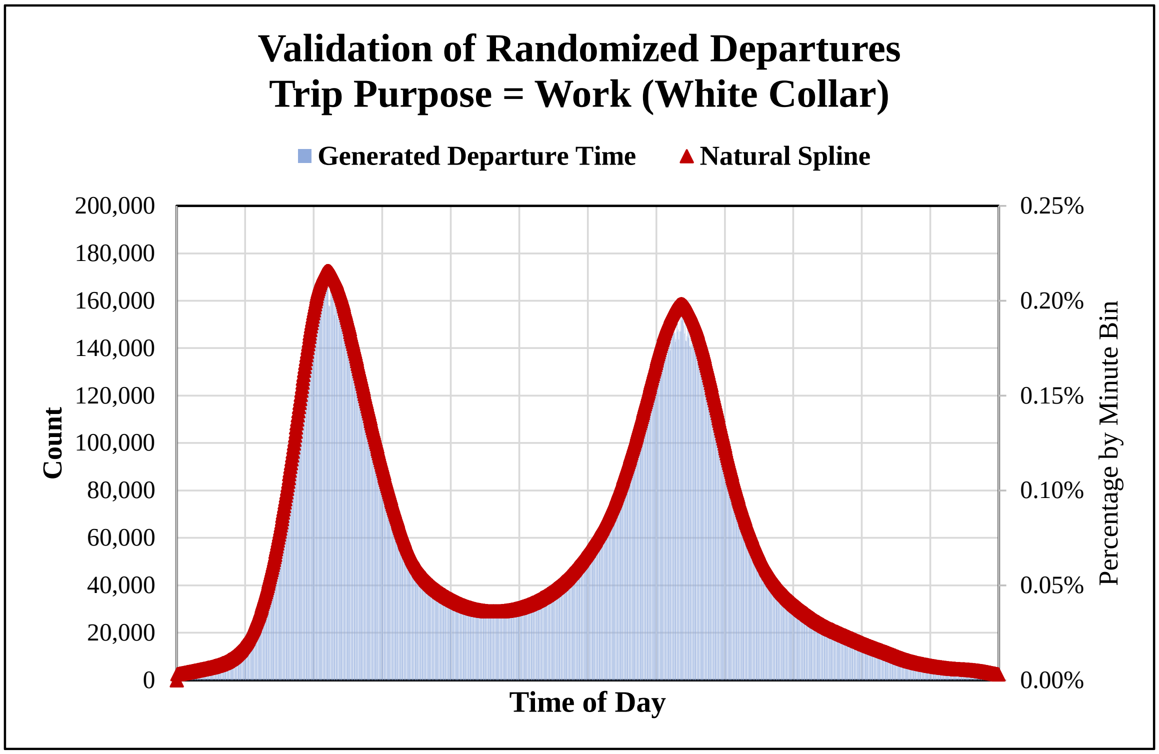

Figure 11. Validated Departure Timestamp Distributions for Work (White Collar)

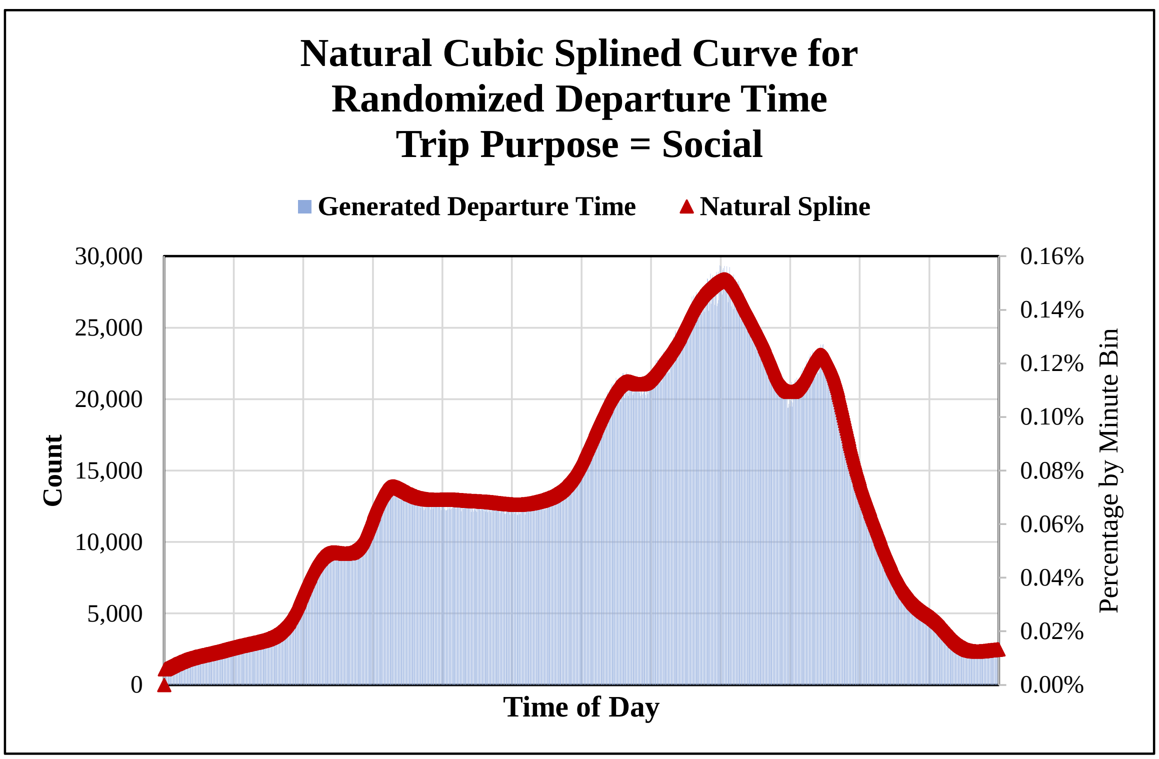

Figure 12. Validated Departure Timestamp Distributions for Social

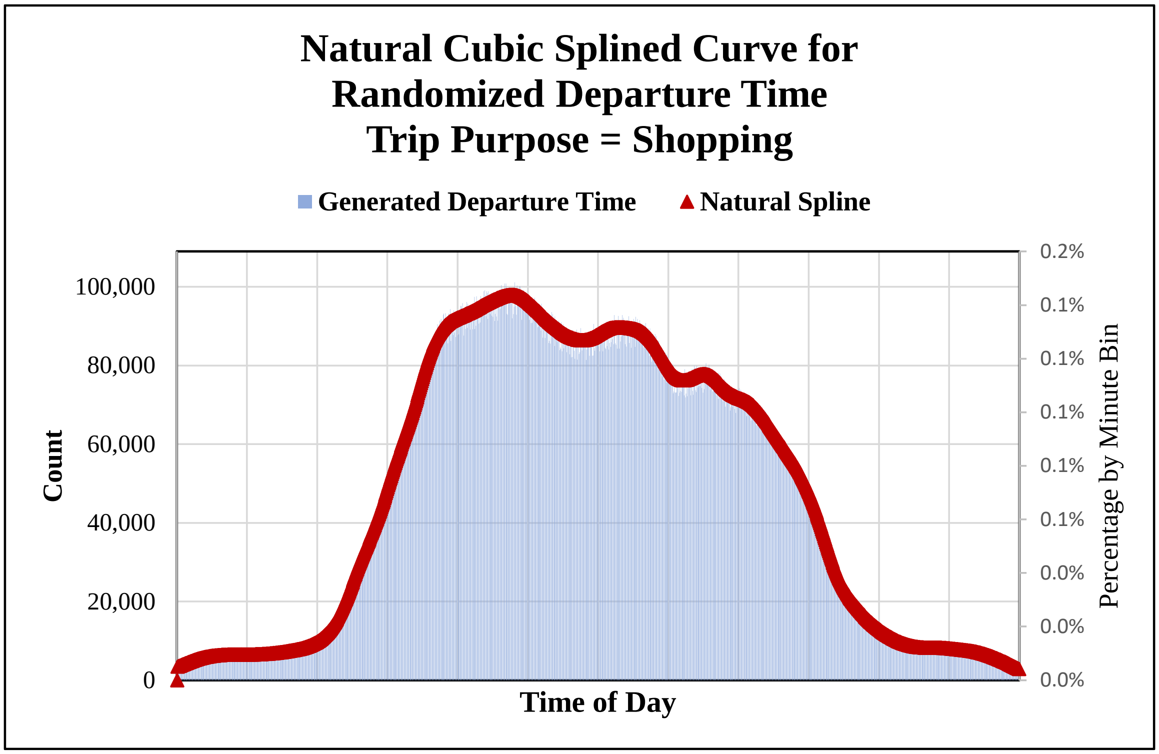

Figure 13. Validated Departure Timestamp Distributions for Shopping

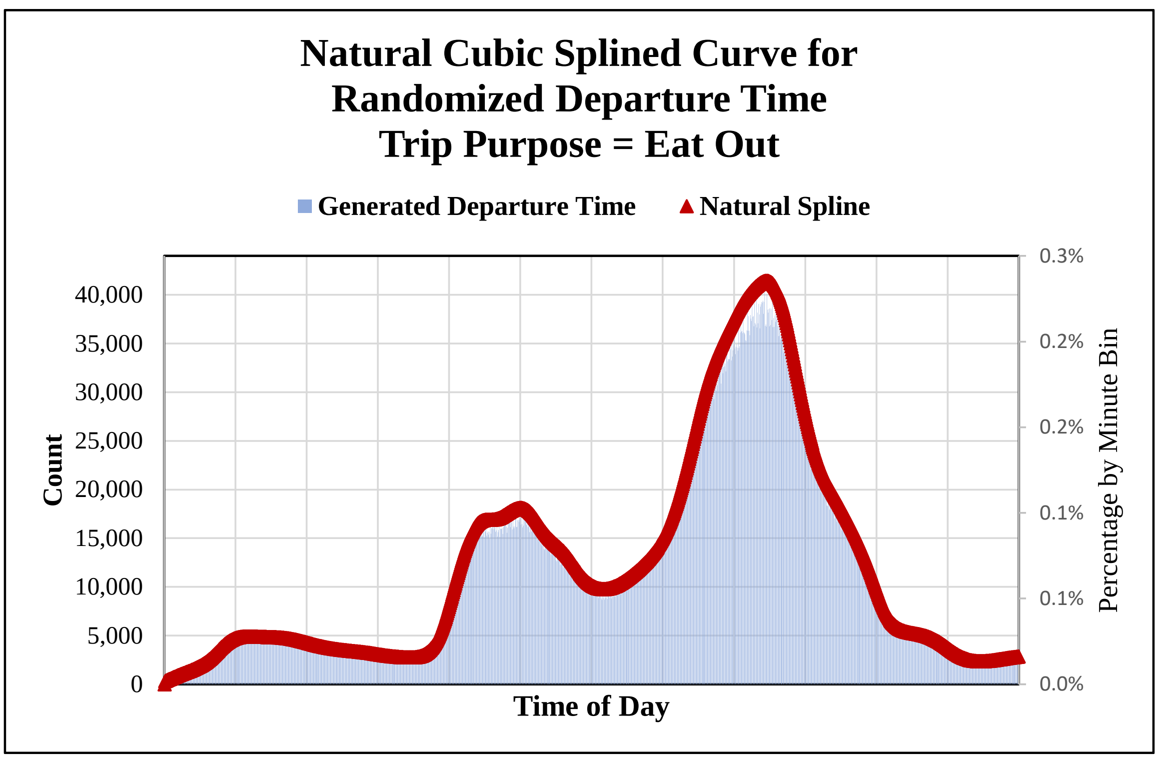

Figure 14. Validated Departure Timestamp Distributions for Eating Out

2.6.2 Second-by-Second Travel Trajectories

Travel trajectories (travel traces on the roadway network) were then generated using Geographic Information System techniques based on the travel paths (path retention) and trip departure time (generated in Section 2.5.2). The at-home locations were generated using a synthetic household generation algorithm by pairing the ABM household locations (TAZ-level) vs. the licensed demographic dataset of Epsilon (longitude and latitude information). The on-road travel trajectories were generated by traversing of the roadway links and assuming no speed variability within each link (same average speed used across all seconds of activity on that link). In this study, only the households that live in the studied subarea of NWC were accounted for (filtered based on their household geographic information). The team is expanding the study to the metro Atlanta area as the final deliverable of the dissertation work.

ARC's ABM provides the TAZ ID for all households, and the longitude and latitude information were generated by randomly pairing the ABM households with the households in the Epsilon household-level demographic dataset. The 2019 Epsilon dataset includes 94 attributes (columns) of demographic information of the metro Atlanta area, including household size (number of adults and number of children), household income (relative index compared to national average), ethnicity of every household member, and the longitude and latitude of the household. The Epsilon households were grouped into TAZs by allocating their geographic information into the roadway network, and for every ABM household used in the assessment, a random Epsilon household location was drawn from the corresponding TAZ (with duplicate).

The team is working on expanding the random allocation into a more comprehensive synthetic fleet and household generation algorithm that accounts for more spatially-resolved demographic information (e.g., income, household size, ethnicity, etc.) in the pairing process. ABM2020 provides 29 variables with respect to household demographics, which can be (partially) paired with a subset of the 94 variables in the Epsilon 2019 dataset, e.g., household size (ABM2020) vs. the number of children plus the number of adults (Epsilon 2019). A list of final variable pairing is shown in Table 6, and the team is expanding the list by examining the rest of the variables.

Table 6. Variables Pairing of ABM2020 vs. Epsilon 2019 for Household Demographics

No	Attributes	ABM2020 Variable	Epsilon 2019 Variable(s)
1	Household size	np	Number of Children, Advantage Number of Adults
2	Number of workers	nwrkrs_esr	Occupation
3	Household income group	hh_inc_bin	Advantage Target Narrow Band Income 2.0
4	Vehicles (1 ton or less) available	veh	N/A
5	Household/family type	hht	Family composition (Enhanced)
6	Units in structure	bld	Advantage Dwelling Type

For each traversed link, the geometry of the roadway was decomposed into points based on the number of seconds that the traveler spends on the link (an even split based on the average speed), and the second-by-second trajectories were assigned accordingly. Given that the link-by-link speed profiles do not change within each ABM period, the travel paths that traverse the same link were only generated once (overlapped trajectories) to minimize computation workload.

The trajectories at 7:00 AM that traverse the managed lane ramps at Roswell Road at I-575 are presented in Figure 15 as an example. The travel trajectories of trips passing through the ramp are displayed as dots (each dot represents one vehicle). The left panel displays all trips, the middle panel color-codes them by vehicle occupancy (SOV, HOV2, HOV3), and the right panel categorizes them by annual income (<$50K, $50K-$100K, >$100K). Note that travel paths often traverse a much larger area than the NWC, but only the NWC sections are retained for this analysis. For the modeled I-575 and I-75 links, and the overpass arterials, the on-the-move trajectories largely fall outside the study area because these trips are long. Hence, population exposure was predominantly from at-home concentrations.

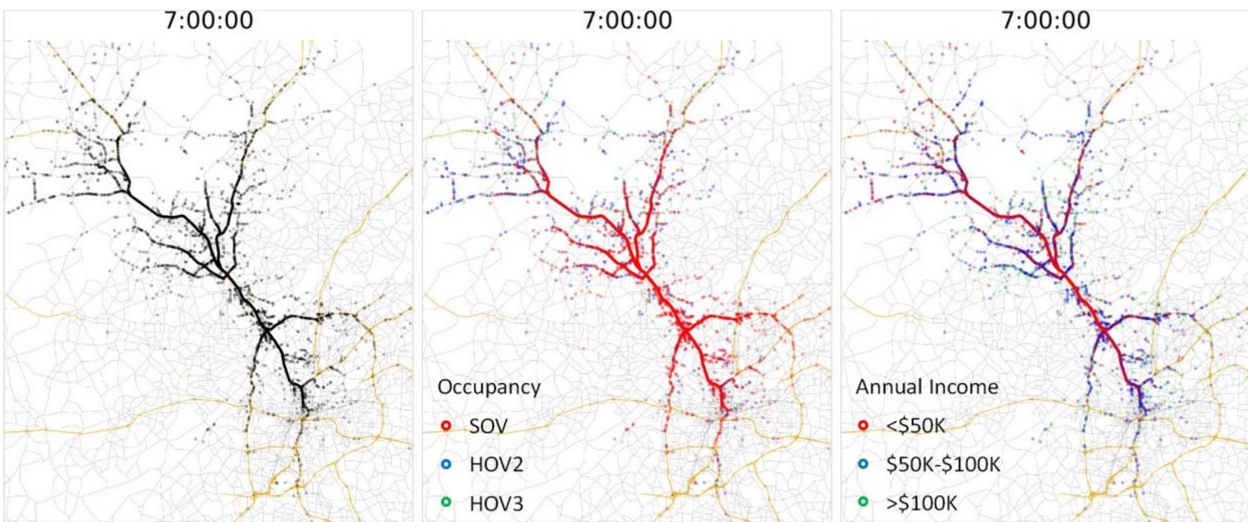

Figure 15. Path Retention Animation Screenshot

2.7 Exposure Modeling by Demographic Group

The exposure modeling of individuals was performed by overlaying the concentration profiles predicted by AERMOD over the hourly position information for each person (all second-by-second exposure profiles (g/m^3) were aggregated to one-hour intervals). The individual-level modeling results were then aggregated into 16 mutually-exclusive groups defined by combinations of demographic variables (vehicle ownership, annual income, household size, number of workers, w/o kid, etc.). The comparison across the demographic groups were performed based on the average hourly person exposure to traffic related $PM_{2.5}$ for the modeled traffic day and meteorology year (i.e., hourly concentration results for the entire year).

The second-by-second travel trajectories (at-home and in-route) were mapped to the roadway network (using Dijkstra's algorithm) and each linked to the closest receptor. For each second, the predicted PM2.5 concentration was retrieved from the hourly concentration profiles, and these second-by-second concentration values were accumulated over 3,600 seconds (one hour) to derive the exposure for that hour (differences in breathing rates under different conditions were not modeled in this study). The receptors placed over the roadway were removed from the analyses via receptor filtering before pairing with the vehicle trajectory data, because EPA does not recommend using AERMOD to assess concentrations over the top of roadway surface (i.e., immediately above the transportation network source links); instead, on-road trajectories were paired to nearest-adjacent receptor."

The ABM2020 demographic information includes household size, number of workers, annual income, number of households, with or without children, etc. The 16 demographic groups used in this study was consistent with the previous study as definitions that are mutually exclusive or collectively exhaustive (Zhao, et al., 2019), and the breakdown of the groups are presented in Table 7. These groupings were established in a collaboration between DOE staff and Georgia Tech researchers for the 2018 ARPA-E TRANSNET-Atlanta project for equity impact assessment (Zhao 2020). The mutually-exclusive groups include lifecycle stages, but were also specifically designed to reveal transportation mobility and accessibility impacts on households that own no vehicles, households with very low-income households, and single parent households with children (i.e., severe constraints to mobility and accessibility). All plus markers ("+") in the table indicate inclusive ranges (e.g., "2+" means "two or more").

The households were first classified by ownership of vehicles, and those with no vehicles were placed into Group #1 (households with no vehicles). The households with at least one vehicle were then examined and divided based on annual income; those below the threshold of $25,000 were classified as Group #2 (low-income households). The remaining households were then categorized based on household size (number of adults and children), and further classified based on number of workers, presence of children, and whether or not the household includes a single parent with children (for households with three or more persons only). The one-person households were placed into Group #3 (non-working households with annual income lower than $60,000), Group #4 (working households with annual income lower than $60,000), and Group #5 (annual income more than $60,000). Households with two or more persons were examined by annual income and categorized into intervals of $25,000 to $60,000 (Groups #6, #7, #8, and #9), $60,000 to $120,000 (Groups #10, #11, #12, #13, and #14), and $120,000 or more (Groups #15 and #16). For households with annual income lower than $60,000 (and with two or more persons), the non-working households were classified into Group #6, and two-person working households were classified into Group #7. Then, the working households with three or more persons were classified based on whether they are single parent with children (Group #8) or without children (Group #9). For the households with annual income from $60,000 to $120,000, the households with either no or one worker were classified based on whether they have at least one kid (Group #10) or not (Group #11), and the two-person households with two workers were classified into Group #12. The three-person households with at least two workers were classified into Group #13, and those households with four or more persons and with at least two workers were classified into Group #14. The households with annual income of $120,000 or more were classified into Group #15 (with no or one worker) and Group #16 (with two or more workers). The classifications result in each household being placed into one, and only one, group.

The team is working on expanding the comparison across demographics by introducing the variables from the Epsilon 2019 dataset (after the synthetic household pairing process), such as ethnicity, as shown in Figure 16. However, race and income are highly correlated in the Atlanta metro area, so the income-based results shown in this work are still revealing.

Table 7. Definition of the 16 Demographic Groups based on ABM Household Information

Group #	Own Vehicles	Low Income	HH Size	Annual Income	Workers	Vehicles	With Kid(s)	Single Parent w/kid(s)	Percentage
1	No	Any	Any	Any	Any	0	Any	Any	2.2%
2	Yes	Yes	Any	$0 - $25k	Any	1+	Any	Any	6.7%
3	Yes	No	1	$25k - $60k	0	1+	Any	Any	0.8%
4	Yes	No	1	$25k - $60k	1	1+	Any	Any	4.1%
5	Yes	No	1	$60k+	0 or 1	1+	Any	Any	3.4%
6	Yes	No	2+	$25k - $60k	0	1+	Any	Any	2.5%
7	Yes	No	2	$25k - $60k	1+	1+	Any	Any	5.4%
8	Yes	No	3+	$25k - $60k	1+	1+	Any	Yes	1.9%
9	Yes	No	3+	$25k - $60k	1+	1+	Any	No	11.7%
10	Yes	No	2+	$60k - $120k	0 or 1	1+	Yes	Any	3.4%
11	Yes	No	2+	$60k - $120k	0 or 1	1+	No	Any	8.8%
12	Yes	No	2	$60k - $120k	2	1+	Any	Any	5.3%
13	Yes	No	3	$60k - $120k	2+	1+	Any	Any	6.4%
14	Yes	No	4+	$60k - $120k	2+	1+	Any	Any	14.0%
15	Yes	No	2+	$120k+	0 or 1	1+	Any	Any	5.9%
16	Yes	No	2+	$120k+	2+	1+	Any	Any	17.4%

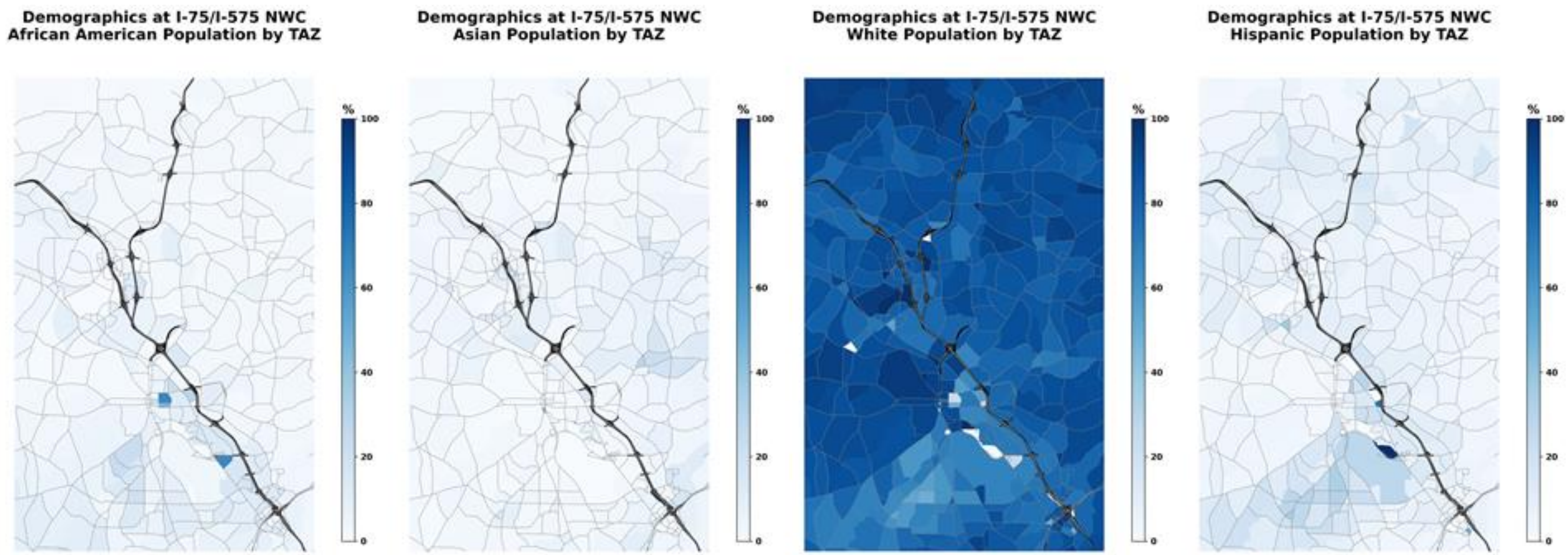

Figure 16. Percentages of Race Distributions by TAZ from the Epsilon 2019 Dataset

Chapter 3. Results and Discussion

This section presents and discusses the results of the modeling framework, including the predicted concentration profiles (in heat maps), the population exposure across demographic groups, and the kernel density maps of household income to help interpret the results.

A sample of the predicted $PM_{2.5}$ concentration profiles is presented in Figure 17, where the temperature, relative humidity, wind speed and wind directions are marked on the heat map. The team is converting the heat maps into animations, and uploading them to the YouTube channel operated by the transportation group of Georgia Tech.

The predicted concentration by receptor is largely dependent on the meteorology settings, especially wind direction and wind speed. Near-road concentrations are generally higher than in areas that are further away. This is consistent with the link screening results where receptors more than about 5 kilometers from the roadway links are barely impacted (concentration impact less than 0.01 μg/m^3 under all meteorological circumstances).

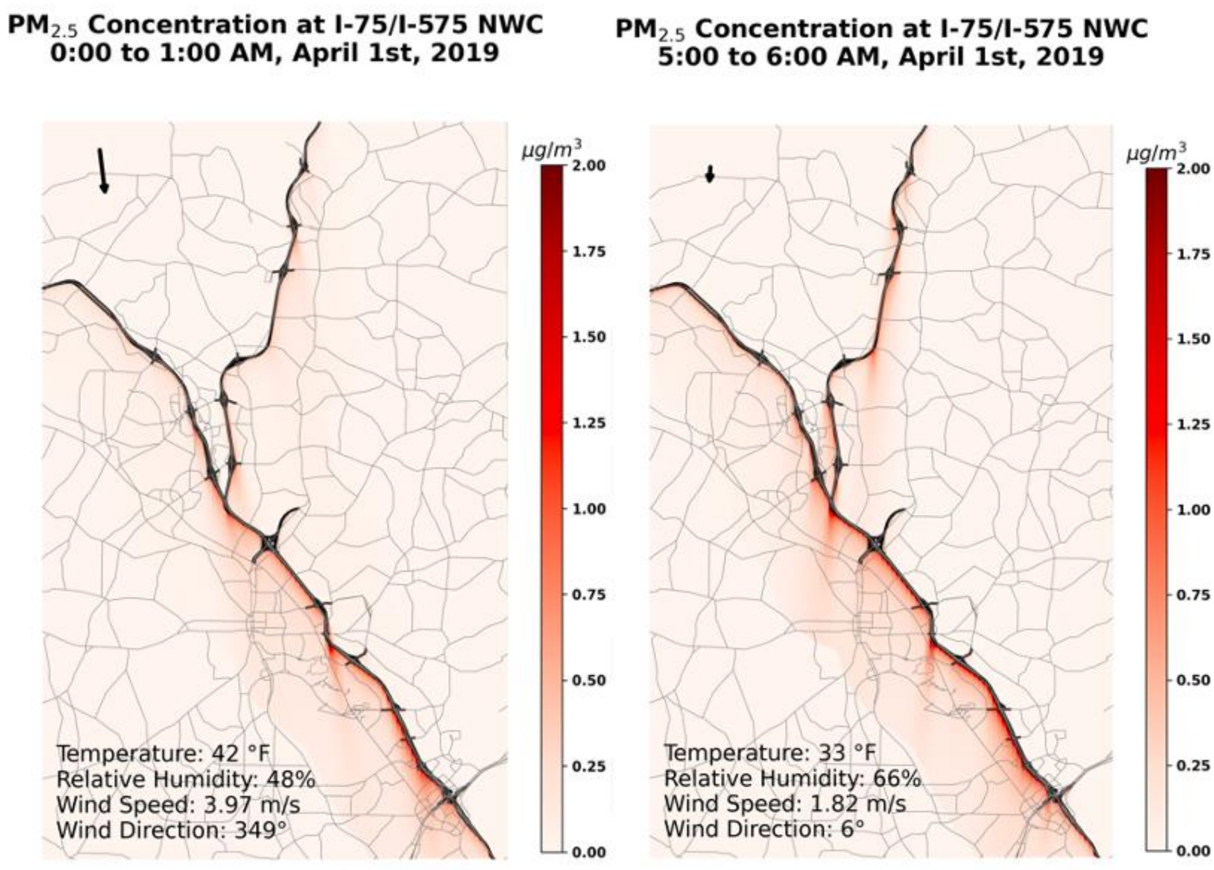

Figure 17. Predicted $PM_{2.5}$ Concentration Profiles of the NWC

The population exposure modeling results (average exposure across 24 hours) by demographic group are presented in Table 8. This assessment only accounted for the I-75/I-75 NWC subarea, and only for the Interstate highway and overpass arterials that were modeled. The modeling results are predominantly based on the at-home exposure (only about 10% to 15% of the travel paths for these households fall inside the modeling area because trips are predominately made from home to other areas in the region); hence, it is likely that the results for the entire subarea will yield somewhat different patterns in terms of the overall exposure profiles once all metro area households and all metro subareas are included in the analyses.

Demographic Group #6 was found to be exposed to the largest amount of traffic-related $PM_{2.5}$, which is not surprising giving that these individuals are likely to spend much more time at home than members of working households. Similarly, Group #10, Group #3, and Group #11 were also found larger than the rest of the groups (either non-working households, or not everyone in the households works), likely because they have smaller number of workers (more at-home time). Larger households (with three or more persons) with two or more workers were found to be exposed to much lower amounts (2.11 $\mu g/m^3$ for Group #13, and 2.14 $\mu g/m^3$ for Group #14), which is also not surprising giving their decent income (over $60,000) and that more workers were going to work (likely outside of the modeling area).

Households with children were exposed to more traffic-related $PM_{2.5}$ than those with no children, with all other demographic information held consistent, which is anticipated given that children are likely to spend more time at home, school, or in short trips where they are likely to stay within the modeling subarea.

The households with no vehicles (Group #1) were exposed to a relatively low amount of traffic-related $PM_{2.5}$ (2.11 $\mu g/m^3$, similar to the high-income households), which could be that these households were close to transit terminals (e.g., Cumberland Mall, which is near the edge of the modeling subarea), and were not located near the modeled roadway.

It might seem counterintuitive that the low-income Group #2 was exposed to a lower amount traffic-related $PM_{2.5}$ (2.18 $\mu g/m^3$), but this is likely due to: 1) these households are located near the Dobbins Air Reserve Base that are either not close to the Interstate (low accessibility to freeway) or near a major branch of I-75 before the split, close to the edge of the study area (low-concentration), as shown in Figure 18. It could also be related to the longer commute distance (e.g., working in downtown Atlanta), and possibly longer working hours (i.e., longer time that is outside the study area). The results for the whole metro area (with all links of freeway and major arterials) could indicate a larger exposure.

The middle-income household groups (Groups #7 with 2.19 $\mu g/m^3$, Group #8 with 2.20 $\mu g/m^3$, Group 9 with 2.19 $\mu g/m^3$, and Group #12 with 2.19 $\mu g/m^3$) fell into the middle of the group exposure metrics, which could be due to that their households were neither too far nor too close to the modeled roadway (see Figure 19 and Figure 20), and these groups could be sharing similar travel patterns (in terms of their working schedule, travel distances, etc.).

Overall, the high-income groups of Group #5 (over $120,000), Group #15 (more than $120,000), and Group #16 (more than $120,000) were found to be exposed to the lowest traffic-related

$PM_{2.5}$, which is not surprising given their resident locations (further away from the Interstate and consequently away from traffic-related air pollutions and noise), as shown in Figure 21.

Although the modeling results are indicating a correlation between household income and exposure, these patterns could change when the whole metro area is modeled, given that the overall spatial distributions of the households (e.g., middle-income and high-income households) and their travel activity will likely be different along other major corridors.

The proposed modeling framework offers a high-efficiency modeling tool for conducting detailed environmental justice and equity assessments of individual and household exposure to traffic-related PM2.5 across demographic groups. Specifically, this framework enables high-resolution modeling of population exposure through second-by-second vehicle travel paths and demographic overlays, which allows for the identification of pollution exposure disparities among groups defined by income, ethnicity, and household characteristics. The framework can also assess the impact of proximity to major roadways, and provide insights into how exposure varies by locations and travel activities.

The purpose of this research is to develop a modeling framework, compatible with a wide-variety of modeling tools, that creates a complete modeling chain that starts with travel behavior, incorporates emission modeling, integrates dispersion modeling, and ends with population exposure. Some items not yet included in this process but relevant for future expansion include: 1) indoor versus outdoor concentration results, 2) in-vehicle versus outdoor concentration results, 3) micro-activities within each TAZ (e.g., walking from a parking lot to a destination, trips within the TAZ, and indoor activities), 4) inhalation rates that vary based on individual attributes (e.g., age, gender) and activities (e.g., walking versus sitting). However, this framework provides flexibility for future research expansions that address these elements as additional data and resources become available.

Table 8. Population Exposure to Traffic-Related $PM_{2.5}$ by Demographic Group

Group #	Own Vehicles	Low Income	HH Size	Annual Income	Workers	Vehicles	With Kid(s)	Single Parent w/kid(s)	Percentage	Average Hourly Exposure ($\mu g/m^3$)
1	No	Any	Any	Any	Any	0	Any	Any	2.2%	2.11
2	Yes	Yes	Any	\$0 - \$25k	Any	1+	Any	Any	6.7%	2.18
3	Yes	No	1	\$25k - \$60k	0	1+	Any	Any	0.8%	2.29
4	Yes	No	1	\$25k - \$60k	1	1+	Any	Any	4.1%	2.17
5	Yes	No	1	\$60k+	0 or 1	1+	Any	Any	3.4%	2.10
6	Yes	No	2+	\$25k - \$60k	0	1+	Any	Any	2.5%	2.38
7	Yes	No	2	\$25k - \$60k	1+	1+	Any	Any	5.4%	2.19
8	Yes	No	3+	\$25k - \$60k	1+	1+	Any	Yes	1.9%	2.20
9	Yes	No	3+	\$25k - \$60k	1+	1+	Any	No	11.7%	2.19
10	Yes	No	2+	\$60k - \$120k	0 or 1	1+	Yes	Any	3.4%	2.29
11	Yes	No	2+	\$60k - \$120k	0 or 1	1+	No	Any	8.8%	2.24
12	Yes	No	2	\$60k - \$120k	2	1+	Any	Any	5.3%	2.19
13	Yes	No	3	\$60k - \$120k	2+	1+	Any	Any	6.4%	2.11
14	Yes	No	4+	\$60k - \$120k	2+	1+	Any	Any	14.0%	2.14
15	Yes	No	2+	\$120k+	0 or 1	1+	Any	Any	5.9%	2.12
16	Yes	No	2+	\$120k+	2+	1+	Any	Any	17.4%	2.04

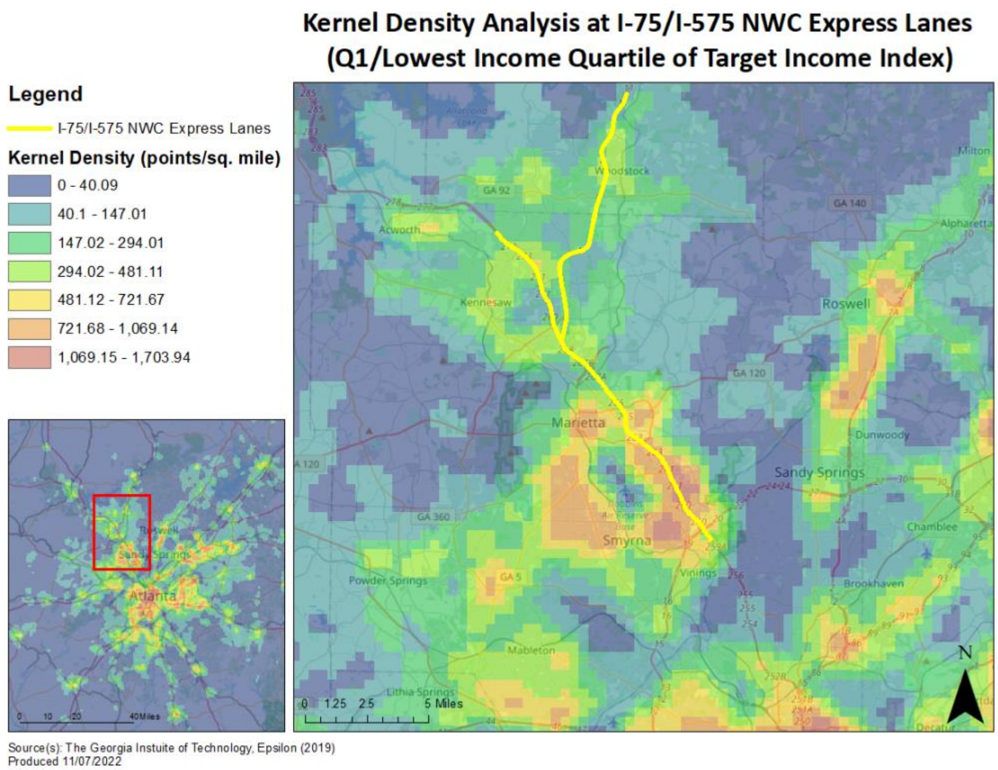

Figure 18. Kernel Density Map of Lowest Quartile (Q1) of Household Income

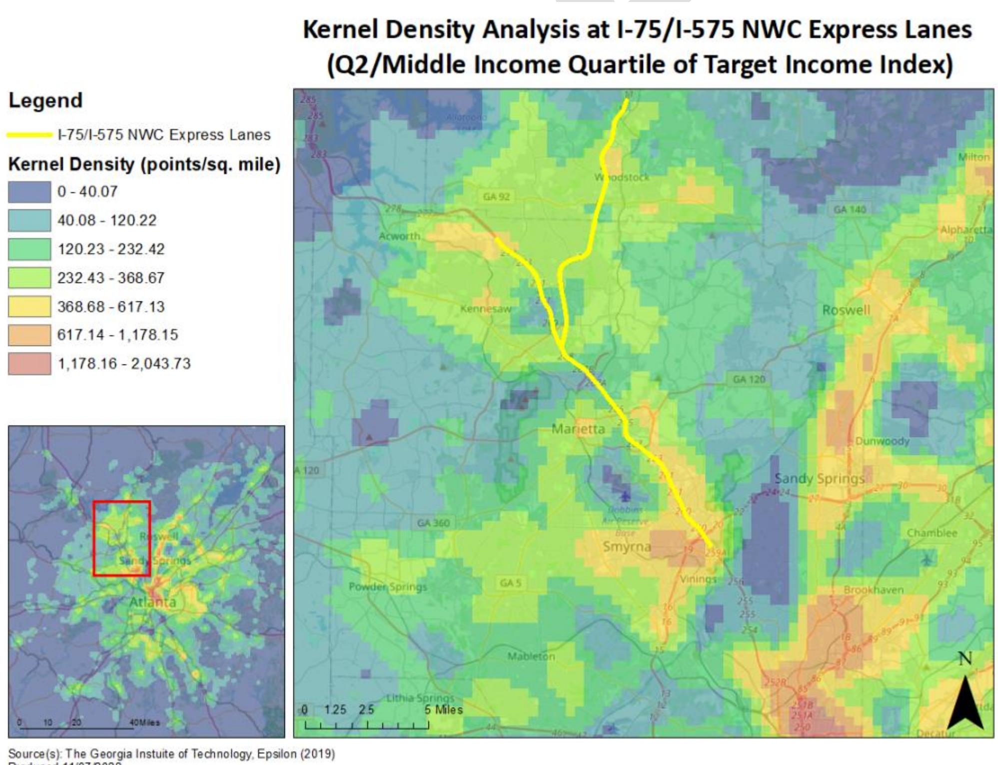

Figure 19. Kernel Density Map of Second Lowest Quartile (Q2) of Household Income

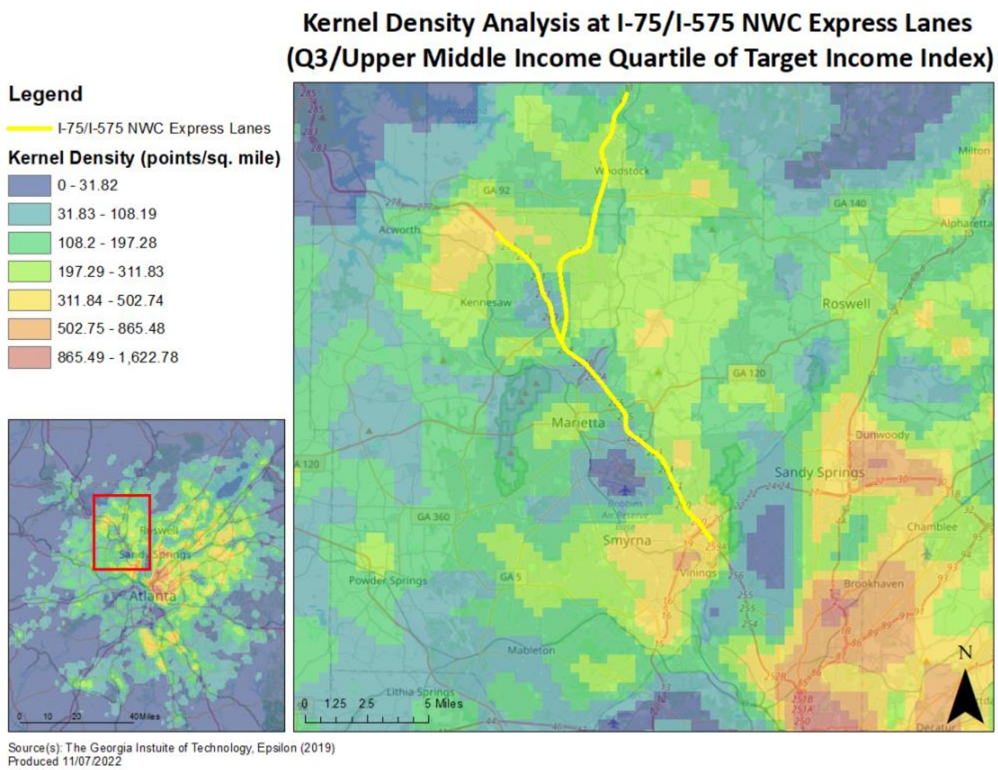

Figure 20. Kernel Density Map of Second Highest Quartile (Q3) of Household Income

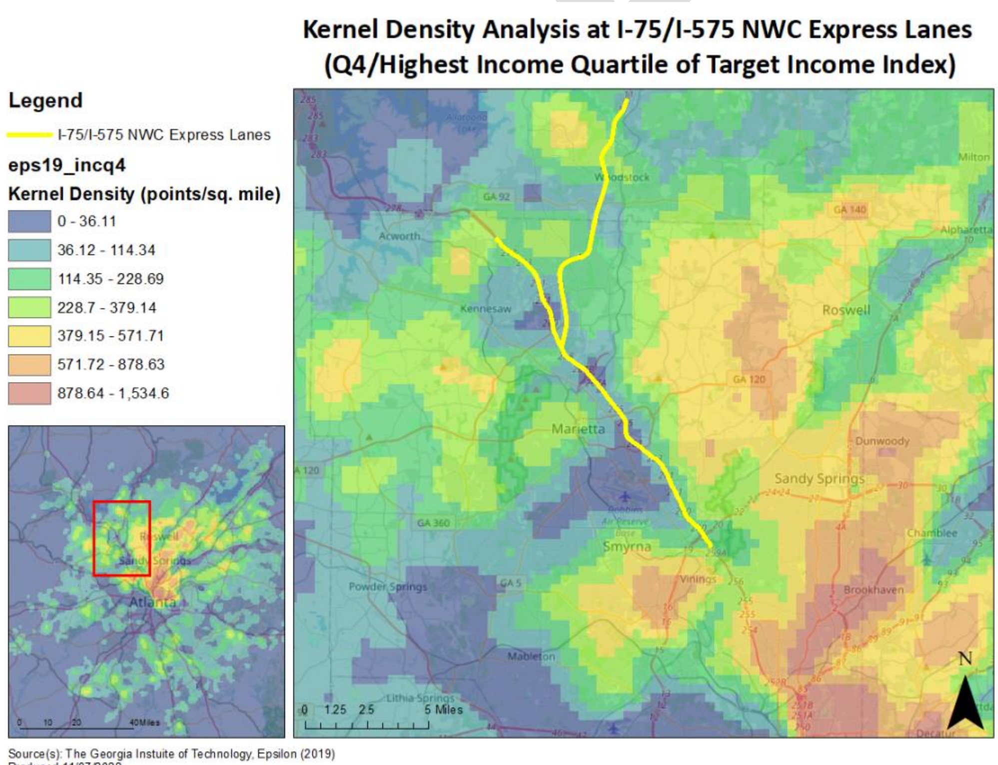

Figure 21. Kernel Density Map of Highest Quartile (Q4) of Household Income

Chapter 4. Conclusions and Future Work

This study developed a modeling framework to integrate predicted person and household activity (person locations and predicted vehicle paths) from the regional activity-based travel demand model (ABM2020), with MOVES-Matrix energy and emission rates, and the AERMOD microscale dispersion model, to assess near-road population exposure to traffic-related $PM_{2.5}$ across demographic groups. All of the modeling components were to run on the Georgia Tech PACE supercomputing cluster. A case study of the I-75/I-575 Northwest Corridor demonstrated the capabilities of the overall modeling framework.

Northwest Corridor population exposures were correlated with distance from homes to the modeled freeway network. High-income households do tend to live further away from the Interstate corridor, and were exposed to slightly less traffic-related $PM_{2.5}$. However, low-income populations were not among the groups with highest daily exposure along this corridor, predominantly due to the fact that household incomes tend to be relatively high throughout the NWC (compared to the rest of the region) and spatial clustering by income is not very notable on this corridor. Based upon the spatial clustering of low-income households throughout other portions of the metro area, the team anticipates that environmental equity disparities in exposure across income groups will be noted in modeling the entire metro area.

The team is currently performing QA/QC processes on recently-procured Polk Data Services vehicle registration dataset that will be used to refine the next set of analyses for enhanced fleet composition to refine emission rates. Addresses were reformatted (e.g., Ave. to Avenue) using a machine learning process to identify common registration locations for demographic linkages and the team is currently performing a spatial completeness check. The team is also running algorithms to condense makes/models into common names (e.g., a 2019 Honda Civic may have 40 different names for the same vehicle), and comparing vehicle make, model, and model year groups with the anticipated values to assess potential sample bias (e.g., older vehicles in the fleet may tend to be missing). The team will also expand the equity assessment by integrating the vehicle profiles from Polk (make, model and model year) into the synthetic household and fleet generation, to pair the vehicle emissions at individual and household levels (modeled using MOVES-Matrix) with the corresponding exposure results.

References

ARC, 2015. The Atlanta Region's Plan – Conformity Determination Report. Atlanta, GA. https://documents.atlantaregional.com/The-Atlanta-Region-s-Plan/rtp/Conformity-Determination-Report.pdf

Beckx, C., L. Int Panis, T. Arentze, D. Janssens, S. Broekx, G. Wets, 2009. A Dynamic Activity-Based Population Modelling Approach to Evaluate Exposure to Air Pollution: Methods and Application to A Dutch Urban Area. Environ. Impact Assess. Rev. 29(3), 179–185. https://doi.org/10.1016/j.eiar.2008.10.001

Clark, L.P., Millet, D.B., Marshall, J.D., 2014. Changes in Transportation-Related Air Pollution Exposures by Race-Ethnicity and Socioeconomic Status: Outdoor Nitrogen Dioxide in the United States in 2000 and 2010 1–10. https://doi.org/10.1289/EHP959

Dias, D., O. Tchepel, 2018. Modelling of Human Exposure to Air Pollution in the Urban Environment: A GPS-Based Approach. Environ. Sci. Pollut. Res. 2014 Mar;21(5):3558-71. https://doi.org/10.1007/s11356-013-2277-6

Frey, H.C., 2018. Trends in On-road Transportation, Energy, and Emissions. J. Air Waste Management Assoc. 68, 1015–1024. https://doi.org/10.1080/10962247.2018.1454357

Guensler, R., H. Lu, W. Reichard, Z. Dai, T. Xia, A. Guin, and M.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, R., Liu, H., Xu, Y. (Ann), Akanser, A., Kim, D., Hunter, M.P., Rodgers, M.O., 2017. Energy Consumption and Emissions Modeling of Individual Vehicles. Transp. Res. Rec. J. Transp. Res. Board 2627, 93–102. https://doi.org/10.3141/2627-11

Hatzopoulou, M., E. J. Miller, 2010. Linking an Activity-Based Travel Demand Model with Traffic Emission and Dispersion Models: Transport's Contribution to Air Pollution in Toronto. Transp. Res. Part D: Transp. Environ. 15(6), 315–325. https://doi.org/10.1016/j.trd.2010.03.007

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

Kim, D., Liu, H., Rodgers, M.O., Guensler, R., 2020a. Development of Roadway Link Screening Model for Regional-level Near-road Air Quality Analysis: A Case Study for Particulate Matter. Atmos. Environ. 237, 117677. https://doi.org/10.1016/j.atmosenv.2020.117677

Kim, D., Liu, H., Xu, X., Lu, H., Wayson, R., Rodgers, M.O., Guensler, R., 2020b. Distributed computing for region-wide line source dispersion modeling. Comput. Civ. Infrastruct. Eng. 1–15. https://doi.org/10.1111/mice.12639

Kim, D., Liu, H., Xu, X., Lu, H., Wayson, R., Rodgers, M.O., Guensler, R., 2019. A Regional Air Quality Impact Assessment Screening Tool based upon MOVES-Matrix and AERMOD, in: 8th Specialty Conference on Air Quality Models - Guideline for Air Quality Models: Planning Ahead. Durham, NC, U.S.

Li, H., Liu, H., Xu, X., Xu, Y. "Ann," Rodgers, M.O., Guensler, R., 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, H., Wang, Y., Xu, X., Liu, H., Guin, A., Rodgers, M.O., Hunter, M.P., Laval, J.A., Guensler, R., 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, H., Guensler, R., Lu, H., Xu, Y., Xu, X., Rodgers, M.O., 2019. MOVES-Matrix for High-Performance On-Road Energy and Running Emission Rate Modeling Applications. J. Air Waste Manage. Assoc. 10962247.2019.1640806. https://doi.org/10.1080/10962247.2019.1640806

Luo, J., Boriboonsomsin, K., Barth, M.J., 2017. Assessing Traffic-Related Primary Fine Particle Exposure for Southern California Residents and the Environmental Inequality Implications.

Marcantonio, Richard A., A. Golub, A. Karner, and L. Nelson. 2017. Confronting Inequality in Metropolitan Regions: Realizing the Promise of Civil Rights and Environmental Justice in Metropolitan Transportation Planning, 44 Fordham Urb. L.J. 1017. https://ir.lawnet.fordham.edu/ulj/vol44/iss4/4

Volk, H.E., Lurmann, F., Penfold, B., Hertz-Picciotto, I., McConnell, R., 2013. Traffic-related Air Pollution, Particulate Matter, and Autism. Arch. Gen. Psychiatry 70, 71–77. https://doi.org/10.1001/jamapsychiatry.2013.266

Wang, M., Hong, B., Puett, R.C., Yanosky, J.D., Laden, F., Hart, J.E., Fisher, J.A., Spiegelman, D., 2014. Particulate Matter Air Pollution Exposure, Distance to Road, and Incident Lung Cancer in the Nurses' Health Study Cohort. Environ. Health Perspect. 122, 926–932. https://doi.org/10.1289/ehp.1307490

WHO, World Health Organization, 2018. Exposure to Ambient Air Pollution from Particulate Matter for 2016. Geneva, Switzerland.

Wu, Y., 2018. Integrated assessment for health effects of sustainable transportation strategies. Diss. Abstr. Int. Sect. B Sci. Eng. 80, No-Specified.

Xu, X., Liu, H., Anderson, J.M., Xu, Y., Hunter, M.P., Rodgers, M.O., 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, X., Liu, H., Guin, A., Rodgers, M.O., Guensler, R., 2018. Regional Emission Analysis with Travel Demand Models and MOVES-Matrix, in: 97th Annual Meeting of Transportation Research Board. Washington, D.C., U.S.

Xu, X., Liu, H., Li, H., Rodgers, M.O., Guensler, R., 2017. Integrating Engine Start, Soak, Evaporative, and Truck Hoteling Emissions into MOVES-Matrix, in: 97th Annual Meeting of Transportation Research Board. Washington, D.C., U.S.

Xu, Y., Li, H., Liu, H., Rodgers, M.O., Guensler, R.L., 2017. Eco-driving for Transit: An Effective Strategy to Conserve Fuel and Emissions. Appl. Energy 194, 784–797. https://doi.org/10.1016/j.apenergy.2016.09.101

Zhang, A., Qi, Q., Jiang, L., Zhou, F., Wang, J., 2013. Population Exposure to PM 2.5 in the Urban Area of Beijing. PLoS One 8. https://doi.org/10.1371/journal.pone.0063486

Zhao, Y., 2020. Distributional Impact Assessment of Transportation Policy Using Activity-based Models with Path Retention. Ph.D. Dissertation. Georgia Institute of Technology.

Zhao, Y., Xu, X., Liu, H., Li, H., Wang, Y., Guin, A., Rodgers, M.O., Guensler, R. (2019). Using ABM15, Path Retetion, and MOVES Matrix to Assess Differences in Energy Use Acorss 16 Demographic Groups in Atlanta, in: 98th Annual Meeting of Transportation Research Board. Washington, D.C., U.S.

Data Summary

As described in this report, the team modeled the traffic-related $PM_{2.5}$ emissions and concentrations, and analyzed the population exposure across demographics.

Products of Research

The traffic volume and speed data used in this study were derived from the Activity-based Model (ABM) by Atlanta Regional Commission (ARC). Under the data user agreement, interested parties need to obtain these data from the corresponding contractors. The energy and emission rate matrices applied are public domain and can be found at https://zenodo.org/records/13878218. The AERMOD modeling results by receptors are also included at the repository.

Data Format and Content

The format and content of the MOVES-Matrix (MOVES2014b) data sets are documented in the NCST MOVES-Matrix overview and training documents at https://github.com/gti-gatech/moves_training/.

Data Access and Sharing

The MOVES-Matrix data are open source and can be downloaded and freely shared from the link provided above.

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 emissions rate matrices should be cited as follows:

> Lu, H. (2024). A Modeling Framework for Near-Road Population Exposure to Traffic-Related PM2.5 and Environmental Equity Analysis: A Case Study in Atlanta, Georgia [Data set]. Zenodo. https://doi.org/10.5281/zenodo.13878218

Appendix. Randomization of Trip Departure Timestamp

This appendix provides the detailed process and an example of the trip departure timestamp randomization based on the trip departure period and the constraints of trip chains. Two iterations were performed for every trip chain to generate the early bounds and late bounds for all trips, respectively. The forward checking generates the early bounds in a sequence from the first trip to the last trip in a chain (from trip 1 to trip n), as shown in (1), and the backward checking generates the late bounds in a sequence from the last trip to the first trip in a chain (from trip n to trip 1), as shown in (2).

$$EB_i \begin{cases} PS_i & i = 1 \\ MAX(PS_i, EB_{i-1} + t_{i-1}) & i = 2, 3, \dots, n \end{cases} \quad (1)$$

$$LB_i \begin{cases} PE_i & i = n \\ MIN(LB_{i+1} - t_{i+1}, PE_i) & i = 2, 3, \dots, n \end{cases} \quad (2)$$

where EB_i is the early bound for trip i, LB_i is the late bound for trip i, PS_i is the start timestamp for the departure period of trip i (e.g., 8:00 AM if the departure period is 8:00 AM to 8:30 AM), PE_i is the end timestamp for the departure period of trip (e.g., 8:30 AM if the departure period is 8:00 AM to 8:30 AM), and t_i is the travel time of trip t_i.

An example of dummy trip chain is presented in Table 9 with the calculation of the legitimate departure periods for a 7-trip chain, and the legitimate periods are illustrated in Figure 22. It is worth noting the legitimate periods across trips could overlap with each other (i.e., in this case trips could be generated in an incorrect sequence), and the departure timestamp randomization algorithm was designed to discard the invalid trip sequences (re-draw of all trips at a time).

Table 9. Sample of the Generation of the Legitimate Trip Departure Time Periods

Trip #	Departure Period #	Assigned Departure Period	Travel Time (Minute)	Early Bound (Forward Checking)	Late Bound (Backward Checking)
1	17	8:00 - 8:30 AM	15	8:00 AM	MIN (8:45 AM - 15 Min, 8:30 AM) = 8:30 AM
2	18	8:30 - 9:00 AM	5	MAX (8:30 AM, 8:00 AM + 15 Min) = 8:30 AM	MIN (8:50 AM - 5 Min, 9:00 AM) = 8:45 AM
3	18	8:30 - 9:00 AM	10	MAX (8:30 AM, 8:30 AM + 5 Min) = 8:35 AM	MIN (9:00 AM - 10 Min, 9:00 AM) = 8:50 AM
4	18	8:30 - 9:00 AM	5	MAX (8:30 AM, 8:35 AM + 10 Min) = 8:45 AM	MIN (9:10 AM - 5 Min, 9:00 AM) = 9:00 AM
5	19	9:00- 9:30 AM	10	MAX (9:00 AM, 8:45 AM + 5 Min) = 9:00 AM	MIN (9:20 AM - 10 Min, 9:30 AM) = 9:10 AM
6	19	9:00 - 9:30 AM	40	MAX (9:00 AM, 9:00 AM + 10 Min) = 9:10 AM	MIN (10:00 AM - 40 Min, 9:30 AM) = 9:20 AM
7	20	9:30 - 10:00 AM	5	MAX (9:30 AM, 9:10 AM + 40 Min) = 9:50 AM	10:00 AM

Ganntt Chart of the Sample Trip Chain

Each Slot Represents a 5-Minute Interval. Legend: Legitimate Period; Legitimate Period Overlapped with Other Trips; Assigned Period That is Not Legitimate

Trip #	Departure Period #	Assigned Departure Period	Legitimate Departure Period
1	17	8:00 AM - 8:30 AM	8:00 AM - 8:30 AM
2	18	8:30 AM - 9:00 AM	8:30 AM - 8:45 AM
3	18	8:30 AM - 9:00 AM	8:35 AM - 8:50 AM
4	18	8:30 AM - 9:00 AM	8:45 AM - 9:00 AM
5	19	9:00 AM - 9:30 AM	9:00 AM - 9:10 AM
6	19	9:00 AM - 9:30 AM	9:10 AM - 9:20 AM
7	20	9:30 AM - 10:00 AM	9:50 AM - 10:00 AM

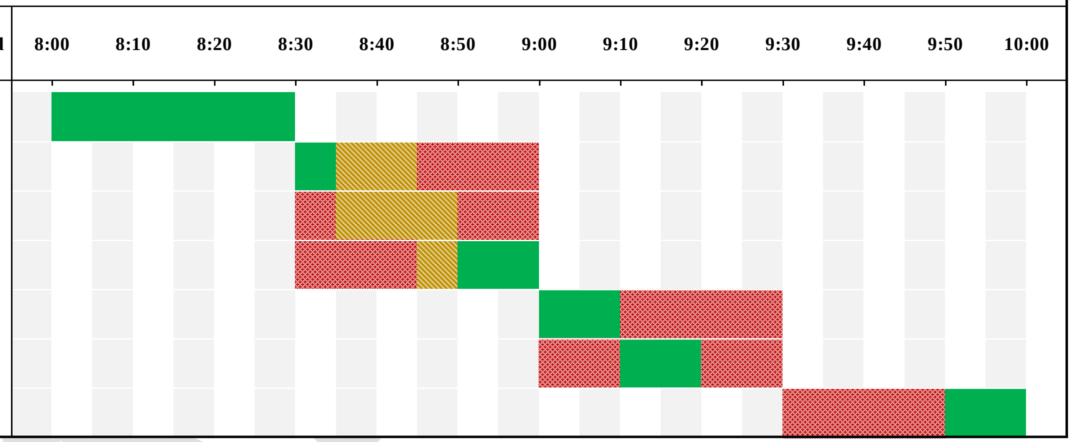

Figure 22. Sample of the Legitimate Trip Departure Time Periods