Texas A&M Transportation Institute

On-road Mobile Air Emissions Reporting Requirement (AERR) Methodology Assessment and Input Sensitivities

DRAFT REPORT

Prepared for the Texas Commission on Environmental Quality (TCEQ)

January 26, 2024

Texas A&M Transportation Institute

Texas A&M Transportation Institute

DRAFT

Proposal for Grant Activities No: 582-22-32060-004

Sub-Task 6.2 – Final Report

DATE: January 26, 2024

TO: Mogwai Turner
Texas Commission on Environmental Quality (TCEQ)

COPY TO: Julie D. Vanden Berg, TCEQ

FROM: Mihir Mandar Kulkarni
Apoorba Bibeka, P.E.
Chaoyi Gu, P.E.
Madhusudhan Venugopal, P.E.
Texas A&M Transportation Institute

FOR MORE INFORMATION:

Madhusudhan Venugopal, P.E.
972-994-2213
m-venugopal@tti.tamu.edu

ABSTRACT: The draft report, titled "On-road Mobile Air Emissions Reporting Requirement (AERR) Methodology Assessment and Input Sensitivities", is prepared for the Texas Commission on Environmental Quality (TCEQ) by the Texas A&M Transportation Institute. It focuses on developing and accessing methodologies for on-road mobile source Air Emissions Reporting Requirements (AERR) inventories. The study aims to evaluate the impact of various input parameters on emission estimates and explore alternative data sources for improving emissions inventories. The report includes extensive reviews of existing methodologies, sensitivity analysis of MOVES3 activity inputs, and identifying Texas-specific activity data sources. Key findings include the importance of accurate data for emission estimates and recommendations for potential data sources to improve future inventories.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS AND ABBREVIATIONS

Acronym	Definition
AERR	Air Emissions Reporting Requirements
APU	Auxiliary Power Unit
ATR	Automatic Traffic Recorder
AVFT	Alternative Vehicle Fuel and Technology
BTS	Bureau of Transportation Statistics
CAP	Criteria Air Pollutants
CDB	County Database
CH_4	Methane
CNG	Compressed Natural Gas
CO	Carbon Monoxide
CO_2	Carbon Dioxide
CO_{2eq}	Carbon Dioxide Equivalent
CT	Combination Truck
DMV	Department of Motor Vehicles
DOE	Department of Energy
E85	Ethanol
EI	Emissions Inventory
EPA	Environmental Protection Agency
ERG	Eastern Research Group
FHWA	Federal Highway Administration
GHG	Greenhouse Gases
GPS	Global Positioning System
GVWR	Gross Vehicle Weight Rating
HPMS	Highway Performance Monitoring System
IHS	IHS-Markit
I/M	Inspection and Maintenance
LBS	Location-Based Services
LDV	Light Duty Vehicles
MCIP	Meteorology-Chemistry Interface Processor
MOVES	Motor Vehicle Emission Simulator
NEI	National Emission Inventory
NO_X	Oxides of Nitrogen
ONI	Off-network Idle Activity
PM	Particulate Matter
PM_{10}	Particulate Matter under 10 Microns
$PM_{2.5}$	Particulate Matter under 2.5 Microns

Acronym	Definition
RMAR	Relative Mileage Accumulation Rate
RPD	RatePerDistance
RPH	RatePerHour
RPHO	RatePerHourONI
RPP	RatePerProfile
RPS	RatePerStart
RPV	RatePerVehicle
SCC	Source Classification Code
SIP	State Implementation Plans
SHEI	Source Hours Extended Idling
SHI	Source Hour Idling
SHO	Source Hour Operating
SHP	Source Hour Parked
SMOKE	Sparse Matrix Operator Kernel Emissions
ST	Single-unit Truck
SUT	Source Use Type
Swkd	Summer Weekday
TCEQ	Texas Commission on Environmental Quality
TDM	Travel Demand Model
TIF	Total Idle Fraction
TPP	Transportation Planning and Programming
TRE	Trinity Railway Express
TTI	Texas A&M Transportation Institute
TxDOT	Texas Department of Transportation
TxLED	Texas Low Emission Diesel
TxDMV	Texas Department of Motor Vehicles
U.S.	United States
VMT	Vehicles Miles of Travel
VIEWS	Vehicle Inventory and Use Survey
VCC	Vehicle Classification Count
VHT	Vehicles Hour Traveled
VIUS	Vehicle Inventory and Use Survey
VOC	Volatile Organic Compounds
VPOP	Vehicle Population
WRF	Weather Research and Forecasting

EXECUTIVE SUMMARY

The main objectives of this study were to assess and improve the Emission Inventory (EI) development process and to identify alternative data sources appropriate for use in the Motor Vehicle Emission Simulator (MOVES) inventory mode County Database (CDBs) for the state of Texas. The research team thoroughly reviewed the regulatory requirements and associated guidance documents. The sensitivity assessment was conducted to determine the MOVES3 activity input parameters with the most significant effect on emissions by assessing the recently completed 2020 Air Emissions Reporting Requirements (AERR) on-road modeling files and the MOVES3 trend information being developed as another recent project[1]. The suitability of using MOVES3 default modeling parameters applied by the Environmental Protection Agency (EPA) in the National Emission Inventory (NEI) CDBs was assessed in place of Texas-specific activity data. The results from the sensitivity assessment were used to identify potential on-road vehicle activity data sources for use in future regulatory and SIP EI development projects.

Figure 1 gives an overview of the project's workflow. This report documents all the pertinent activities performed by Texas A&M Transportation Institute (TTI) to complete this study. Tasks 1 through 6 of the project's Grant Activity Description (GAD) detailed the work and products. Task 1 is the TTI's GAD and Quality Assurance Project Plan (QAPP) as approved by TCEQ. Task 2 is TTI's monthly progress reports to the Texas Commission on Environmental Quality (TCEQ). The technical work and documentation requirements are detailed in Tasks 3 through 6:

- Task 3 – Literature and Methodology Review.
- Task 4 - Sensitivity Analysis of MOVES3 Activity Inputs and Defaults.
- Task 5 – Identification of Potential Texas-Specific Activity Data Sources.
- Task 6 - Draft and Final Reports.

[1] MOVES3 On-road Trend Emissions Inventories for 1990 and 1999 through 2060. Final deliverables submitted to TCEQ in April 2023.

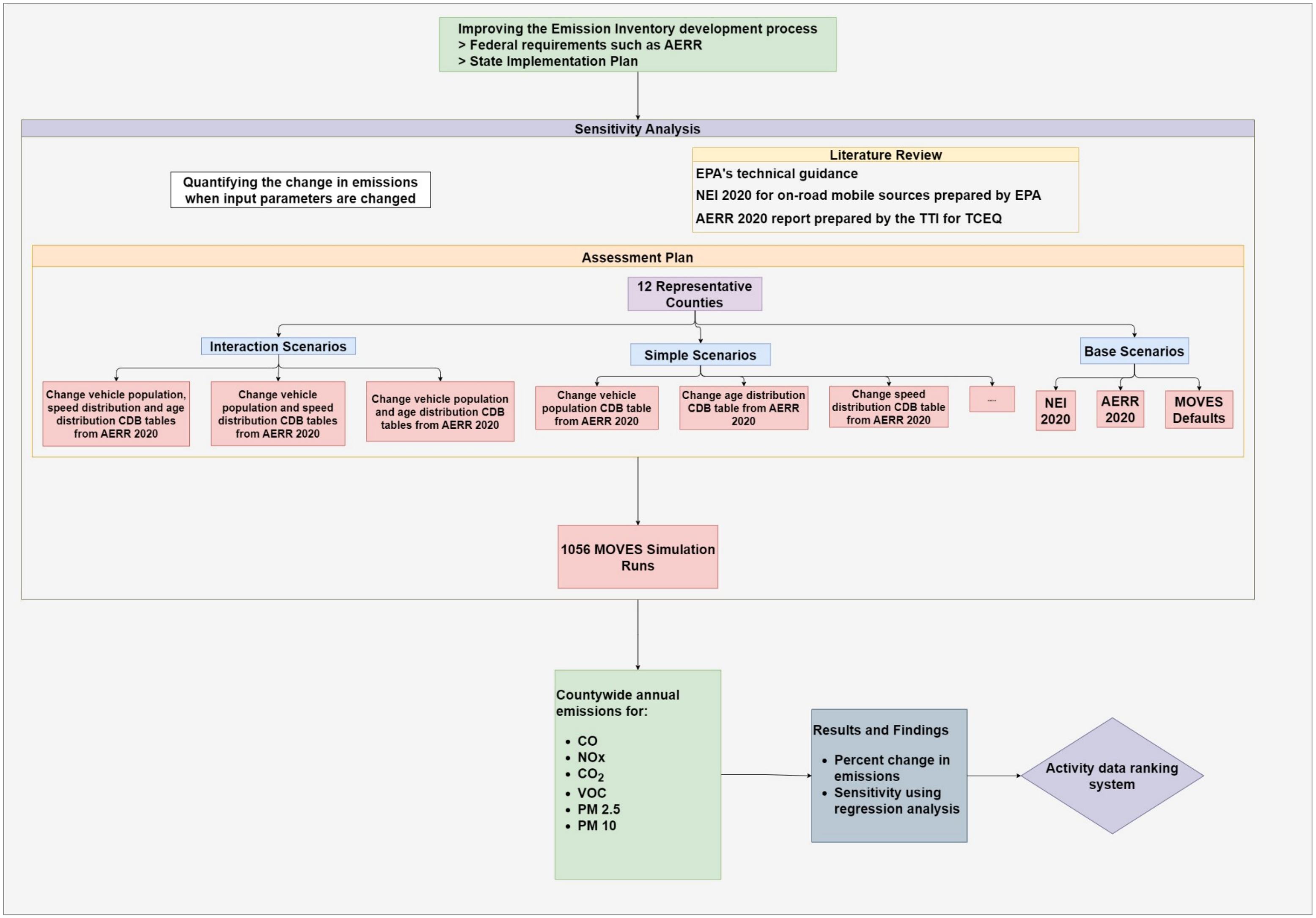


Figure 1. Project Workflow Overview

As a part of task 3, the TTI research team extensively reviewed methodologies used in developing 2020 AERR and NEI CDBs and identified key differences between the NEI CDBs and the CDB developed by TCEQ with local activity data sources. Critical inputs such as source type population, age distribution, speed distributions, and meteorology and fuel properties were considerably different and were the candidates of the parameters for the sensitivity assessment. For example, TTI estimated the source type population from Texas Department of Motor Vehicle (TxDMV) registration data, while EPA estimated it from HIS vehicle registration data. TTI used speed estimation equations based on the travel demand model (TDM), MOVES defaults for the speed distribution, while EPA used telematics data. The differences in the data sources led to variability in the CDBs developed by TTI and the NEI, highlighted in this task.

For task 4, the TTI research team developed an assessment plan in consultation with TCEQ to test the sensitivity of MOVES emissions results with respect to various inputs. The sensitivity scenarios were classified into three categories – base scenarios, simple scenarios, and interaction scenarios. Base scenarios consisted of the benchmarks, including AERR 2020, NEI 2020, and MOVES defaults. As a part of the base scenarios analysis, several input parameters were initially tested to assess the level of emissions differences in the AERR 2020 resulting from swapping, in particular, NEI input values (individually for fuels, meteorology, hotelling activity, and allocation of Vehicles Miles of Travel (VMT) by vehicle category; and for fuels and meteorology together) while keeping other AERR CDB inputs constant. In the simple and interaction scenarios, one or more inputs were changed in AERR 2020 CDBs, keeping the other inputs the same. The TTI research team categorized MOVES source types[2] into four groups: source types 21, 31, and 32 as Group 1; source types 41, 42, and 43 as Group 2; source types 51, 52, 53, and 54 as Group 3; and source types 61 and 62 as group 4. The emission difference between AERR 2020 and the considered scenario was utilized to determine the sensitivity of that input. A key feature of the assessment plan, which is different from other related studies, is that inputs used in the sensitivity scenarios were directly derived from actual data sources (NEI and AERR); hence, the sensitivity analysis is bounded within the range of realistic values. The team performed a total of 1056 simulation runs on a selection of twelve representative counties and various scenarios. Key findings from the study are as follows.

[2] Source types are vehicle categories. For example, MOVES assign source type 21 to passenger cars. For the list of source types, refer to https://github.com/USEPA/EPA_MOVES_Model/blob/master/docs/MOVES4CheatsheetOnroad.pdf

- For considered pollutants, NEI showed higher emissions than AERR. When MOVES default inputs were used, emissions were considerably higher.
- Emission estimates using MOVES defaults were neither close to NEI nor AERR. Hence, using MOVES defaults is not recommended. It is crucial to develop county-specific inputs instead of using defaults representing national averages.
- For all the source type groups, the TTI research team observed that fuel inputs, meteorology inputs, and road distribution have negligible effects on pollutant.
- Higher emissions were observed when VMT were distributed according to MOVES source types instead of Highway Performance Monitoring System (HPMS) source types[3] while keeping the total VMT the same. Hence, it was essential to focus on how the VMT was distributed and entered for various source types.
- Emissions were more sensitive to the parameters of light-duty vehicles (LDVs) and combination trucks (CTs) as compared to buses and single-unit trucks (STs); hence, more attention should be given to improving inputs for these source types.
- The population of LDVs was found to have a substantial (>20 percent) impact on emissions and hence was concluded to be a highly sensitive parameter. When only changed, the age and speed distribution had a low (0-5 percent) to medium (5-20 percent) impact on emissions. However, when age, population, and speed distribution were changed simultaneously, their interaction effect substantially impacted emissions (>20 percent).
- For combination trucks, the age distribution was the most critical input with a medium (5-20 percent) impact on emissions.

For task 5, based on the study's findings, the TTI Research team developed a ranking system for potential data sources that could be used and explored to improve current emission inventories. These considered datasets include the TxDMV data, his Markit–Polk registration data (S&P Global, 2023; S&P Global Mobility, 2023), and 2021 Vehicle Inventory and Use Survey (VIUS) data (U.S. Department of Transportation, Bureau of Transportation Statistics & U.S. Department of Commerce, U.S. Census Bureau, 2023). TxDMV data has been in use in Texas for over 20 years. It is reliable for regional emission estimateshisIHS Markit–Polk dataset can provide more information, such as

[3] For HPMS source types, refer to https://github.com/USEPA/EPA_MOVES_Model/blob/master/docs/MOVES4CheatsheetOnroad.pdf

vehicle make, model, Gross Vehicle Weight Rating (GVWR), fuel type, vehicle type, and registration type. 2021 VIUS data can provide information for differentiating activity of some source use types, such as light commercial trucks and personal trucks. For population and age input tables, the TxDMV dataset is good. TCEQ can consider purchasing IHS Markit-Polk data and analyzing the 2021 VIUS dataset. TDM, MOVES defaults, and NEI tables can be used for the average speed distribution input table. Alternatively, TCEQ can consider buying telematics data.

For task 6, TTI wrote this report documenting all of the pertinent activities related to the completion of this study, including a summary of the results of the literature and methodology review, a description of methodology and model input information used in the sensitivity analysis; and an outline of processes used to gather and evaluate Texas-specific data sources for the development of annual inventory mode MOVES3 CDBs.

1 INTRODUCTION

The Texas Commission on Environmental Quality (TCEQ) is responsible for developing emission inventories (EIs) of on-road emission sources to support state implementation plan (SIP) development to meet federal EI requirements such as the Air Emissions Reporting Requirements (AERR) by the United States Environmental Protection Agency (EPA) and for emissions trend analyses in Texas. The Texas A&M Transportation Institute (TTI) assisted TCEQ with developing the MOVES CDBs for 2020 on-road mobile source AERR EIs (2020 AERR) for the 254 Texas counties based on various Texas-specific data sources[4]. The EPA uses the 2020 AERR submittals from different states and other data sources to build the 2020 National Emission Inventory (NEI).

In the context of the 2020 NEI, the EPA developed the 2020 NEI county databases (CDBs) for each county. For the state of Texas, the EPA partially used the data submitted by TCEQ to create the corresponding NEI CDBs[5]. There were some critical differences in the input data between the CDBs submitted by TCEQ and the final NEI CDBs developed and used by the EPA. These differences in the inputs led to variability in the final emission estimates. Emissions estimated by the EPA NEI CDBs were higher than those from TCEQ CDBs for the major pollutants, including carbon dioxide (CO_2), carbon monoxide (CO), particulate matters (PM), oxides of nitrogen (NO_X) and volatile organic compounds (VOC). Statewide total CO_2 emissions were around 15 percent higher, and total oxides of nitrogen NO_X emissions were around 23 percent higher. For state agencies, it is essential to understand the impact of the differences in activity inputs on the difference in emissions estimates.

The main objectives of this study were to assess and improve the EI development process and to identify alternative data sources appropriate for use in AERR Motor Vehicle Emission Simulator (MOVES) inventory mode CDBs for the state of Texas. The research team thoroughly reviewed the regulatory requirements and associated guidance documents. The sensitivity assessment was conducted to determine the MOVES3 activity input parameters with the most significant effect on emissions by assessing the recently completed 2020 AERR on-road modeling files and the MOVES3

[4] 2020 On-road Mobile Source Annual, Summer Weekday and Winter Weekday Emissions Inventories – Final Report. https://www.tceq.texas.gov/downloads/air-quality/research/reports/on-road/5822111197019-20220425-tti-2020-aerr.pdf

[5] Instructions and Best Practices for Development and Submittal of Onroad Inputs for the 2020 National Emissions Inventory. https://www.epa.gov/system/files/documents/2021-07/instructions-and-best-practices-for-submitting-mobile-source-data-20210722.pdf

trend information being developed as another recent project[6]. The suitability of using MOVES3 default modeling parameters applied by the EPA in the NEI CDBs was assessed in place of Texas-specific activity data. The results from the sensitivity assessment were used to identify potential on-road vehicle activity data sources for use in future regulatory and SIP EI development projects.

This report documents all the pertinent activities performed by TTI to complete this study. Tasks 1 through 6 of the project's Grant Activity Description (GAD) detailed the work and products. Task 1 is the TTI's GAD and Quality Assurance Project Plan (QAPP) as approved by TCEQ. Task 2 is TTI's monthly progress reports to TCEQ. The technical work and documentation requirements are detailed in Tasks 3 through Task 6:

- Task 3 – Literature and Methodology Review.
- Task 4 - Sensitivity Analysis of MOVES3 Activity Inputs and Defaults.
- Task 5 – Identification of Potential Texas-Specific Activity Data Sources.
- Task 6 - Draft and Final Reports.

Corresponding to the work from Task 3 through Task 5, the report is divided into the following chapters (Chapter 2 to Chapter 6).

- Chapter 2 discusses the research team's literature review (Task 3). The team reviewed the MOVES3 technical guidance[7], associated MOVES3 documentation provided by the EPA, and the 2020 AERR NEI documentation about the methodology used to produce annual MOVES3 inventory mode CDBs. The findings from the literature review were summarized, and the information gathered from this review helped identify where the method for developing the annual MOVES CDBs may be updated or improved. In addition, the chapter briefly summarizes the reviews on the methodologies used by other states.
- Chapter 3 includes the developed assessment plan with a detailed list of the activity parameters assessed as part of the Task 4 work. Excluding well-established parameters like meteorology and fuels, the parameters in the test plan are source type population, age distribution, speed distribution, hotelling hours, and vehicle miles of travel (VMT) road type distribution. The sensitivity analysis of the MOVES3

[6] MOVES3 On-road Trend Emissions Inventories for 1990 and 1999 through 2060. Final deliverables submitted to TCEQ in April 2023.

[7] MOVES Onroad Technical Reports. https://www.epa.gov/moves/moves-onroad-technical-reports.

inventory mode CDB inputs will be conducted using the most recent 2020 AERR and the MOVES3 trend information developed by TTI. The expected value changes and the explanations on the nature of the activity parameter would be analyzed to identify which input parameters have the greatest sensitivity and impact on final emission estimates. The comparison with 2020 NEI CDBs and 2020 AERR CDBs would be able to determine how appropriate the use of MOVES3 default data when compared with Texas-specific local data.

- Chapter 4 discusses details of the sensitivity study, as well as the results and findings of the study as part of Task 4 work. A total of 1056 MOVES runs were carried out for twelve representative counties. The first section delves into how emissions sensitivity varies with different datasets. It contrasts emission discrepancies across various scenarios, mainly focusing on substituting NEI or MOVES default datasets for AERR's Commercial Database data. Key areas of difference include population, age and speed distribution, meteorology, fuel inputs, VMT distributions, and hotelling inputs. This section primarily emphasizes the relative emission disparities across scenarios rather than quantifying input differences. The final section examines how different input variables, particularly vehicle population, affect pollutant emissions. The analysis, focusing on group 1 vehicles (source types 21, 31, and 32), revealed notable variations in emission responses to population changes among these types. Regression models applied to data from various counties confirmed the statistical significance of these findings, suggesting the model's applicability across diverse county characteristics. Key observations and conclusions were summarized in the final section of the chapter.

- Chapter 5 delves into a detailed explanation of the results from Chapter 4, exploring the relationship between emissions and activities and providing insights into why emissions from the NEI are different (mostly higher) despite having the same total HPMS VMT as TCEQ’s AERR 2020 submittal as part of the Task 4 work.

- Based on the findings in Chapters 4 and 5, Chapter 6 presents the Task 5 work conducted by the team by introducing a ranking system and resulting matrix to prioritize data sources for improving sensitive activity inputs and emission estimates in the MOVES model. This system evaluates various datasets based on their weighted ranking attributes with the potential to enhance the accuracy of emission inventories.

- Chapter 7 summarizes the findings from the study, followed by a reference section and Appendix A.

2 LITERATURE REVIEW

The following section discusses the literature review (Task 3) conducted by the research team. The team reviewed the MOVES3 technical guidance, associated MOVES3 documentation provided by the EPA, and the 2020 AERR NEI documentation about the methodology used to produce annual MOVES3 inventory mode CDBs. The findings from the literature review were summarized, and the information gathered from this review helped identify where the method for developing the annual MOVES CDBs may be updated or improved. In addition, the chapter briefly summarizes the reviews on the methodologies used by other states.

The literature reviewed as a part of task 3 includes:

- EPA's technical guidance. Office of Transportation and Air Quality. 2020.
- MOVES3 Technical Guidance: Using MOVES to Prepare Emission Inventories for State Implementation Plans and Transportation Conformity. (USEPA. 2021). https://nepis.epa.gov/Exe/ZyPDF.cgi/P1010LY2.PDF?Dockey=P1010LY2.PDF
- MOVES Onroad Technical Reports. (USEPA. 2021) https://www.epa.gov/moves/moves-onroad-technical-reports#moves3.
- Instructions and Best Practices for Development and Submittal of Onroad Inputs for the 2020 National Emissions Inventory (NEI). (USEPA. 2021). https://www.epa.gov/system/files/documents/2021-07/instructions-and-best-practices-for-submitting-mobile-source-data-20210722.pdf.
- *2020 National Emissions Inventory Technical Support Document: Onroad Mobile Sources.* U.S. Environmental Protection Agency (USEPA. 2023)
- *2020 On-Road Mobile Source Annual, Summer Weekday, and Winter Weekday Emissions Inventories.* Texas A&M Transportation Institute (Venugopal, Madhusudhan. 2022)

The literature review section is divided into six subsections. Section 2.1 includes a detailed look at the MOVES3 modeling framework, emphasizing the importance of accurate activity inputs in inventory and emission rate modes for reliable EI development. It also discusses the various domains of MOVES inputs, such as activity characteristics (on-network and off-network), fleet characteristics, fuel properties, and meteorology, and the significance of using localized data to enhance accuracy. Section 2.2 describes the EPA's approach to calculating the 2020 NEI, including using MOVES rates mode, telematics data, and representative county methodology for efficiency and accuracy. Section 2.3 overviews the methods used by TCEQ for the 2020 AERR, focusing on the detailed hourly link-level analysis and the incorporation of various local and default data sources. Section 2.4 briefly reviews the methodologies used by different states, like Pennsylvania and Kentucky, and metropolitan areas, like Atlanta and Chicago, for developing input CDBs for MOVES. Section 2.5 provides a comparison in terms of the Texas county-level MOVES input database tables in 2020 NEI developed by EPA and AERR 2020 TCEQ submittal developed by TTI and their impact on emission estimates. Section 2.6 summarizes the entire literature review work by the research team.

2.1 Emission Modeling Overview and Guidance

This section includes a detailed look at the MOVES versions and modeling framework, emphasizing the importance of accurate activity inputs in inventory and emission rate modes for reliable EI development. It also discusses the various domains of MOVES inputs, such as activity characteristics (on-network and off-network), fleet characteristics, fuel properties, and meteorology, and the significance of using localized data to enhance accuracy.

2.1.1 MOVES Versions

The EPA's official emission model is the MOVES. It calculates emissions for mobile sources, providing national, county, and project-level estimates for criteria air pollutants, greenhouse gases, and air toxics. MOVES4, released in September 2023, is the latest version of EPA's MOVES emission model that replaced the previous MOVES 3.1. EPA's release of MOVES4 for SIPs and transportation conformity analyses becomes effective on September 12, 2023. This marks the beginning of a two-year transportation conformity grace period, concluding on September 12, 2025. After this deadline, using MOVES4 becomes mandatory for new transportation conformity analyses outside of California for regional emissions and hot-spot analyses.

MOVES4 was released toward the end of this study. At the same time, most works were developed using MOVES3 as the input parameters developed in this study for 2020 NEI and 2020 AERR are compatible with MOVES3; the protocols and scripts developed in this study were not altered with the version change.

2.1.2 Overview of MOVES Modeling Modes: Emission Inventory (EI) and Emission Rate (ER) Methods

Modeling emissions in the context of environmental assessment and policy-making involves complex methodologies to estimate emissions from various sources accurately. This is particularly true for on-road mobile sources, which significantly contribute to overall emissions. Two primary methods are used for such modeling: Emission Inventory (EI) and Emission Rate (ER) methods, both of which are integral components of the MOVES3 model[8]:

- *Inventory Mode*: MOVES3 calculates this mode's total emission or energy estimates. This is done using input data such as VMT, vehicle population, and other inputs fed into the model by users. The output is a comprehensive view of total emissions or energy consumption for a given area or scenario. The following activity and emission quantity tables will be populated in this mode:
    - *movesactivityoutput*: vehicle on-network and off-network activity[9].
    - *movesoutput*: vehicle on-network and off-network emission quantity and energy consumptions.
- *Emission Rates Mode*: Contrasting the Inventory mode, the Emission Rates mode of MOVES3 is focused on determining emission rates per unit of activity. The model provides emissions or energy consumption rates per unit, such as per mile traveled. These emission rates are then applicable for detailed analysis and can be integrated with activity data to construct an emission inventory. The following emission rate tables will be populated in this mode:
    - *rateperdistance*: running exhaust, idling, crankcase running, brake wear, tire wear, running evaporative emission rates, and energy consumption rates associated with on-network vehicle activity.
    - *rateperhour:* hotelling and crankcase hotelling emission rates associated with hotelling activity.

[8] Overview of EPA's Motor Vehicle Emission Simulator (MOVES3). https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P1011KV2.pdf

[9] On-network vehicle activity is a term in the TTI EI process to describe all the vehicle activities when the vehicle is running on the network. Off-network vehicle activity is a term in TTI EI process to describe all the vehicle activities when the vehicles are not running, including off-network idle activity, engine start activity, hotelling activity, park activity.

- *ratepervehicle:* Start, hotelling, evaporative, and refueling processes associated with off-network vehicle activity.
- *rateperprofile:* Evaporative vapor venting emission rates associated with off-network vehicle activity.
- *rateperstart*: start exhaust and crankcase start emission rates associated with engine start activity.

Whether using the Inventory or Emission Rates mode, the accuracy of vehicle activity inputs is critical. These inputs influence the calculated on-road and off-network emission rates. This is because MOVES3 uses these inputs to proportionate vehicles' running and resting activities, which in turn impacts emissions related to starts, evaporative processes, and extended idling. The significance of having precise activity estimates cannot be overstated. Regardless of the chosen mode in MOVES3, these estimates are essential for modeling a scenario accurately.

2.1.3 MOVES3 CDB Input Guidance

The EPA's guidance for using the MOVES model in on-road vehicle emissions estimation emphasizes the county-scale approach for SIPs and regional conformity analyses, aligning with the 2020 NEI and 2020 AERR methodologies. While MOVES-Onroad includes a default U.S.-wide database covering meteorology, fleet, activity, fuel, and emissions control data, this data may not always be the most current or optimal for specific counties. Local data is generally preferred for accurate emissions estimates, though reliance on default data is acceptable in particular scenarios. The quality of inputs significantly influences model outcomes, and specific county data enhances the precision of MOVES emissions estimates. In cases where MOVES default data is not the most up-to-date for a county, its use is still recommended when it does not compromise result quality. The MOVES inputs can be classified into the following broad domains[10]:

- On-network activity characteristics:
    - VMT by calendar year and vehicle type.

[10] Instructions and Best Practices for Development and Submittal of Onroad Inputs for the 2020 National Emissions Inventory (NEI). (USEPA. 2021). https://www.epa.gov/system/files/documents/2021-07/instructions-and-best-practices-for-submitting-mobile-source-data-20210722.pdf.

   - Average speed distributions.
   - VMT temporal distributions on the hour of the day, day of the week, and month of the year.
   - VMT spatial distributions of source type by road type.
- Off-network activity characteristics:
   - Vehicle populations by calendar year.
   - hotelling activity.
   - Engine start activity.
   - Off-network idle activity.
- Fleet characteristics:
   - Vehicle age-related characteristics.
   - Inspection and Maintenance (I/M) coverage.
   - Alternative Vehicle Fuel and Technology (AVFT).
- Fuel Properties and Supply: properties of the fuel types used, the different fuel types used in a region, and the distribution of fuels used.
- Meteorology: details on barometric pressure, relative humidity, and temperature.
- Supporting tables such as state, year, county, and allocation information.

Table 1 lists the essential CDB tables and describes their content.

Table 1. MOVES3 CDB Table Descriptions

CDB Table Type	CDB Table	Description of Content
On-road Activity Characteristics	hpmsvtypeyear	VMT by calendar year by Highway Performance Monitoring System (HPMS) vehicle type.
	sourcetypevmtyear	VMT by calendar year by source use type.
	hpmsvtypeday	VMT by day by HPMS vehicle type.
	sourcetypevmtdday	VMT by day by source use type.
	avgspeeddistribution	Average speed distributions by 16 MOVES speed bins.
	dayvmtfraction	Fractions to distribute VMT between day types
	hourvmtfraction	Fractions to distribute VMT across hours in a day
	monthvmtfraction	Fractions to distribute VMT across 12 months of the year
	roadtypedistribution	Fractions to distribute VMT across the road types

CDB Table Type	CDB Table	Description of Content
Vehicle population	sourcetypeyear	Vehicle populations by calendar year
Fleet Characteristics	avft	Fuel type fractions
	imcoverage	Description of the inspection and maintenance program
	sourcetypeagedistribution	Distribution of vehicle population by age
Hotelling	hotellinghoursperday	Total hours of hoteling per day, including total time spent in all of the four operating modes defined in the hotellingactivitydistribution table.
	hotellingactivitydistribution	Fraction of hoteling hours in which the power source is the main engine, diesel APU, electric APU, or engine-off
	hotellingagefraction	fraction of hoteling hours by age (e.g., to account for newer trucks having more hoteling activity). Fractions should sum to 1.0.
	hotellinghourfraction	fraction of hoteling in hours of the day. Fractions should sum to 1.0 for each day type.
	hotellingmonthadjust	Adjustment factors to vary hoteling activity between different months. A factor of 1.0 for each month will model a situation where annual hoteling hours are evenly divided among months. A value of 1.1 for month ID 1 will increase the hoteling hours per day in January by 10%.
Engine Starts	startsperdaypervehicle	total number of starts per vehicle in a day by source type
	startsageadjustment	numbers reflecting relative differences in the number of vehicles starts by age.
	startsmonthadjust	*Fractions to vary the vehicle start by month of year.*
	startshourfraction	fractions to distribute starts across hours in a day
Fuel Properties and Supply	fuelsupply	Fuel differences by month of the year
	fuelformulation	Fuel properties
	fuelusagefraction	Fraction of the time that ethanol (E85) vs. gasoline is used in flex-fuel engine vehicles
Meteorology: County Pressure	county	Description of the county, including barometric pressure
Meteorology: Hourly Temperature & Humidity	zonemonthhour	Temperature and relative humidity values
Off-network Idling	totalidlefraction	Fraction of vehicle operating time when speed is zero
Support table	state	Description of the state
Support table	year	Year of the database
Support table	zone	Allocations of starts extended idle, and vehicle hours parked to the county.
Support table	zoneroadtype	Allocation of source hours operating to the county

The following sections include recommendations from the EPA to prepare the emission rate mode or inventory mode CDBs from the above MOVES inputs.

2.1.4 Activity Characteristics: On-network

This section discusses the key on-network emissions source activity inputs and guidance on using local data to develop these inputs.

2.1.4.1 VMT by Calendar Year and Vehicle Type.

The EPA advises that local entities, such as metropolitan planning organizations (MPOs) and state departments of transportation (DOTs), create their own VMT estimates. The primary source for these estimates is often travel demand models. These models, utilized by transportation modelers in MPOs and state DOTs, are adjusted to align with the HPMS estimates or other local vehicle count data. This practice ensures the generation of consistent VMT estimates across various roadway functional classes within the HPMS, which are crucial for SIP analysis.

The latest version of the MOVES3 offers flexibility in data input, allowing for either annual VMT or daily VMT. The EPA recommends that entities with access to average annual daily VMT data use the daily VMT input option for enhanced precision. When VMT is input by the HPMS class to the HPMSVtypeday table or the *HPMSVtypeyear* table, MOVES automatically allocates VMT to different source types based on default assumptions. Alternatively, users who can develop VMT data specifically by MOVES source types can input this data directly as imported to *sourceTypeYearVMT* table or *sourceTypeDayVMT* table. This approach bypasses the default allocation process, potentially leading to more accurate source-type VMT estimates. Table 2 lists the MOVES source type information with the associated HPMS vehicle type in MOVES3. Both methods are deemed acceptable for SIP and conformity assessments. However, it should be noted that discrepancies may arise between the results of the default HPMS class-to-source type allocations in MOVES and those derived from user-supplied source-type VMT data.

Table 2. Source Use Type Information with Associated HPMS Vehicle Type in MOVES3

sourceTypeID	HPMSVtypeID	sourceTypeName
11	10	Motorcycle
21	25	Passenger Car
31	25	Passenger Truck
32	25	Light Commercial Truck
41	40	Other Buses
42	40	Transit Bus

sourceTypeID	HPMSVtypeID	sourceTypeName
43	40	School Bus
51	50	Refuse Truck
52	50	Single Unit Short-haul Truck
53	50	Single Unit Long-haul Truck
54	50	Motor Home
61	60	Combination Short-haul Truck
62	60	Combination Long-haul Truck

2.1.4.2 VMT Spatial and Temporal Distribution

As listed in Table 1, MOVES allows users to define VMT distributions across various day types, months, hours, and road types. The temporal distribution tables help users distribute the VMT by calendar year by vehicle type to different months of the year, days of the week, and hours of the day. The other critical aspect of this distribution is the allocation of VMT across different road types, which significantly impacts emissions from on-road mobile sources. Table 3 lists the road type definitions in the MOVES onroad model.

Table 3. Road Type Information in MOVES3

roadTypeID	roadTypeDescription
1	Off-network
2	Rural Restricted Access
3	Rural Unrestricted Access
4	Urban Restricted Access
5	Urban Unrestricted Access

EPA mandates that states generate and apply their specific VMT estimates per road type. In MOVES, the Road Type Distribution table within the input database records the VMT distribution by road type for each vehicle source type, such as the proportion of passenger car VMT on each road type.

For inventory modeling purposes in SIPs and regional conformity analyses, the VMT fractions by road type must align with the latest transportation planning data. Recognizing the challenge in developing local road type distributions for all 13 vehicle source types listed in Table 2, the EPA permits states to apply a uniform distribution across all source types within an HPMS vehicle class if specific data is unavailable.

In scenarios where the Inventory mode is applied, MOVES will produce the activities based on the road type distributed VMTs and calculate emission quantities for each source type across different MOVES speed bins. Table 4 lists the speed bins being used in the MOVES3 model. In the emission Rates mode, MOVES will generate a table outlining running emission rates for the road type from 2 to 5 (see Table 3) if the road type distribution is not 0. While these distributions do not directly influence the on-network emission rates, they are vital for Emission Rate runs involving off-network processes. MOVES utilizes these inputs to determine the relative volumes of on-network and off-network activity, consequently impacting the rates of off-network processes.

Table 4. Speed Bin information with Descriptions in MOVES3

avgSpeedBinID	avgBinSpeed	avgSpeedBinDesc
1	2.5	speed < 2.5mph
2	5	2.5mph <= speed < 7.5mph
3	10	7.5mph <= speed < 12.5mph
4	15	12.5mph <= speed < 17.5mph
5	20	17.5mph <= speed < 22.5mph
6	25	22.5mph <= speed < 27.5mph
7	30	27.5mph <= speed < 32.5mph
8	35	32.5mph <= speed < 37.5mph
9	40	37.5mph <= speed < 42.5mph
10	45	42.5mph <= speed < 47.5mph
11	50	47.5mph <= speed < 52.5mph
12	55	52.5mph <= speed < 57.5mph
13	60	57.5mph <= speed < 62.5mph
14	65	62.5mph <= speed < 67.5mph
15	70	67.5mph <= speed < 72.5mph
16	75	speed >= 72.5mph

2.1.4.3Speed Distribution

Vehicle-specific power (VSP), speed, and acceleration significantly affect vehicle emissions. Inherited from the Mobile 6 model, the MOVES model applied the concept of drive cycles and operating modes to describe the vehicle condition in on-network emissions modeling. A drive cycle[11] (called drive schedule) is a second-by-second

[11] For detailed information of MOVES default drive cycles, see DriveSchedule table in the MOVES default database.

vehicle speed trajectory used in the MOVES model to determine the appropriate operating mode distributions for emissions and energy consumption processes. Table 5 lists the operating mode bin definitions for running emissions.

Table 5. Operating Mode Bin information with Descriptions in MOVES3

Operating mode ID	VSP (kw/Tonne)	Operating mode description	Vehicle speed (v, mph)	Acceleration (a, mph/sec)
0		Braking		a<= -2.0 or (for consecutive three seconds a<-1.0)
1		Idle	-1<= v <1	
11	VSP < 0	Coast	0<= v <25	
12	0<= VSP <3	Cruise/Acceleration	0<= v <25	
13	3<= VSP <6	Cruise/Acceleration	0<= v <25	
14	6<= VSP <9	Cruise/Acceleration	0<= v <25	
15	9<= VSP <12	Cruise/Acceleration	0<= v <25	
16	VSP >= 12	Cruise/Acceleration	0<= v <25	
21	VSP < 0	Coast	25<= v <50	
22	0<= VSP <3	Cruise/Acceleration	25<= v <50	
23	3<= VSP <6	Cruise/Acceleration	25<= v <50	
24	6<= VSP <9	Cruise/Acceleration	25<= v <50	
25	9<= VSP <12	Cruise/Acceleration	25<= v <50	
27	12<= VSP <18	Cruise/Acceleration	25<= v <50	
28	18<= VSP <24	Cruise/Acceleration	25<= v <50	
29	24<= VSP <30	Cruise/Acceleration	25<= v <50	
30	VSP >= 30	Cruise/Acceleration	25<= v <50	
33	VSP < 6	Cruise/Acceleration	v>=50	
35	6<= VSP <12	Cruise/Acceleration	v>=50	
37	12<= VSP <18	Cruise/Acceleration	v>=50	
38	18<= VSP <24	Cruise/Acceleration	v>=50	
39	24<= VSP <30	Cruise/Acceleration	v>=50	
40	VSP >= 30	Cruise/Acceleration	v>=50	

In county scale MOVES modeling, the average speed distribution table, *avgSpeedDistribution*, helps the model select the corresponding drive cycles and the operating mode distribution calculated from the drive cycles and then calculate the emission rates. For SIP development and regional conformity analyses, where activity is averaged over various driving patterns, a local speed distribution by road type and source type is necessary. The Average Speed Distribution Importer in MOVES calls for a speed distribution in 16-speed bins by each road type, source type, and hour of the day included in the analysis. The EPA recommends users develop the most detailed local speed information that is reasonable to obtain. However, EPA acknowledges that average speed distribution may not be available at the level of detail that MOVES allow. One recommended approach for estimating average speeds is to post-process the output from a travel demand model. In most transportation models, speed is estimated primarily to allocate travel across the roadway network. An alternative approach to developing a local average speed distribution is to process on-vehicle Global Positioning System (GPS) data.

2.1.5 Activity Characteristics: Off-network

This section discusses the key off-network emissions source activity inputs and guidance on using local data to develop these inputs.

2.1.5.1Source type population

MOVES uses the source type (vehicle type) population in the SourceTypeYear table to calculate start and evaporative emissions. Start and evaporative emissions depend more on how many vehicles are parked and started than on how many miles they are driven. In MOVES, start and resting evaporative emissions are related to the population of vehicles in an area. Because vehicle population directly determines start and evaporative emissions, users must develop local data for this input. MOVES categorizes vehicles into thirteen source types, subsets of five HPMS vehicle types in MOVES, as shown in Table 2.

EPA recommends that states should be able to develop population data for many of these source type categories from state motor vehicle registration data (e.g., motorcycles, passenger cars, passenger trucks, light commercial trucks) and local transit agencies, school districts, bus companies, and refuse haulers (intercity, transit, and school buses, and refuse trucks).

2.1.5.2Engine Start Activity

MOVES3 uses a set of tables in its default database to estimate daily vehicle start activities, soak times, and their temporal distributions. These tables include:

- *StartsPerDayPerVehicle,*
- *StartsAgeAdjustment,*
- *StartsHourFraction,*

- *StartsMonthAdjust, and*
- *StartsOpModeDistribution.*

The *StartsPerDayPerVehicle* table contains a factor that calculates daily starts when multiplied by the total number of vehicles of a specific type. This factor represents the average daily starts for each vehicle type and day type (weekday or weekend). MOVES3 introduces adjustments for vehicle age, recognizing that start activity decreases as vehicles age. For instance, light-duty vehicles show higher start activity at age 0 than a fleet-average, with a significant decrease by age 30.

In MOVES3, the effect of vehicle age on engine starts is accounted for using age adjustment factors in the *StartsAgeAdjustment* table. Based on mileage accumulation rates, these factors assume constant starts per mile over a vehicle's lifespan. As vehicles age and travel fewer miles, they have fewer starts. The ageAdjustment factor is set to one at age zero and decreases with age. MOVES scales these factors to maintain average begins as in the *StartsPerDayPerVehicle* table. The *StartsMonthAdjust* table adjusts daily starts for monthly variations, and the *StartsHourFraction* table distributes starts across hours of the day. The *StartsOpModeDistribution* table deals with engine start soak times. MOVES allows users to update these tables with more specific data.

2.1.5.3Hotelling Activity

In MOVES, "hotelling" refers to heavy-duty truck drivers spending in their vehicles during mandatory rest periods, often in built-in sleeping berths, during long hauls. This time is included in MOVES to account for energy consumption and pollution from amenities like air conditioning and heating. These amenities can be powered by the truck's main engine, auxiliary power units, or external power hookups at truck stops. In MOVES, only the long-haul combination truck type (sourceTypeID 62 in Table 2) is considered for hotelling activities, all being diesel-fueled. Other vehicle types in MOVES are set to have zero hotelling activity.

Table 6 shows the hotelling operating modes available in MOVES. The *HotellingActivityDistribution* contains the MOVES default values for the distribution of hotelling activity to the operating modes. In Inventory Mode, the users can update the hotelling tables based on local information on vehicle start activity in the *HotellingHoursPerDay* table.

Table 6. Hotelling Activity Operating Modes in MOVES3

OpModeID	Description
200	Extended Idling of Main Engine
201	Hotelling Diesel Auxiliary Power Unit (APU)
203	Hotelling Battery or AC (plug-in)
204	Hotelling All Engines and Accessories Off

2.1.5.4 Off-network Idle Activity

Off-network Idle Activity (ONI) is the new activity type introduced in MOVES3. ONI is defined in MOVES3 as when a vehicle engine is running in idle mode and the vehicle is somewhere other than on the road, such as in a parking lot or driveway. MOVES2014 primarily assigned vehicle running emissions to four road types, with idle emissions linked to driving schedules and varying by average speed and road type. However, it was found that these driving schedules significantly underestimated idle times during trips. For example, the total idle fraction in MOVES2014 was about 14% for certain source types, while data from Verizon Telematics indicated it should be 18–31%[12]. This discrepancy was partly due to excluding activities like parking and queuing in drive cycle development. MOVES3 addresses this by adding a new calculation for idle emissions occurring off the road network for all source types, providing a more accurate reflection of real-world idling and idling emissions that occur off the road network (i.e., on roadTypeID=1), especially in light of increased congestion and idling in recent years. The details of the off-network idle activity calculation methodology will be described in the AERR 2020 section. MOVES allows users to update the *TotalIdleFraction* table with more specific data.

2.1.6 Fleet Characteristics

2.1.6.1 Age distribution

The age distribution of vehicle fleets can vary significantly from area to area. Fleets with a higher percentage of older vehicles will have higher emissions. Surveys of registration data indicate considerable local variability in vehicle age distributions, which is not reflected in the default age distributions in MOVES. MOVES uses the same national default age distribution in the SourceTypeAgeDistribution table for each vehicle type in

[12] Population and Activity of Onroad Vehicles in MOVES3. https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P1011TF8.pdf.

each year for every county. Therefore, the EPA recommends and encourages states to develop local age distributions to specific calendar years for SIP and conformity purposes. EPA recommends compiling data according to MOVES vehicle classifications and model years in the same format of 0 to 30 model years (in a total of 31 years). In Texas, local age distributions can be estimated from local vehicle registration data.

2.1.6.2AVFT

The *AVFT* (fuel type and vehicle technology) table allows users to modify the fraction of vehicles capable of using different fuels and technologies in each model year. Specifically, the *AVFT* table allows users to define the engine fraction split between diesel, gasoline, E85, Compressed Natural Gas (CNG), and electricity for each vehicle type and model year. The default VMT split between diesel, gasoline, CNG, and E85 should be used in many cases. For all 13 source types, the use of local information may be more important because of a higher likelihood that national defaults are not consistent with the local fleet.

2.1.6.3I/M Programs

For I/M programs, EPA recommends that users review the inputs in the default *IMCoverage* table and make any necessary changes to match the actual local program.

2.1.7 Fuel Properties

MOVES has four tables related to fuel —*fuel supply, fuel formulation*, *fuelusagefraction*, and *AVFT*—that interact to define the fuels used in the modeled area. MOVES calculates fuel adjustments based on the attributes defined in the *fuelformulation* table. MOVES then uses the *marketshare* field from the *fuelsupply* table to appropriately weigh the fuel adjustment factors. Finally, the emission rates are applied to the appropriate activity defined through the *fuelusagefraction* and *AVFT* tables.

MOVES has default diesel, gasoline, E85, and CNG fuel formulation and supply information for every county-year-month combination that can be selected. These default tables are based on volumetric fuel data for thousands of batches of fuel in each of the fuel regions. For MOVES3, the EPA developed new fuel properties by region based on averages of survey data and data provided to the EPA at the refinery gate as part of EPA fuel compliance programs. These new data provide consistent and maintainable fuel defaults that account for fuel production and distribution networks, natural borders, and regional/state/local variations in fuel policy.

Users should review the default fuel formulation and supply and then make changes only where specific local volumetric fuel property information is available or local fuel requirements have changed. Where local requirements have not changed, EPA strongly recommends using the default fuel properties for a region unless a complete local fuel property study exists.

2.1.8 Meteorology

Local temperature and humidity data are required inputs for SIP and regional conformity analyses with MOVES. Ambient air temperature is a key factor in estimating emission rates for on-road vehicles due to its substantial effects on most pollutant processes. The relative humidity is also important for estimating NOx emissions from motor vehicles. MOVES requires a temperature (in degrees Fahrenheit) and relative humidity (0 to 100 percent) input for each analyzed hour in the *zoneMonthHour* table. The local barometric pressure data should be provided in the *county* table. Therefore, MOVES requires a 24-hour temperature and humidity profile to model a full day of emissions on an hourly basis. Temperature assumptions used for regional conformity analyses must be consistent with those used to establish the emission estimates in the SIP. The MOVES default temperature, humidity, and barometric data are based on average temperatures for each county from the National Climatic Data Center for the period from 2001 to 2011. EPA does not recommend using these default values for SIP EIs.

2.1.9 Summary for CBD Input Guidance

The MOVES3 model requires various data inputs to estimate on-road vehicle emissions. For County Scale analysis, it recommends using local data over defaults for more accurate results. Inputs include on-network and off-network activities, fleet characteristics, fuel properties, and meteorology. Key tables detail VMT by vehicle type and year, speed distributions, vehicle populations, start activity, off-network idling activity, and hotelling activity. MOVES3 also allows for updated local data to enhance model precision. MOVES defaults can be used if the local data is unavailable for off-network activity. This guidance ensures that the model reflects local conditions, which is crucial for State Implementation Plans and regional analyses.

2.2 2.2 National Emission Inventory (NEI 2020)

The EPA estimated the on-road mobile source portion of the 2020 NEI based on MOVES3 modeling using some activity data provided by state, local, and tribal entities. Emissions were calculated at the county level. EPA-developed default activity-based data from the Federal Highway Administration (FHWA) and other sources in cases where state, local, or tribal data was unavailable. EPA also developed default data for all other inputs required by MOVES for use in cases where local data were deemed of insufficient quality or unavailable.

2.2.1 Overview of Emission Estimation

The EPA used *Emission Rates* mode to run MOVES3 with NEI CDBs with activity data in the MOVES-SMOKE framework to compute the EIs for 2020 NEI. MOVES was run for a subset of the representative county to mitigate the computation time and cost. The representative county approach is also supported by the concept that the most important emissions-determining differences among counties can be accounted for by assigning counties to groups with similar properties, such as fleet age, a shared inspection and maintenance (I/M) program, and shared fuel controls. The MOVES input data for county grouping included state, altitude, fuel region, presence of an I/M program, and average light-duty vehicle age. After grouping similar counties, the county with the highest VMT in each group was selected as the representative county.

For emission calculations, EPA first developed lookup tables of emission rates for representative counties. To generate the MOVES emission rates for counties in each state across the U.S., the EPA used an automated process to run MOVES to produce emission factors by temperature and speed for a set of representative counties to which every other county was mapped in Sparse Matrix Operator Kernel Emissions (SMOKE).

To prepare the NEI emissions, the EPA first generated emissions at an hourly resolution using more detailed source classification codes (SCCs) than are found in the NEI (i.e., by MOVES road type and aggregate processes). The MOVES-SMOKE MOVESmrg program performs this function by combining activity data, meteorological data, and emission factors to produce gridded, hourly emissions. SMOKE selected appropriate emissions rates for each county, hourly temperature, SCC, and speed bin and multiplied the emission rate by activity (VMT, vehicle population, or hoteling hours) to produce emissions. These calculations were done for every county, grid cell, and hour in the continental U.S. and aggregated by county and SCC for use in the 2020 NEI.

EPA ran MOVESmrg for each set of emission factor tables RatePerDistance (RPD), RatePerVehicle (RPV), RatePerHour (RPH), RatePerHourONI (RPHO), RatePerProfile (RPP), and RatePerStart (RPS). During the Movesmrg run, the program used the hourly, gridded temperature (for RPD, RPV, RPH, RPS, and RPHO) or daily, gridded temperature profile (for RPP) to select the proper emissions rates and compute emissions. These calculations were done for all counties and SCCs in the SMOKE inputs, covering the continental U.S.

2.2.2 Inputs for NEI 2020

EPA received 1,565 MOVES CDB submittals from state, local, and tribal agencies and applied them based on the priority level for the 2020 NEI[13,14]. IHS-Markit (IHS) vehicle registration data was adapted to compute vehicle populations (VPOP), vehicle age distributions, and fuel type fractions. FHWA provided VMT data by county and road type. EPA also received 2020 vehicle telematics data from StreetLight Data, Inc., which EPA transformed into MOVES- and SMOKE-ready input files describing the distributions of vehicle speeds and fractions of VMT by the hour, day of the week, and month.

EPA purchased data on vehicles in operation across the nation as of July 1, 2020, from IHS, which was further processed to develop CDB tables of vehicle population, age distribution, and fuel fractions. IHS receives registration records from each state's Department of Motor Vehicles (DMV) and decodes vehicle identification numbers to assign each vehicle a MOVES source type code.

2.2.3 Activity Characteristics: On-network

2.2.3.1 VMT by Calendar Year Source Type, and Speed Distribution

For the 2020 NEI, the FHWA provided VMT data by county and road type. Determination of the speed distribution was a challenge for 2020 because of the effects of COVID-19 on travel patterns. Some states responded that they did not have the resources to provide inputs that would characterize the effects of COVID on the 2020 on-road activity inputs. In response to state needs, EPA purchased county-level telematics data from StreetLight to characterize vehicle speed profiles and VMT temporal distributions. StreetLight uses Location-Based Services (LBS) data from cellular phones as a surrogate

[13] 2020 National Emissions Inventory Technical Support Document: Onroad Mobile Sources. U.S. Environmental Protection Agency (USEPA. 2023)

[14] Instructions and Best Practices for Development and Submittal of Onroad Inputs for the 2020 National Emissions Inventory (NEI). (USEPA. 2021). https://www.epa.gov/system/files/documents/2021-07/instructions-and-best-practices-for-submitting-mobile-source-data-20210722.pdf.

for personal vehicles and in-vehicle GPS data for medium and heavy-duty commercial trucks.

For the state of Texas, the total VMT is directly applied by the EPA from the submitted CDBs. TCEQ provided VMT according to HPMS source types. The EPA further distributed the provided VMT to MOVES source types. For the speed distribution, EPA's starting point CDBs have national average placeholder values for this table that will change when 2020-specific data become available. Differences exist in the *averagespeedfraction* field of 2020 AERR and 2020 NEI CDBs. An example is illustrated in Appendix A.

2.2.3.2 VMT Distribution

For the year 2020, the VMT sample from Streetlight represented over nine percent of the continental US total VMT 2020, as estimated by the FHWA. EPA accepted agency submittals for month VMT fractions if the patterns clearly showed 2020 pandemic effects for the expected months. However, EPA used telematics-based data in all counties for the hour and day VMT fractions as well as the speed distributions because of the higher resolution (by month and seven-day types of the week) available from StreetLight.

EPA did not use any state, local, and tribal agency data submittals for *hourvmtfraction* and *dayvmtfraction* because 2020 was a unique year where the activity varied by month. The months of January and February look a lot different than March through December. However, two data tables (*hourvmtfraction* and *avgspeeddistribution*) are the annual average in the CDB submittal framework. Therefore, EPA developed month-specific data for use with MOVES and SMOKE.

EPA reviewed agency data submittals for *monthvmtfraction* and found that many states provided VMT distributions that reflected the actual conditions in 2020. For these agencies, the EPA used the submitted *monthvmtfraction*. Elsewhere, EPA developed statewide averages for personal vehicles separately from commercial trucks based on the monthly StreetLight data sample normalized by the number of devices installed in the trucks. Due to sample size differences by month, the commercial truck data's *monthvmtfraction* had an unrealistic spike in the profile in May. EPA corrected this by dropping that month and interpolating it from April to June instead.

EPA used MOVES3 default distributions of VMT across four road types by county for the *roadtypedistribution* table.

For the state of Texas, EPA defaults were used for *monthvmtfraction* and *dayvmtfraction*. There is a considerable difference in *monthvmtfraction* for source type 62. There are differences in hourvmtfraction submitted by the TCEQ and used by the EPA.

2.2.3.3Quality Assurance Checks

EPA developed two decision tree flowcharts to establish procedures for filling gaps, one for the VMT distributions and a different process for the speed distributions. In general, the EPA preferred to let the local data from each county stand on its own, representing only itself, even in low data areas with resulting "noisy" data profiles that may have missing hours of data. For the VMT distributions with missing hours, EPA set those hours values to zero (0), interpreting the missing coverage as low or no vehicle activity. As telematics data samples continue to grow into the future, the instances of missing coverage are expected to lessen. In contrast, for the speed distributions, EPA did not allow missing hours of data in the modeling profiles due to the potential for data loss in the SMOKE-MOVES system and representative county approach.

After the vehicle population and VMT data were finalized, the population and VMT were compared by county and source type to look for inconsistencies between the population and VMT datasets. Specifically, counties and source types with unreasonably high miles per year per vehicle average (VMT divided by VPOP) were identified and addressed. For counties and source types with a VMT/VPOP ratio above a certain threshold, the vehicle population was increased so that the new VMT/VPOP ratio would equal the maximum allowable ratio.

2.2.4 Activity Characteristics: Off-network

2.2.4.1Source type population

Vehicle population was one of the most commonly provided local data types. States without any CDB submittals received EPA default populations. Some states with submittals were overridden and decided on a case-by-case basis. In areas where submitted vehicle population data were accepted for the NEI, the relative populations of cars vs. light-duty trucks were reapportioned (while retaining the magnitude of the light-duty vehicles from the submittals) using the county-specific percentages from the IHS data.

IHS's registration data reflected higher light-duty vehicle (LDV) populations than corresponding state agency analyses of the same DMV data. EPA found that IHS populations for 2020 were higher than the state data by 10.8 percent on average for LDVs. EPA also found that the discrepancies in the 2020 data between IHS and states are larger for older vehicles. Adjustment factors were applied to the IHS data to create EPA defaults for vehicle population and age distribution. In addition to removing the older and antique plate vehicles from the IHS data, the EPA also removed outlier age distributions that showed excessively "new" fleets, usually for light commercial trucks, in 28 counties. The most extreme example of this was a light commercial truck age distribution where over 85 percent of the commercial light-duty truck population in the county is 0 or 1 year old. This situation where the registration data reflects a young fleet occurs when the headquarters of a leasing or rental company owns a large fraction of the vehicles in the county.

2.2.4.2 Starts, Hotelling, and Idling Activities

In the case of off-network activity, EPA has empty tables in the CDBs for *hotellingagefraction* and *hotellinghourfraction*. EPA has used a different profile for *hotellingmonthadjust* than the submitted data. EPA has an empty CDB table for *startsperdaypervehicle*. These off-network activity input tables use national average EPA estimates in 2020 NEI.

2.2.5 Fleet Characteristics

2.2.5.1 Age distribution

IHS data for 2020 were used to derive updated age distributions adjusted to remove older vehicles (MOVES *sourcetypeagedistribution* table) and to split by source type, fuel type, and model year (MOVES *AVFT* table) in the CDBs. These data were computed at the county level for "all CDBs." They were a weighted average over county groups for the set of representative CDBs used in the MOVES runs for the NEI. In both cases, EPA preferred to use local data where they were acceptable. Local data were used preferentially and supplemented with EPA-developed information where needed. The source registration data in the EPA-developed data does not reliably distinguish between short-haul and long-haul activity. So source types 52 and 53 (single unit trucks) have the same age distributions as those of source types 61 and 62 (combination trucks). In addition, all age distributions for long-haul trucks (source types 53 and 62) are a national average because these vehicles are expected to travel long distances from the county where they are registered.

2.2.5.2 AVFT

Because the *AVFT* table is optional in a MOVES CDB, it was not always populated in a submitted database. In cases where data was not provided, the EPA used a default of MOVES national distributions of fuel types and/or E85 availability to fill in the missing data using the SampleVehiclePopulation tables of the model default database.

2.2.5.3 I/M Programs

The 2020 NEI for Texas used 2020 AERR values with updated compliance rates for counties with I/M programs.

2.2.6 Fuel Properties

Like representative county, fuel month was used for aggregation concerning fuel type and month. A "fuel month" indicates when a particular set of fuel properties should be used in a MOVES simulation. Similar to the representative county, the fuel month reduces the computational time of MOVES by using a single month to represent a set of months during which a specific fuel has been used in a representative county. Because there are winter and summer fuels, the EPA used January to represent October through April and July to represent May through September.

For 2020, the nationwide fuel supply assumed a 100% market share of E10 ethanol blends in gasoline, which was also true for Texas. All diesel was assumed to be six ppm sulfur, and on-road diesel was 100% market share B5 biodiesel blends nationwide. EPA's starting point CDBs already have values for this table from MOVES3 that will likely remain the final 2020 values for the *fuelusagefraction* table.

2.2.7 Meteorology

Ambient temperature can have a large impact on emissions. Low temperatures are associated with high start emissions for many pollutants. High temperatures and high relative humidity are associated with greater running emissions due to increased heat index and higher engine load for air conditioning. High temperatures also are associated with higher evaporative emissions. EPA applied the SMOKE program Met4moves to the gridded, hourly meteorological data (output from Meteorology-Chemistry Interface Processor [MCIP]) to generate a list of the maximum temperature ranges, average relative humidity, and temperature profiles that are needed for MOVES to create the emission-factor lookup tables. "Temperature profiles" are arrays of 24-hour temperatures that describe how temperatures change over a day. They are used by

MOVES to estimate vapor venting emissions.Met4moves computes the range of temperatures needed by each representative county for each fuel month (i.e., 5-month summer season [May to September] or 7-month winter season [October to April]). When the emission factors are applied by SMOKE, the appropriate temperature bin and fuel month are used to compute the emissions. EPA used a five °F temperature bin size for RPD, RPV, RPH, RPHO, and RPS.

For the state of Texas, there are differences in temperature and relative humidity in the *zonemonthhour* table in the 2020 NEI compared to TCEQ CDBs with local data since the EPA uses Weather Research and Forecasting (WRF) meteorology for the 2020 NEI.

2.2.8 Summary of EPA's NEI Development

EPA calculated the on-road emissions 2020 for all states using the most recent version of MOVES, MOVES3.0.4, with default database *movesdb20220802*. EPA used programs within the SMOKE modeling system that use data output from MOVES to generate the emission inventories in all 50 states for each hour of the year. These emissions were summed over all hours and across road types to develop the county-level emissions estimates for the NEI.

The data selection hierarchy for 2020 favored local input data over EPA-developed information, except for the three MOVES tables *hourvmtfraction*, *dayvmtfraction*, and *avgspeeddistribution* where county-level, telematics-based EPA Defaults were adopted for the NEI universally due to unique activity patterns by month during 2020.

2.3 2.3 TCEQ's 2020 AERR Submittal

The 2020 AERR *annual* EIs use MOVES in *inventory mode to directly get the emission estimates by SCCs for different counties*. However, the input CDBs for the *annual* inventories used an earlier set of summer weekday (*swkd*) Emission Rates MOVES results and external activity estimates. Thus, the basic methodology of 2020 NEI and 2020 AERR is similar. The *Emission Rates* mode-based swkd output was also used for 2020 AERR development; thus, it is relevant for some discussion in the results section.

The methods used to calculate the *swkd* EIs are an extension of historically consistent traffic activity and emission rate methods developed by TTI (Venugopal 2022). This section details the data sources, methods, and annual and swkd EIs. The description captures the essence of the methodology, but implementation was done differently.

The emissions inventory calculations described in this section were based on an hourly, link-level analysis that uses the outputs of the regional TDM or HPMS, as well as other local data sources (e.g., seasonal, day type, and hourly travel factors; vehicle population data; and environmental inputs) consistent with the region, and MOVES default inputs.

2.3.1 Overview of Emission Estimation

The 2020 AERR swkd and annual EI were performed using five basic steps as follows:

- **Step 1** – Estimate Emission Rates: MOVES3 was used to estimate regional *swkd* emission rates (or factors) relevant to the analysis area. The rates were calculated based on local inputs to MOVES, such as temperature and humidity, fuel formulation, etc.
- **Step 2** – Estimate Traffic Activity: The local HPMS and TDM data were processed to derive 24 hourly VMT and speed estimates for all virtual HPMS links and TDM links (as well as for added TDM intrazonal links). Further processing was used to convert VMT-based HPMS and seasonal and daily adjustment factors. Local automatic traffic recorder (ATR) traffic count data was used to process the HPMS and TDM data. After the on-network activity was estimated, off-network activity was calculated using outputs from the processed HPMS and TDM data, vehicle population data, and MOVES default inputs. The traffic activity was processed to replicate the operating conditions for each EI.
- **Step 3** – Develop *swkd* Emissions: The *swkd* emission rates calculated in Step 1 were multiplied by the on- and off-network activity calculated in Step 2. This yielded emission estimates in units of mass calculated at a spatial scale of each link (on-network) or county (off-network) for each hour of the day.
- **Step 4** – Develop Annual Emissions: The *swkd* on- and off-network activity were used in the development of the *annual* activity, which in turn was used to develop the county-level CDBs used in the annual inventory runs using MOVES.
- **Step 5** – Post-Process EI Outputs: Outputs for each pollutant were post-processed into various formats and electronic deliverables for reporting purposes and downstream air quality planning.

swkd inventories were estimated using the detailed hourly link (roadway segment)-based method. CDBs for MOVES *annual inventory mode* runs were prepared using local input data (from the *swkd* EI activity data and various conversion factors) and some default input data.

Annual EIs were produced from MOVES *inventory* mode runs using the local, *annual inventory* mode CDBs. The *swkd emission rates* and *annual inventory*-mode runs were performed using the EPA's latest version of MOVES.

The EIs were calculated using a detailed MOVES rates-per-activity estimation method based on the areas described in Table 8. The counties were divided based on the availability of travel demand model data. swkd rates per activity approach calculate on-network emissions for each link defined by the regional TDM or HPMS outputs and formats results as needed for subsequent uses. It also calculated off-network activity and emissions.

Table 7. County Groups for CDB Development (Adapted from Venugopal 2022)

Area[1]	Counties	Activity Basis
1. Austin	Bastrop, Burnet, Caldwell, Hays, Travis, Williamson	TDM
2. Beaumont-Port Arthur	Jefferson, Hardin, and Orange	TDM
3. Dallas-Fort Worth	Collin, Denton, Dallas, Ellis, Hood, Hunt, Johnson, Kaufman, Parker, Rockwall, Tarrant, Wise	TDM
4. El Paso[2]	El Paso	TDM
5. Houston-Galveston-Brazoria	Brazoria, Chambers, Fort Bend, Galveston, Harris, Liberty, Montgomery, Waller	TDM
6. San Antonio	Bexar, Comal, Guadalupe, Kendall, Wilson	TDM
7. Tyler-Longview-Marshall	Gregg, Smith, Harrison, Rusk, Upshur	TDM HPMS
8. Remainder of Texas	214 Counties	HPMS
Totals by Activity Basis	**37**	**TDM**
	217	**HPMS**

[1] The 40 counties listed as (1) through (7) were modeled using county-level emission rates. In contrast, the remaining 214 counties (8) were modeled using the statewide inventory methodology, which produces emission rate estimates by county groups.
[2] El Paso is the only county for which a winter weekday inventory was produced.

The TDM and HPMS data were post-processed to estimate hourly, directional, link-level VMT, and operational speeds for the emission calculations. The hourly off-network activity factors were estimated for the off-network emission calculations using estimates of vehicle operating hours (also known as vehicle hours traveled [VHT]), vehicle type populations, combination long-haul truck hotelling, and other data. These off-network activity factors are ONI hours, source hours parked (SHP), starts, and source hours extended idling (SHEI) and APU hours—where SHEI and APU are components of hotelling hours for combination long-haul trucks. Postprocessing was performed using MOVES input, output, and default data to produce the off-network evaporative emission rates for mass/SHP (currently not directly provided by MOVES). These post-processed emission rates were compiled with the other rates produced directly by MOVES emission rate mode runs, yielding final emission rate lookup tables with all rates in terms of mass per vehicle activity unit (i.e., mass/mile, mass/SHP, mass/start, mass/ONI hour, mass/SHEI, mass/APU hour). Table 8 shows the mapping between the activity and emission rate categories.

Table 8. Emission Processes and Emission Rates

Emissions Process (Rates Category)	Activity	Emission Rates
Running Exhaust	VMT	mass/mile (mass/mi)
Crankcase Running Exhaust	VMT	mass/mi
Brake Wear	VMT	mass/mi
Tire Wear	VMT	mass/mi
Start Exhaust	Starts	mass/start
Crankcase Start Exhaust	Starts	mass/start
Extended Idle Exhaust	SHEI	mass/hour
Crankcase Extended Idle Exhaust	SHEI	mass/hour
Auxiliary Power Exhaust	APU Hours	mass/hour
Running exhaust (1) – Road Type 1 off-network	ONI Hours	mass/hour
Evaporative Permeation Evaporative Fuel Vapor Venting Evaporative Fuel Leaks	VMT, SHP	mass/mi, mass/hour

2.3.2 Inputs for AERR 2020

This section provides an overview of the input development for the 2020 AERR study.

2.3.2.1 Activity Characteristics: On-network

2.3.2.1.1 VMT by Calendar Year and Source Type and VMT Distribution

TDMs are the main sources to get total VMT. However, for the 2020 AERR study, TDMs were unavailable for all the counties. In the counties where TDMs were not available, TTI used HPMS data. For these counties, 21 road links based on 3 area types (rural to a different degree of urban) and 7 FHWA functional classifications are defined as virtual links in the model. VMT is distributed for each virtual link using adjustment factors. See Table 7 for TDM and non-TDM counties. Note that for 2020, HPMS VMT was available for all counties. Based on standard practice, the VMT of the TDM counties was scaled up or down to make the total county VMT match the HPMS data.

The 2020 AERR study uses ATR station-based seasonal, day type, and hourly factors to convert TDM or HPMS VMT to appropriate season, month, and day type (e.g., weekday) and distribute the link VMT estimates to each hour of the day.

By applying the ATR-based factors, TTI could convert the link VMT to a one-hour resolution for any time of year and day type. TTI further splits the hourly VMT based on a "VMT-Mix"[15] to get the VMT by MOVES source use types and fuel types.

Hourly VMT estimates by vehicle type were combined with the appropriate emission factors in the link-emissions calculations. The datasets used to calculate the VMT mix include local vehicle classification count and ATR data, MOVES defaults, and local registration data. The hourly, link-based emissions process requires VMT estimates by hour and direction for each link in the TDM or virtual link in HPMS.

The above VMT by source and fuel types are then reaggregated to the required level. For 2020, AERR swkd estimates used the VMT for summer weekdays directly and used the emission rate output from MOVES. In contrast, the annual estimates aggregate the VMT to annual resolution and by HPMS vehicle types. Figure 2 shows the above workflow of transforming the TDM or HPMS VMT to hourly VMT by source type and fuel type at the temporal resolution required. The reader is referred to the 2020 AERR report section 2.3.1for the detailed equations.

[15] The VMT mix procedure correlates the TxDOT vehicle classification count (VCC) data for FHWA (HPMS) vehicle classes to MOVES on-road SUT groupings, then applies a series of factors (based on MOVES defaults and TxDMV registration data) and assumptions to expand the HPMS category data into the 13 individual MOVES SUTs. Application of gasoline and diesel fuel fractions results in 22 MOVES vehicle types. To produce a robust set of VMT mixes, TTI aggregated multiple years of the hourly, weekday, VCC data, in four time periods (AM peak, midday, PM peak, overnight), into four functional classification groups (correlated to MOVES road types) and combined adjacent counties into regional estimates (based on TxDOT Districts).

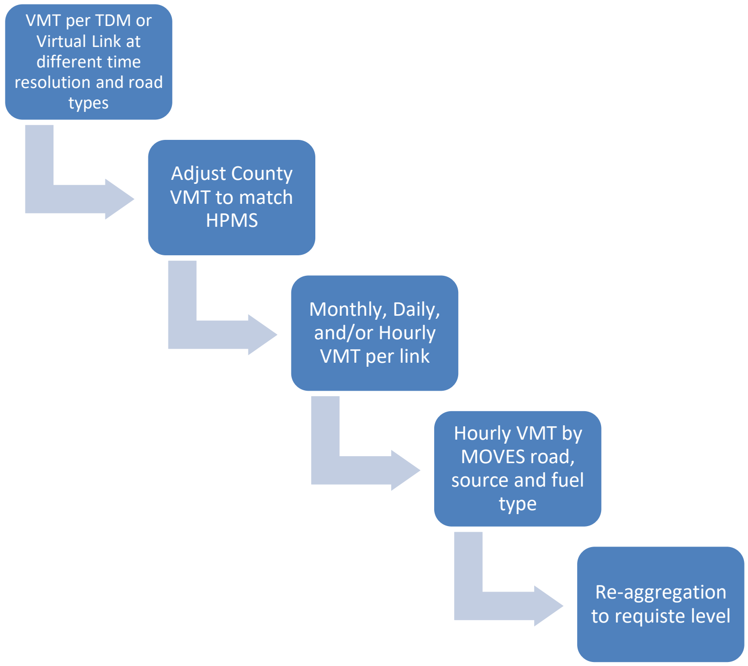


Figure 2. VMT Breakdown Flowchart

In addition to getting the total VMT, the above procedure provides the required data to populate the *hourvmtfraction*, *dayvmtfraction*, *monthvmtfraction*, and *roadtypedistribution*.

2.3.2.1.2 Average Speed

On-network emission factors are based on each link's congested (or operational) speed. Three different speed models were used to estimate the speed for each TDM link[16] or virtual link in the analysis: The TTI speed model, the Houston speed model, and the virtual link speed model. The TTI speed model was used for all metropolitan areas that provided a TDM, except for the Houston/Galveston TDM area, which uses the Houston speed model designed specifically for the Houston/Galveston TDM. All HPMS-based virtual link areas use the virtual link speed model. Congested speed for links was calculated from free flow speed and delays. The reader is referred to the 2020 AERR

[16] Travel demand models explicitly model vehicle speed (through volume-capacity relationships) as part of the process of estimating the route choice of travelers through a network. However, most travel demand models only provide speed for aggregate, non-hourly time periods (e.g., a full day, peak, or off-peak). The standard on-road emissions inventory process is based on hourly VMT and speeds. This requires disaggregating travel model VMT to each hour of the day, and then re-estimating operational speed for each hour of the day using the hourly VMT and speed-volume-capacity relationships like those used in the original travel modeling process.

report for the detailed equations.

TTI used a two-step procedure to estimate congested link speeds. The first step is the *v/c* ratio calculation. The second step is applying the congested speed model to estimate the congested speed[17]. *v/c* ratios are generated for each combination of a period (hour), roadway functional classification, area type, and direction using the hourly lane capacities and VMT. The congested speed model calculates the delay on the link and then applies this delay to the link free-flow speed to calculate the link operational congested speed estimate.

These speed models were applied to all functional classes, excluding the TDM centroid connector and intrazonal functional classes. For these functional classes, capacity data were not used. Average trip length and time within the zone were used to estimate the speed.

2.3.2.1.3 Speed Distribution

The speeds were converted to VHT, which can be transformed to other resolutions, along with VMT to re-compute speeds. The VHTs and corresponding VMTs were grouped into the source type, road type, analysis hour, day, and month categories. Thus, given VMT and VHTs for an hour for a day type (e.g., weekday) for a MOVES source type and road type, the user can allocate these quantities into 16-speed bins. The advantage of this disaggregation is that it allows the user to obtain *avgspeeddistribution* for higher temporal resolution, such as an entire day or a year. The VHTs at desired aggregation by MOVES road type were used to get the average speed distribution. by MOVES's average speed distribution table, *avgspeeddistribution* by day type, hour, road type, source type, and 16 average speed bins.

2.3.2.2Activity Characteristics: Off-network

The 2020 AERR study computed off-network activity: ONI hours, SHP, starts, and long-haul combination truck hotelling hours. These quantities are estimated for each hour of the day at a spatial scale of a county and for each MOVES source use type (SUT) and fuel type combination.

[17] The virtual link speed model was applied to the 214 Texas counties that are not included in a TDM. There are three critical parameters for estimating operational speeds on virtual links: hourly lane capacity, free-flow speed, and hourly volume by direction. The virtual link speed model was applied to estimate a link's directional, time-of-day congested speed. The capacities and free-flow speeds used in the virtual link speed model procedure are based on the Highway Capacity Manual (HCM).

Vehicle population data were used to estimate SHP and vehicle starts off-network activity. The disaggregation of TxDMV registration data into MOVES SUT population data is described in the section on Source type population. Estimations for each of the above off-network activities are described below.

- **ONI** activity for each hour was determined from the VHT (alias source hour operating [SHO] in MOVES) on each link, total source hour idling (SHI), and a total idle fraction (TIF).

$$SHI = \left(ONI + \sum_{i=2}^{5} VHT\right) \times TIF$$

$$ONI = \frac{TIF \times \sum_{i=2}^{5} VHT - SHI}{1 - TIF}$$

where,

TIF is a constant obtained from the MOVES *totalidlefraction* table. It is the ratio of total idling by total vehicle operating hours on-road and off-network,

SHI is the on-road idling hours. It is computed from RIF and VHT_{cat}. *RIF* is the road idle fraction from the *roadidlefraction* table. It is the fraction of SHI and VHT. *roadidlefraction* provides *RIF* by month, day, road type, source type, and average speed bin. VHT_{cat} is the *VHT* corresponding to the *RIF* month, day, road type, source type, and average speed bin. The user can compute the SHI by month, day, road type, source type, and average speed bin based on the VHT estimates (described in the previous section) and aggregate it to get the on-road idling hours and

ONI is the off-network idling.

- **County-level vehicle type SHP** was calculated for each hour of the day and each SUT as the difference between the local vehicle population (total available vehicle hours) minus VHT.
- **Vehicle starts** for *swkd* were directly obtained from the emission rate starts output from the *startspervehicle* table. The input CDB used default starts tables. For the annual case, TTI uses the output of *startspervehicle* to get the hourly start by county, hour, source type, and fuel type, and the MOVES default *startsperdaypervehicle* to

split the starts into hourly starts. The starts per vehicle per hour were also adjusted based on the local fleet's age distribution and fuel fractions. TTI used local age distributions and fuel fractions inputs to MOVES combined with MOVES default parameters (*startsageadjustment* [three-month seasonal average]) to produce age-adjusted hourly starts per vehicle per hour. The age-adjusted hourly starts per vehicle were further modified to account for the month using the MOVES default parameter: *startsmonthadjust* (three-month seasonal average).

- **Hotelling hours** were calculated for heavy-duty, long-haul trucks only (i.e., SUT 62) in several steps. The first total 2017 winter daily hotelling hours were obtained from a 2017 TCEQ extended idling study. VMT-based scaling factors were then used to convert these 2017 hotelling hours to those relevant to the analysis scenario (defined by analysis year, season, and day type), which were then allocated to each hour of the day based on the inverse of CLhT hourly VHT; a surrogate for hotelling activity. Estimations were then made of the proportions of hotelling hours that occur in each of the four hotelling categories: idling using the main engine (SHEI), idling using a diesel APU, idling using an electric APU, or idling with no engine or auxiliary power.

Similar to the on-road activity, TTI methodologies for off-network aim to get the off-network activity at the hourly level by source type and fuel type for different seasons. This allows TTI to either directly use this activity or aggregate them as needed.

2.3.2.3Fleet Characteristics

2.3.2.3.1 Source type population

At the end of the year 2018, TxDMV vehicle registration data was provided in the form of total vehicles registered by county, aggregated by vehicle categories. These TxDMV vehicle categories were disaggregated to MOVES SUT and fuel type aggregations.

The following steps were used to disaggregate the TxDMV vehicle registration data to vehicle population data by vehicle type.

- Step 1 – VMT-Mix data was used to calculate the proportional representation of each MOVES vehicle type within each TxDMV aggregation class.
- Step 2 – The proportional fractions calculated in Step 1 were multiplied by the total number of vehicles reported in each TxDMV vehicle registration category to obtain the estimated number of vehicles (populations) for each modeled MOVES vehicle type.

- Step 3 – The long-haul truck vehicle type populations were estimated as an extension of their estimated short-haul vehicle type population counterparts by multiplying a long-haul-to-short-haul ratio derived from the weekday vehicle type VMT mix by the associated short-haul truck vehicle type populations from Step 2.

The VMT mix data used in these calculations was the TxDOT district-level, 24-hour weekday VMT mix.

2.3.2.3.2 Age distribution and AVFT

Vehicle registration data from TxDMV and MOVES defaults were used to develop age distributions for various source types. Regarding motorcycles and passenger cars, the registration data was directly available. However, truck registration data were divided into MOVES source types according to the truck's weight available in the registration data. MOVES defaults were used for age-type distribution for some source types[18].

The local TxDMV registration data provides fuel type fractions (proportion of gasoline or diesel-powered vehicles) for heavy-duty vehicles but not for light-duty vehicles. MOVES default fuel fractions were therefore applied to estimate light-duty fuel fractions. Only gasoline and diesel vehicles were included in the CDBs.

2.3.2.3.3 I/M Programs

2020 AERR used an updated compliance rate provided by TCEQ for counties with I/M programs.

2.3.2.4Fuel Properties

TTI used various data sources to produce the best available Texas summer and winter fuel formulation inputs to MOVES. Four MOVES fuel input tables must be consistent between the fuel types in the scope of the inventory analysis. These are:

- *AVFT* (source type population fuel type distributions by model year).
- *fuelformulation* (fuel properties for each fuel subtype supplied in the study area).
- *fuelsupply* (market shares of each fuel sub-type formulation).
- *fuelusagefraction* (flex fuel vehicle fuel type usage).

[18] For more details on mapping registration data categories and MOVES source use types, the reader is referred to Table 23 of the 2020 AERR report.

The fuel types in the scope of the inventory analysis were gasoline and diesel, with alternative fuels assumed to have an insignificant impact. With solely gasoline and diesel set by the *AVFT* and *fuelusagefraction* tables, the *fuel formulation and fuel supply* table's gasoline and diesel fuel properties and market shares were then specified. The alternative fuels available in MOVES3 were treated as negligible and excluded from the analysis (using the MOVES *AVFT*, *fuelusagefraction* tables, and *fuelfraction* inputs). Additionally, MOVES defaults were used as needed. This was the case for winter conventional gasoline formulations and for winter RFG RVP, for which local data were unavailable.

The best available local fuel survey data by season for the study year were used, supplemented as needed by MOVES defaults and other data (e.g., DOE annual fuel sales statistics). Six MOVES fuel regions for Texas performed the fuel formulation development procedures. In general, the sample data were aggregated and averaged by fuel grade within each MOVES fuel region (e.g., consistent with Texas fuel regulation jurisdictions and distribution networks), then weighted into gasoline composite averages using relative sales volumes by grade (results of this procedure were available directly from the TCEQ 2020 survey summary for the summer season). For the MOVES RFG region, TTI developed separate RFG formulation estimates for the DFW and HGB RFG counties for the summer and winter seasons. Summer and winter fuel formulations were applied in the seasonal weekday emission rates via month ID, where MOVES month IDs 1 and 7 (January and July) were used to represent the winter and summer seasons.

The fuel inputs to MOVES were supplied in the CDB *fuelsupply* and *fuelformulation* tables. The local fuel supply for each county, year, and month (or season) consisted of one gasoline and one diesel formulation (except for the other MOVES default alternative fuels required to run MOVES). Therefore, each gasoline and diesel formulation market share in the fuel supply was 1.0.

2.3.2.5Meteorology

Texas statewide AERR inventory analyses use local meteorological input data prepared by 25 TxDOT districts. In contrast, the individual county EI analyses use county-level meteorological inputs. District and county-level meteorological inputs were prepared for the four seasonal periods of spring (March through May), summer (June, July, August), fall (September, October, November), and winter (December, January, February) for all districts and individual counties. The “county” table contains barometric pressure, and the *zonemonthhour* table houses temperature and relative humidity data. TCEQ produced the hourly temperature, hourly relative humidity, and 24-hour barometric pressure averages by season and year, using the latest available 2019 calendar year hourly data from numerous weather stations within each district and county. TTI used the seasonal averages for temperature, relative humidity, and barometric pressure for the summer weekday analyses. Two sets of meteorological data were used: county-level data and one based on district-level data. TTI assigned the district-level meteorological inputs to their corresponding individual counties for all 254 counties for use in building the county group CDBs.

2.3.3 Summary of AERR Development

Based on this section and section three, it can be seen that the basic methodology of 2020 NEI and 2020 AERR are similar. Similar to how 2020 NEI computed on-road and off-network activity and rates and used them to compute EIs and develop *Inventory* mode CDBs, 2020 AERR conducted *Emission Rates* mode runs, obtained *swkd* activity, and used the two to develop annual CDBs. Following is a brief overview:

- *swkd* inventories were estimated using the detailed hourly link (roadway segment)-based method. It uses MOVES *emission rates* mode CDBs and TTI EI utilities (Python scripts) to process various on-road activity inputs, off-network inputs, and MOVES defaults to produce granular hourly link-based emission estimates.
- *Annual* EIs were produced from MOVES *inventory* mode runs using the local, *annual* inventory mode CDBs. To build the *annual* CDBs consistently, the *swkd* activity was converted to annual activity based on the TTI EI utilities to process the data into the MOVES3 *inventory* mode inputs for the *annual* runs. TTI EI utilities access the data sources, perform the needed data processing into MOVES input form, and organize the resulting MOVES input files in folders by county, year, period, and day type for populating the CDBs.

2.4 Methodologies Adapted by Other States

This section briefly reviews data sources and methodologies adapted by other states to develop input CDBs for MOVES. Specifically, the following four methodologies are studied:

- Atlanta (The Atlanta Region's Plan 2015)
- Chicago area (Chicago Metropolitan Agency for Planning 2018)

- Pennsylvania (Pennsylvania Department of Transportation 2018)
- Kentucky (Clarksville Urbanized Area MPO 2019)

The methodologies used by the above four states were consistent for important CDBs. All four states used TDM to estimate VMT. Atlanta and Kentucky used default data, and Chicago and Pennsylvania used local data to determine VMT distribution fractions. All four states used TDM data provided by state DOTs to determine road type distributions. Similarly, vehicle registration and TDM data were used for source type and vehicle age distribution. Kentucky used mixed TDM and local data to determine speed distribution, while the other three states used TDM data. Lastly, for fuel, Atlanta found that the default does not match the actual fuel distributions and observed. Adjustments were made to the fuel type distributions. For Chicago, the data was provided by the Illinois EPA. Pennsylvania adjusted fuel inputs according to the state's clean energy programs. Kentucky used MOVES fuel input defaults.

2.5 2020 NEI and TCEQ's 2020 AERR Comparison

This section compares emission and CDB inputs developed by EPA and TTI. Figure 3 gives total emission differences for important pollutants estimated by TTI and EPA for Texas. Emissions estimated by the EPA are higher than TTI estimates for all the pollutants. The total emissions estimated by the EPA are higher than those developed by TTI. Total CO_2 emissions are around 15% higher, and total NOx emissions are around 23% higher.

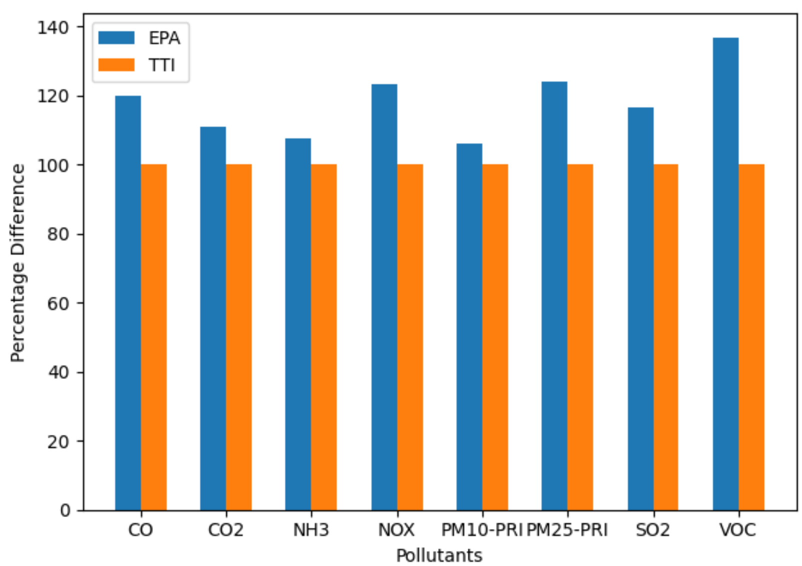


Figure 3. Emissions Comparison for the State of Texas

Figure 4 shows the percentage difference between EPA and TTI NOx emissions for ozone non-attainment zones in Texas. On average, the EPA estimates for these areas are 19% greater than TTI estimates. The percentage difference for particulate matter under 10 microns (PM_{10}) non-attainment zones is 1.54%.

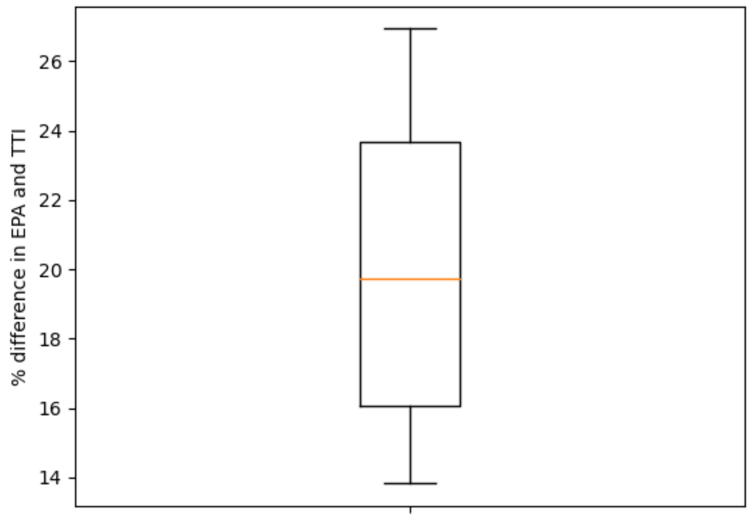


Figure 4. NOx Emissions: 2020 NEI vs. 2020 AERR for 28 Texas Non-Attainment Ozone Counties

The above differences are expected given that the source of the input data sources and methodology are quite different between the two studies. The following section provides a methodology comparison between the two studies. The subsequent section details the input differences.

2.5.1 Methodology

For the state of Texas, some important similarities and differences can be pointed out between CDB submitted by TTI and those used by EPA for emission estimations. For the CDBs related to traffic activity, EPA has directly adapted the total VMT from the submitted data. EPA has used the data submitted by I/M. However, the EPA has used different VMT adjustment factors than the submitted data. EPA utilized IHS registration and telematics data to develop CDBs related to source type population and age distribution. EPA has used more fuel formulations than TTI. However, some off-network activity-related CDBs are empty in the EPA data, e.g., hotelling and vehicle starts. In summary, there are considerable differences in the CDBs submitted by TCEQ and the CDBs used by EPA, which has led to significant differences in total emissions. Table 9 lists the CDB inputs and provides comments on data sources and input development methodology for the 2020 NEI and 2020 AERR study.

Table 9. 2020 AERR vs. 2020 NEI Data Source Comparison

CDB Table Type	CDB Table	2020 AERR Data Source	2020 NEI Data Source
On-road Activity Characteristics	hpmsvtypeyear	Based on TDM and HPMS VMT and local MVC data-based vehicle classification factors and ATR data-based temporal conversion factors.	N/A
	sourcetypevmtyear	N/A	It is based on 2020 AERR HPMS VMT and *sourcetypeyear* population proportions within an HPMS category.
	avgspeeddistribution	Weekday: Based on speed estimation equations using estimated volume from TDM or HPMS VMT and miles and capacity. Weekday: MOVES *avgspeeddistribution* table.	StreetLight telematics data
	dayvmtfraction	ATR Data Average.	StreetLight telematics data
	hourvmtfraction	ATR Data Average.	StreetLight telematics data
	monthvmtfraction	ATR Data Average.	StreetLight telematics data statewide averages to source types except for source type 62, which instead used the ratio of the number of days in each month by 365.
	roadtypedistribution	TDM or HPMS Data mapped to MOVES road types.	MOVES3 default distributions of VMT across four road types by county.
Fleet Characteristics	sourcetypeyear	TxDMV Registration Data	2020 IHS registration data
	avft	TxDMV Registration Data and MOVES default *samplevehiclepopulation* table.	2020 IHS registration data
	imcoverage	MOVES default *IMCoverage* table and updated compliance rate for counties with I/M programs.	2020 AERR data for Texas.
	sourcetypeagedistribution	TxDMV Registration Data and MOVES default *sourcetypeagedistribution* table.	2020 IHS registration data
Hotelling	hotellinghoursperday	2017 Hotelling hours scaled to 2020 based on 2017 and 2020 CLhT VMT.	National hotelling rate per VMT and CLhT restricted road VMT to estimate hotelling hours.
	hotellingactivitydistribution	MOVES default.	MOVES default.
	hotellingagefraction	Empty Table.	Empty Table.

CDB Table Type	CDB Table	2020 AERR Data Source	2020 NEI Data Source
	hotellinghourfraction	Empty Table.	Empty Table.
	hotellingmonthadjust	Empty Table.	The ratio of the number of days in each month by (365/12).
Starts	startsperdaypervehicle	MOVES default.	Empty Table.
	startsageadjustment	Empty Table.	Empty Table.
	startsmonthadjust	*monthvmtfraction* * 12	*monthvmtfraction* * 12
	startshourfraction	Weekday: Based on the start output from the startspervehicle table, Weekday: MOVES default.	Empty Table.
Fuel Properties	fuelsupply	MOVES default (except usage for fueltype 5 = 0).	100% market share of the E10 ethanol blend in gasoline. Diesel had six ppm sulfur and 100% B5 biodiesel.
	fuelformulation	SIP fuel data for monthID 1 used for months 1-3, 11-12; monthID 7 data for months 5-9; MOVES default fuel formulation for non-winter and -summer months: 4, 10.	It is derived from refinery production compliance data, market fuel survey data, and known federal and local regulatory requirements.
	fuelusagefraction	Market shares of fuel formulations set to reflect Texas modeling assumptions of gasoline and diesel only, although all MOVES default fuels were included as required to run MOVES3 (i.e., CNG, E85, and electric are included but are not used as specified in the AVFT and fuel usage configurations).	MOVES default.
Meteorology: County Pressure	county	Averaged values from those reported by the National Oceanic and Atmospheric Administration's (NOAA) National Climatic Data Center (NCDC) and TCEQ's Texas Air Monitoring Information System (TAMIS).	MOVES default.
Meteorology: Hourly Temperature & Humidity	zonemonthhour	Averaged values from those reported by the National Oceanic and Atmospheric Administration's (NOAA) National Climatic Data Center (NCDC) and TCEQ's Texas Air Monitoring Information System (TAMIS) by county.	2020 temperature profile and average humidity data from the Weather Research and Forecasting Model (WRF) averaged by county using the Met4moves module in EPA's SMOKE program.

CDB Table Type	CDB Table	2020 AERR Data Source	2020 NEI Data Source
N/A	totalidlefraction	-	-
N/A	auditlog	-	-
N/A	countyyear	-	-
N/A	dayofanyweek	-	-
N/A	monthofanyyear	-	-
N/A	state	-	-
N/A	year	-	-
N/A	zone	-	-
N/A	zoneroadtype	-	-

2.5.2 Inputs

Figure 5 shows the percentage difference in CDB inputs of final and submitted inputs (final - submitted) of some of the important CDBs considering the data of all the counties. It can be seen that the total VMT considered by TTI and EPA are almost the same. The differences are negligible. The average car age in EPA data is almost 15% greater than the average car age in TTI submitted data. A large variation is observed in the total population for SUT 21 and 31 (passenger cars and passenger trucks) in submitted and final CDBs. The average speed bin of EPA is, on average, 1.5% less than the average speed bin of TTI input. On average, the hotelling hours used by EPA are 5% greater than those submitted by TTI.

Appendix A shows disaggregated differences in inputs for one county. The key takeaways with respect to the inputs are as follows:

- EPA directly used the total VMT submitted by TTI, which is further distributed according to MOVES source type. EPA redistributes the source type population of passenger cars and passenger trucks.
- The number of SUT 62 (long haul combination trucks) is greater in TTI-submitted CDBs.
- There are no significant differences between road-type distributions of VMT.
- For passenger cars (source type 21), TTI estimated larger fractions for newer vehicles. On the other hand, the EPA estimated a larger fraction for vehicles 5 to 10 years older.
- For passenger trucks, EPA estimated a larger fraction of newer vehicles, and TTI estimated a larger fraction of older vehicles.
- The speed bin distribution for passenger cars is observed to be almost similar. Minor differences are observed in the *AVFT* tables.
- The meteorological inputs are different in the submitted and final CDBs.

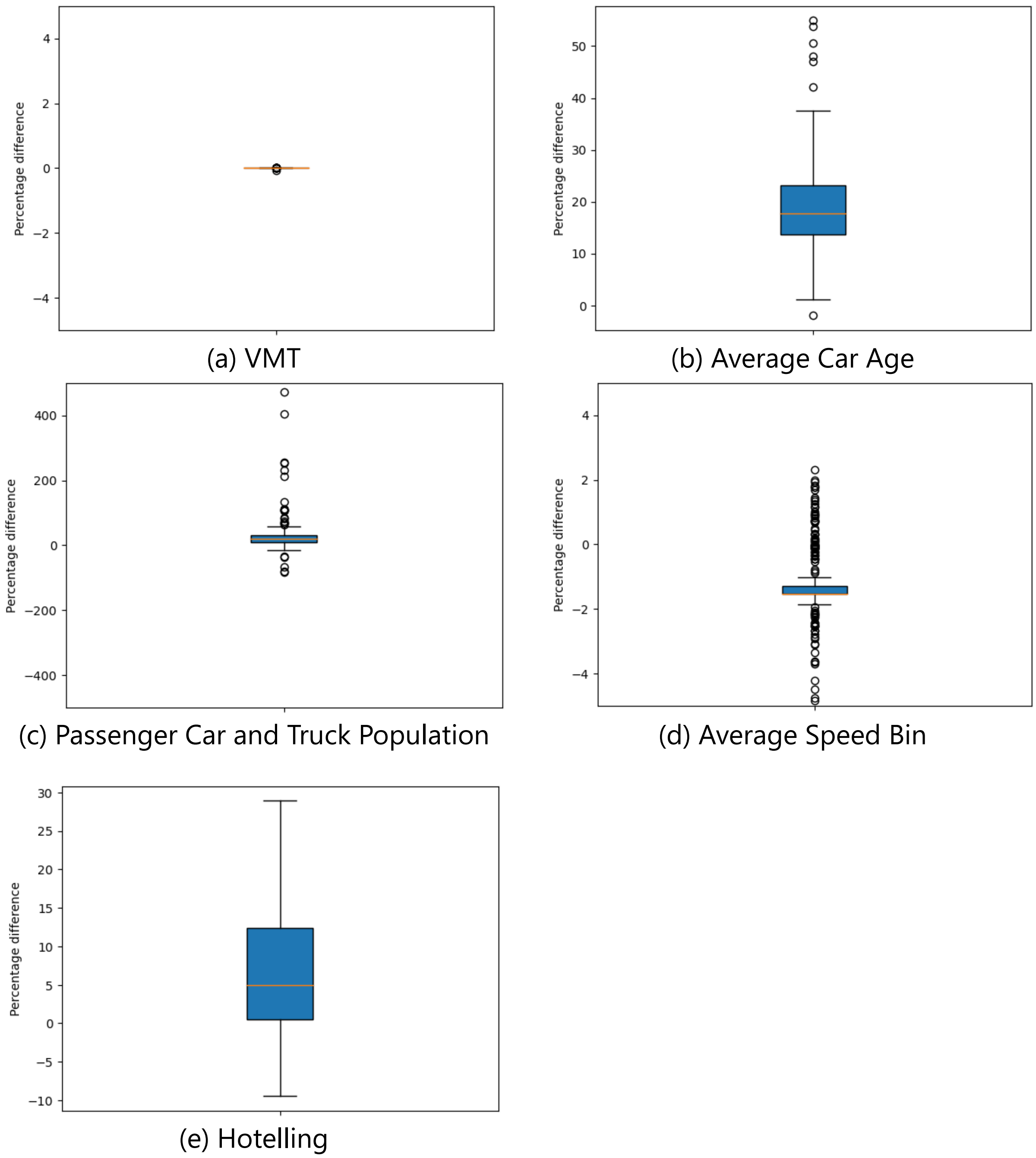


Figure 5. 2020 NEI Inputs Compared with 2020 AERR Inputs for 254 Texas Counties

2.6 Chapter Summary

The literature review provides a comprehensive literature review related to the MOVES3 technical guidance and methodology for developing annual MOVES inventory mode CDBs. The MOVES modeling framework is discussed, emphasizing the importance of accurate activity inputs in both inventory and emission rate modes. It covers the various domains of MOVES inputs and the significance of localized data. It provides guidance on using the MOVES model for populating county scale input database for on-road vehicle emissions estimation for SIP inventories and conformity analysis. EPA's CDBs for 2020 NEI were compared with the CDBs from 2020 AERR for the state of Texas submitted by TCEQ using TTI methodologies. The key differences in activity inputs and emission estimations between EPA and TTI methodologies in terms of CDB developments are highlighted as follows:

- EPA's spatial and temporal scope analysis was 12-kilometer gridded cells and an hour. TTI's were TDM links or county-level HPMS-based 21 "virtual links." For TDM counties, TTI has higher resolution, while for non-TDM counties, EPA analysis was more detailed.
- Among the traffic activity-related inputs, EPA directly adapted the total VMT from the submitted data. However, the VMT adjustment factors used to distribute VMT to the resolution of an hour are different.
- EPA adopted various data sources like IHS market data and StreetLight telematics data to develop the adjustment factors. Among the source type populations, EPA redistributed many of the passenger cars as passenger trucks.
- EPA used StreetLight telematics to develop average speed distributions. In contrast, TTI used MOVES default and TDM-based speed distributions. EPA methodology captured COVID impacts and was based on real-world data as opposed to TTI's methodology of using TDM outputs, HPMS-based "virtual links," and volume-capacity-based speed equations.
- EPA used national hotelling rates. In contrast, TTI used local hotelling data.
- EPA has adopted different meteorological inputs like temperature and relative humidity. EPA has also used different fuels, fuel usage fractions, and fuel formulations. Both fuels and meteorology impact emission rates and thus will produce different emission estimates.
- The combined effect of the differences in CDB development led to differences in emission estimates. For all the important pollutants like CO, CO_2, NO_X, and PM (both under 10 microns [PM_{10}] and 2.5 microns [$PM_{2.5}$]), the estimates of 2020 NEI are higher than those of 2020 AERR. For the non-attainment zones in Texas, NOx estimates by 2020 NEI are, on average, 20% greater than 2020 AERR.

3 ASSESSMENT PLAN

The previous chapter pointed out the differences in the emissions estimation methods of the important pollutants between the 2020 AERR and 2020 NEI studies. The emissions estimates of the EPA were up to 40% higher for some pollutants. To explain the differences in the emission estimates, the differences in the input CDBs of MOVES developed by TTI and EPA were analyzed. This chapter includes the developed assessment plan with a detailed list of the activity parameters assessed as part of the Task 4 work. Excluding well-established parameters like meteorology and fuels, the parameters in the test plan are source type population, age distribution, speed distribution, hotelling hours, and VMT road type distribution. The sensitivity analysis of the MOVES3 inventory mode CDB inputs will be conducted using the most recent 2020 AERR as well as the MOVES3 trend information developed by TTI. The expected value changes and the explanations on the nature of the activity parameter would be analyzed to identify which input parameters have the greatest sensitivity and impact on final emission estimates. The comparison with 2020 NEI CDBs and 2020 AERR CDBs would be able to determine how appropriate the use of MOVES3 default data when compared with Texas-specific local data.

This chapter details the assessment plan for the sensitivity analysis.

3.1 Simulation Details

The details of the MOVES simulation are as follows:

- MOVES Version 3.1
- Emission Inventory Mode
- Evaluation period: For each month of the year 2020
- Resolution: Per hour

The TTI research team considered the criteria air pollutants (CAP) and greenhouse gas (GHG) pollutants listed in Table 10 for the sensitivity analysis. Notably, for PM_{10} and $PM_{2.5}$, the aggregate emission estimates are based on different pollutant processes that are represented by separate pollutant IDs. For example, PM_{10} is the combination of MOVES pollutant ID 100: primary exhaust, ID 106: brakewear, and ID 107: tirewear.

Table 10. Pollutants Studied

Pollutant Name	MOVES Pollutant ID
CO	2
NO_x	3
VOC	87
PM10	100 + 106 + 107
PM2.5	110 + 116 + 117
CO_{2eq} (which includes CH_4(5), CO_2(90), and N_2O(6))[1]	98

[1]CO_{2eq} – carbon dioxide equivalent, CH_4 – methane, N_2O – nitrous oxide.

3.2 Representative Counties

Table 11 shows the representative counties which were used for the simulation. Counties from urban districts like Dallas, El Paso, Forth-Worth, Houston, Austin, and San Antonio were selected. Non-attainment areas were included. Coke County was selected to see the sensitivity for smaller counties.

Table 11. Representative Counties

District	County	County ID
Dallas	Collin	48085
El Paso	El Paso	48141
Fort Worth	Erath	48143
Bryan	Brazos	48041
Fort Worth	Tarrant	48439
Houston	Brazoria	48039
Houston	Harris	48201
Austin	Travis	48453
Corpus Christi	Aransas	48007
San Antonio	Atascosa	48013
San Angelo	Coke	48081
El Paso	Jeff Davis	48243

3.3 Scenario Development

Methodologies and outputs of NEI 2020 and AERR 2020 were compared as a part of the literature review. Differences in the key inputs, which were major contributors to the different emission estimates, were highlighted. These differences in the input CDB formed the basis of the sensitivity analysis.

This section explains the scenarios tested for the sensitivity analysis. The TTI research team divided the scenarios into three parts:

- base scenarios,
- simple scenarios, and
- interaction scenarios.

As shown in Table 12, base scenarios consisted of the benchmarks, which include original CDBs from AERR 2020 and NEI 2020. From the literature review, it was observed that fuel inputs and meteorology inputs were different for the NEI 2020 and AERR 2020 studies. Hence, the sensitivity of fuel and meteorology inputs were tested as a part of base scenarios. The inputs for hotelling activity and VMT allocation by vehicle category (i.e., MOVES source type versus HPMS vehicle type) were found to differ as well and were also tested within the base scenarios. Lastly, MOVES defaults were considered as one of the base scenarios.

Table 12. Base Scenarios with Descriptions and Short Names

Scenario Name	Description	Short name
AERR 2020	CDBs developed for the AERR 2020 study were used.	aerr
NEI 2020	CDBs developed for the NEI 2020 study were used.	nei
AERR 2020, except fuel inputs from NEI 2020	avft, fuelsupply, fuelusagefraction, and fuelformulation in AERR 2020 were replaced by corresponding CDB tables in NEI 2020. All other CDB tables were unchanged.	neifuels
AERR 2020, except meteorology inputs from NEI 2020	zonemonthhour and county table in AERR 2020 were replaced by corresponding CDB tables in NEI 2020. All other CDB tables were unchanged.	neimet
Hotelling hours	Replaced AERR 2020 hoteling hours with NEI 2020 hotelling hours.	neihotelling
Vehicle Miles of Travel (VMT): HPMS and MOVES source type.	Replaced the HPMS vehicle type VMT in AERR 2020 with the MOVES source type VMT from NEI 2020.	neivmt
AERR 2020, except fuel and meteorology inputs	avft, fuelsupply, fuelfraction, fuelformulation, zonemonthhour, and county in AERR 2020, were replaced by corresponding CDB tables in NEI 2020. All other CDBs were unchanged.	neifuelsmet
MOVES default	MOVES defaults were used. MOVES Runs were prepared through GUI by selecting appropriate inputs.	moves-defaults

Table 13 shows the simple scenarios with the important CDB tables selected for the sensitivity analysis. In simple scenarios, one input is changed at a time while keeping other inputs consistent with the AERR 2020. The range for each input varies between AERR 2020 and NEI 2020 values. The rationale behind limiting the inputs between these two benchmarks is to design scenarios between two real sets of values. Source type population, age distribution, speed distribution, hoteling hours, and VMT are the selected input CDB tables. The total VMT submitted by the TTI was directly used by the EPA for NEI 2020. However, EPA divided the total VMT submitted by TTI according to MOVES source type, whereas in the AERR 2020, TTI input VMT by HPMS vehicle type.

For the development of simple and interaction scenarios, the MOVES source types were divided into four groups as follows:

- LDV: Source types 21, 31, and 32 (Group 1)
- Bus: Source types 41, 42, and 43 (Group 2)
- Single Unit Trucks (ST): Source types 51, 52, 53, and 54 (Group 3)
- Combination Trucks (CT): Source types 61 and 62 (Group 4)

Table 13. Simple Scenarios with Descriptions and Short Names

Scenario Name	Description	Short name (i, group number)
Source type population	The population of source type in the source type group (LDV, Bus, ST, and CT) varied between the AERR 2020 and the EPA source type population. Each scenario reduced the difference in population for a source type group. All other CDB tables and **group populations were** kept the same as AERR 2020.	-
(Source type population 75% AERR and 25% NEI)	For each source type group, **75%** weightage was given to the AERR 2020 population, and **25%** was given to the NEI 2020 population. During one simulation, only one source type group population was changed, keeping the source type population of other groups equal to AERR 2020.	neipop25grp(i)
(Source type population 50% AERR and 50% NEI)	For each source type group, **50%** weightage was given to the AERR 2020 population, and **50%** was given to the NEI 2020 population. During one simulation, only one source type group population was changed, keeping the source type population of other groups equal to AERR 2020.	neipop50grp(i)
(Source type population 25% AERR and 75% NEI)	For each source type group, **25%** weightage was given to the AERR 2020 population, and **75%** was given to the NEI 2020 population. During one simulation, only one source type group population was changed, keeping the source type population of other groups equal to AERR 2020.	neipop75grp(i)

Scenario Name	Description	Short name (i, group number)
(Source type population 100% NEI)	For each source type group, **100%** weightage was given to the NEI 2020 population. During one simulation, only one source type group population was changed, keeping the source type population of other groups equal to AERR 2020.	neipop100grp(i)
Age distribution	For each source type group, the age distribution in the AERR 2020 was replaced by different distributions (see Figure 6), keeping other CDB tables and age distributions for other source type groups unchanged. **Note that age distribution was only changed for source types in one source type group for each simulation.**	-
Age distribution	For each source type group, the AERR 2020 age distribution was replaced by the NEI 2020 'sourcetypeagedistribution' table for each source type. Other CDB tables were kept the same as AERR 2020.	neiagedistgrp(i)
Age distribution	For each source type group, an average age distribution was calculated for each source type as shown in Figure 6. Other CDB tables and source-type group age distributions were kept the same as AERR 2020.	neiaerravgagedistgrp(i)
Speed distribution	For each road type, hourday ID, and source type in source type groups, the speed distribution of AERR 2020 was replaced with the speed distribution of NEI 2020. An example is shown in Figure 7. **Note that speed distribution was only changed for source types in one source type group for each simulation.**	neispeeddistgrp(i)
Road type distribution	For each source type group, the AERR 2020 roadtypedistribution table was replaced with the NEI 2020 roadtypedistribution table.	neiroaddistgrp(i)

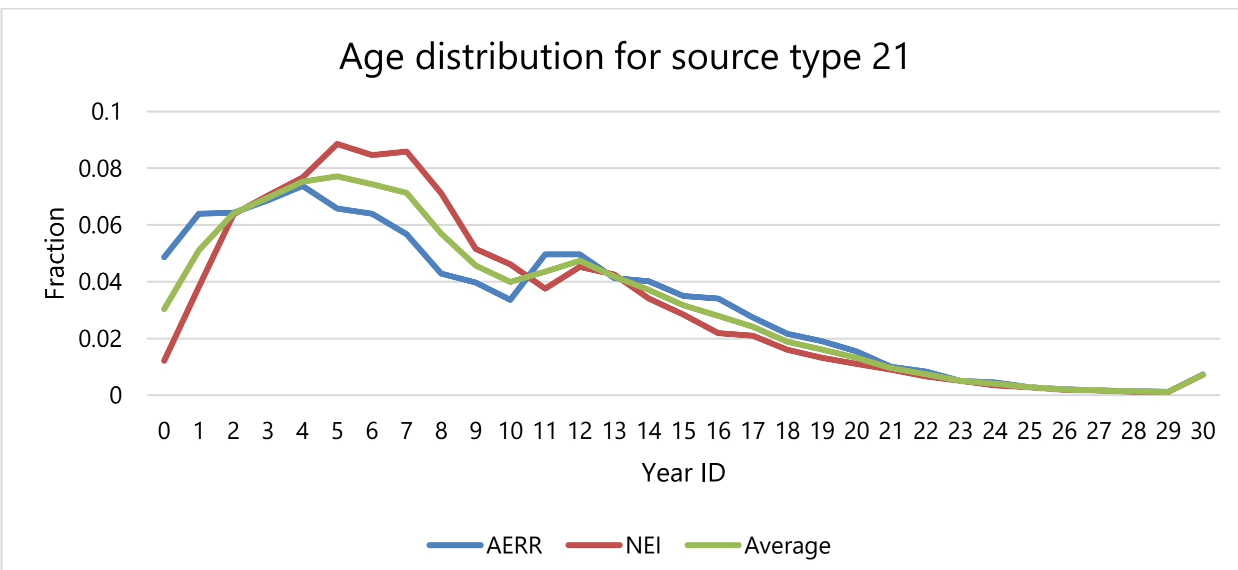


Figure 6. Average Age Distribution

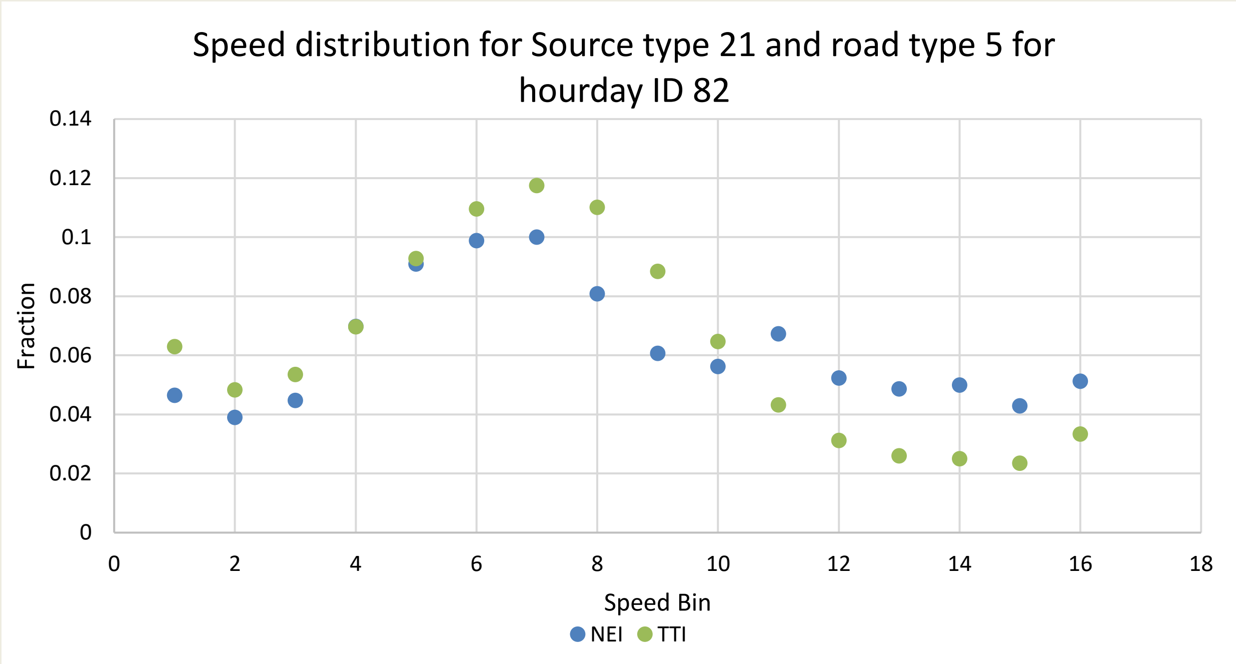


Figure 7. Speed Distribution Example

As shown in Table 14, two or more input CDB tables were varied simultaneously for the interaction scenarios. The effect of interactions between source type population, age distribution, speed distribution, and VMT was considered in the development of these scenarios.

Table 14. Interaction Scenarios with Descriptions and Short Names

Scenario Name	Description	Short name (i, group number)
Interaction of source type population and age distribution	For each source type group, the source type population and the age distribution were simultaneously varied.	-
(Source type population 75% AERR and 25% NEI) * Age distribution NEI	For each source type group, **75%** weightage was given to the AERR 2020 population, 25% weightage was given to the NEI 2020 population, and the NEI 2020 age distribution was used.	neipop25agedistgrp(i)
(Source type population 50% AERR and 50% NEI) * Age distribution NEI	For each source type group, **50%** weightage was given to the AERR 2020 population, 50% weightage was given to the NEI 2020 population, and the NEI 2020 age distribution was used.	neipop50agedistgrp(i)
(Source type population 25% AERR and 75% NEI) * Age distribution NEI	For each source type group, **25%** weightage was given to the AERR 2020 population, 75% weightage was given to the NEI 2020 population, and the NEI 2020 age distribution was used.	neipop75agedistgrp(i)

Scenario Name	Description	Short name (i, group number)
(Source type population 100% NEI) * Age distribution NEI	For each source type group, **100%** weightage was given to the NEI 2020 population, and the NEI 2020 age distribution was used.	neipop100agedistgrp(i)
Interaction of source type population and speed distribution	The source type population and speed distribution were simultaneously varied.	-
(Source type population 75% AERR and 25% NEI) * Speed distribution NEI	For each source type group, **75%** weightage was given to the AERR 2020 population, 25% weightage was given to the NEI 2020 population, and NEI 2020 speed distributions were used.	neipop25speeddistgrp(i)
(Source type population 50% AERR and 50% NEI) * Speed distribution NEI	For each source type group, **50%** weightage was given to the AERR 2020 population, 50% weightage was given to the NEI 2020 population, and NEI 2020 speed distributions were used.	neipop50speeddistgrp(i)
(Source type population 25% AERR and 75% NEI) * Speed distribution NEI	For each source type group, **25%** weightage was given to the AERR 2020 population, 75% weightage was given to the NEI 2020 population, and NEI 2020 speed distributions were used.	neipop75speeddistgrp(i)
(Source type population 100% NEI) * Speed distribution NEI	For each source type group, **100%** weightage was given to the NEI 2020 population, and NEI 2020 speed distribution was used.	neipop100speeddistgrp(i)
Interaction of source type population, age distribution, and speed distribution	The source type population and speed distribution were simultaneously varied.	
(Source type population 75% AERR and 25% NEI) * (Age distribution NEI) * Speed distribution NEI	For each source type group, **75%** weightage was given to the AERR 2020 population, and **25%** weightage was given to the NEI 2020; NEI 2020 age distribution was used, and NEI 2020 speed distributions were used.	neipop25speedagedistgrp(i)
(Source type population 50% AERR and 50% NEI) * (Age distribution NEI) * Speed distribution NEI	For each source type group, **50%** weightage was given to the AERR 2020 population, and **50%** weightage was given to the NEI 2020 population; NEI 2020 age distribution was used, and NEI 2020 speed distributions were used.	neipop50speedagedistgrp(i)
(Source type population 25% AERR and 75% NEI) * (Age distribution NEI) * Speed distribution NEI	For each source type group, **25%** weightage was given to the AERR 2020 population, and **75%** weightage was given to the NEI 2020 population; NEI 2020 age distributions were used, and NEI 2020 speed distributions were used.	neipop75speedagedistgrp(i)
(Source type population 100% NEI) * (Age distribution NEI) * Speed distribution NEI	For each source type group, **100%** weightage was given to the NEI 2020 population, NEI 2020 age distributions were used, and NEI 2020 speed distributions were used.	neipop100speedagedistgrp(i)

3.4 Chapter Summary

Chapter 3 of the document focuses on the assessment plan for sensitivity analysis using MOVES3. Here's an extended summary of the plan:

- **Overview of the difference**: Outlines the differences in emissions estimation methods between the 2020 AERR and 2020 NEI studies, highlighting the significant variances in the input CDBs of MOVES developed by TTI and EPA.
- **Setup simulation details**: Describes the MOVES3.1 model setup used for the analysis, focusing on the evaluation period, resolution, and the pollutants considered for the sensitivity analysis.
- **Select representative counties**: Lists the selected counties for simulation, including urban districts and non-attainment areas, to gauge the sensitivity of smaller and larger regions.
- **Develop scenarios:** Discuss the methodologies and key input differences between NEI 2020 and AERR 2020 studies. It details the division of scenarios into base, simple, and interaction scenarios for sensitivity testing, including variations in fuel inputs, meteorology, hotelling activity, and VMT allocation.
- **Design detailed scenario analysis**: Provides in-depth analysis of different scenarios, including altering source type populations, age distributions, speed distributions, and road type distributions. It also examines the interaction effects between these variables.

The chapter is integral in understanding the nuances and variations in emission estimates, illustrating the importance of accurate and localized data inputs in environmental modeling. All MOVES3 CDBs and MRS files for all the scenarios are included in Appendix B. The next chapter discusses the results of the assessment plan.

4 SENSITIVITY ASSESSMENT RESULTS

This chapter discusses the details of the sensitivity study and its results and findings as part of Task 4. A total of 1,056 MOVES runs were carried out for twelve representative counties. The first section delves into how emissions sensitivity varies with different datasets. It contrasts emission discrepancies across various scenarios, particularly focusing on the substitution of NEI or MOVES default datasets for AERR's Commercial Database data. Key areas of difference include population, age and speed distribution, meteorology, fuel inputs, VMT distributions, and hotelling inputs. This section primarily emphasizes the relative emission disparities across scenarios rather than quantifying input differences. The final section examines how different input variables, particularly vehicle population, affect pollutant emissions. The analysis, focusing on group 1 vehicles (source types 21, 31, and 32) and group 4 vehicles (source types 61 and 62), revealed notable variations in emission responses to population changes among these types. Regression models applied to data from various counties confirmed the statistical significance of these findings, suggesting the model's applicability across diverse county characteristics. Key observations and conclusions were summarized in the final section of the chapter. All sensitivity results figures are included as electronic files in Appendix C.

4.1 SENSITIVITY WITH RESPECT TO SCENARIOS

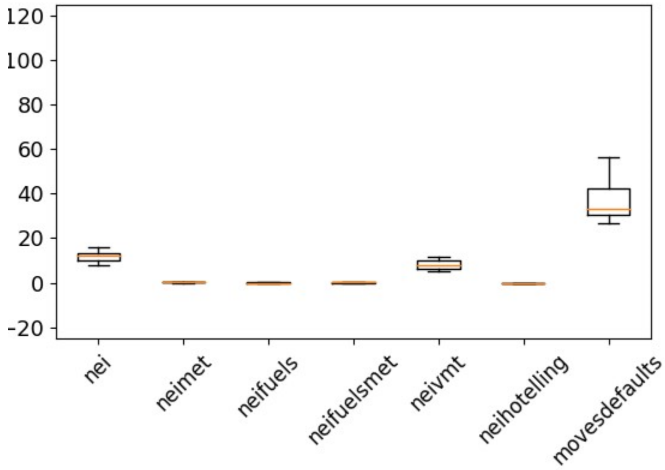
The sensitivity with respect to scenarios is based on the following formula:

$$Percentage\ Difference = \frac{Scenario\ Emissions - AERR\ Base\ Scenario\ Emissions}{AERR\ Base\ Scenario\ Emissions} \times 100\%$$

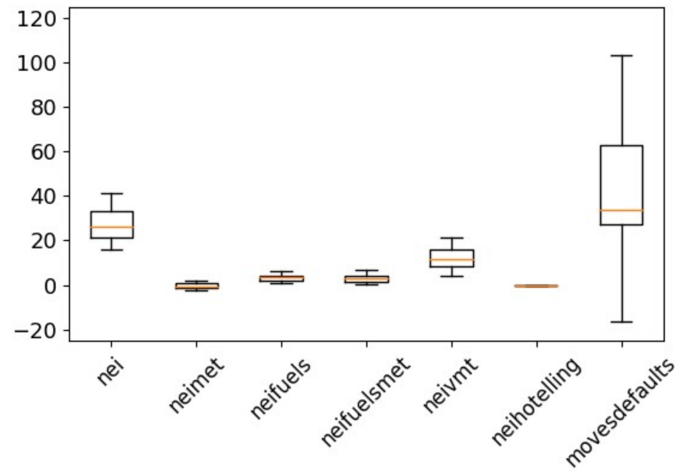
The following sections describe the findings based on the formula for all base, simple, and interactive scenarios.

4.1.1 Base Scenarios

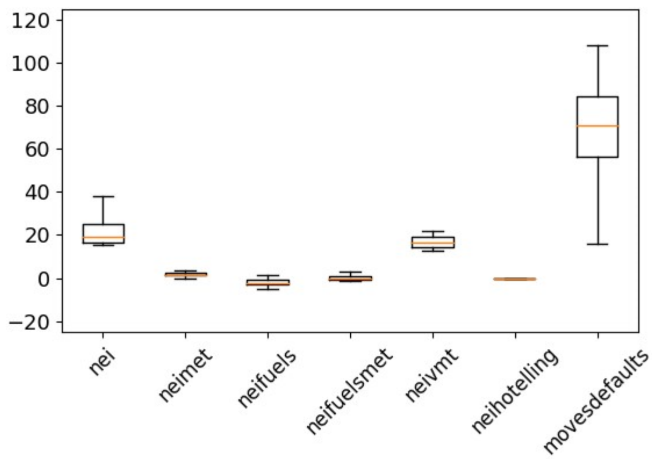
Figure 8 gives an overview of base scenario results for six pollutants. The x-axis contains labels of the short names of scenarios as described in Table 12. The y-axis is the percent difference between AERR emissions and the scenario emissions for that specific pollutant. The percentage difference of each base scenario is plotted as a range box for twelve representative counties. The leftmost box shows the NEI scenario, and the rightmost box shows MOVES defaults.



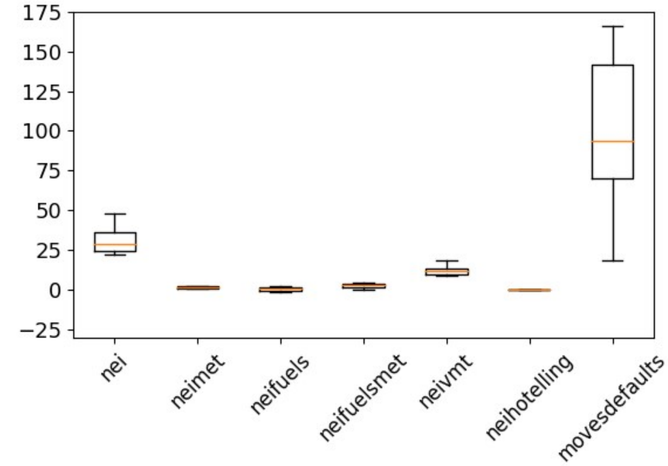
(a) CO2 Equivalent



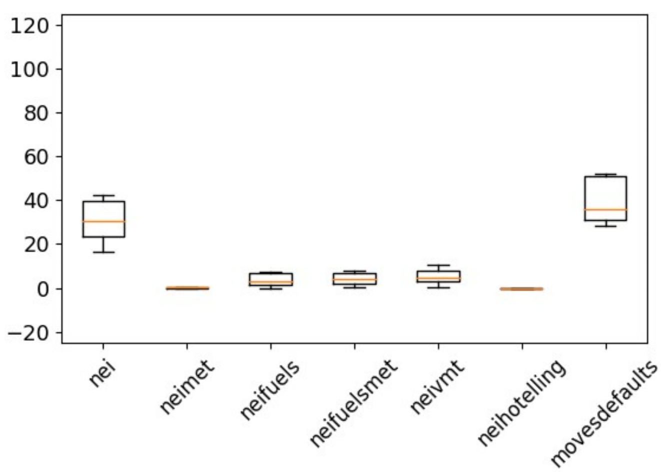
(b) NOx



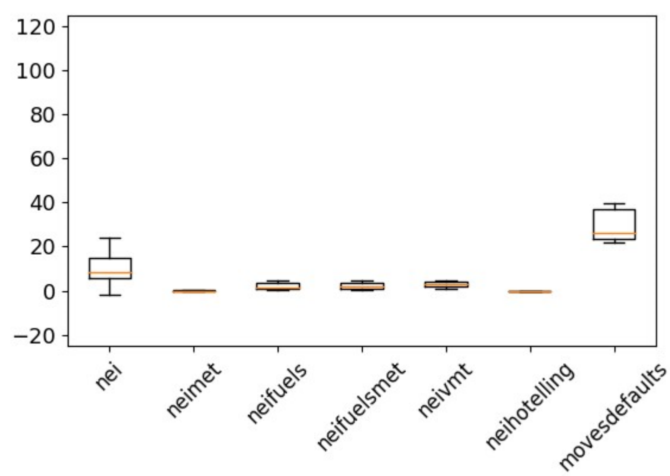
(c) CO



(d) VOC



(e) PM2.5



(f) PM10

Figure 8. Percentage Difference of Different Pollutants of Base Scenarios (Percent Difference on Y-axis)

The following conclusions can be drawn from the plots:

- A trend was observed for all the pollutants. The highest percent difference was observed for the MOVES defaults scenario, followed by the NEI scenario.
- For all the pollutants, the median difference was greater than 10% for NEI and greater than 25% for MOVES defaults.
- The difference for VOCs in the MOVES defaults scenario was close to 100%. This shows that the difference in the datasets is a major factor accounting for the difference in emission estimates. MOVES defaults use national averages for some inputs. The large differences in MOVES default scenarios can be attributed to national averages and county-specific data differences.
- An interesting observation was noted in the *neivmt* scenario where EPA kept total HPMS VMT the same as AERR but inputted the VMT by MOVES source type. This caused a greater than 10% difference for CO, NO_x, and VOCs.

For meteorology, fuels, and hoteling scenarios, the emission difference was less than 5% for all the pollutants. This shows that these inputs have little impact on the emissions, though they differ in NEI and AERR datasets. For the base scenarios, a major difference in emissions relative to the AERR scenarios was observed for MOVES defaults and NEI scenarios. There could be several possible reasons. The emission difference could be the interaction effect of several inputs. For example, if the TTI research team compare NEI and AERR, the source type populations are derived from different data sources and are considerably different. Speed and age distributions are different. Though total VMT for source types are the same, they are distributed differently.

On the other hand, when meteorology and fuel inputs CDB tables were changed, the emission difference was negligible. The literature review showed that data sources used for fuels and meteorology differ for NEI and AERR. However, the differences in these datasets did not lead to considerable differences in emissions. This suggests that these inputs have lesser sensitivity towards emissions.

4.1.2 Quality Assure Source Type Group Results for Simple and Interaction Scenarios

The simple and interaction scenarios were designed based on source type groups and input parameters. Before looking at differences in emissions in various scenarios, the TTI research team designed that if the input was changed for a specific source type group, emissions and activity for only that source type group were changed in the output. For example, in the *neipop100grp1* scenario, only the population inputs of source type group 1 (source types 21,31, and 32) were changed from the AERR 2020 populations to the NEI 2020 populations. All the other input tables for groups 2, 3, and 4 were kept constant. To ensure that the emissions and activity of only the scenario's test group were varied, emissions and activity outputs from movesoutput tables were aggregated to obtain quantities as described in Figure 9. For each scenario, the TTI research team used this technique to verify that the differences in emissions results were only caused by the group for which input data were changed.

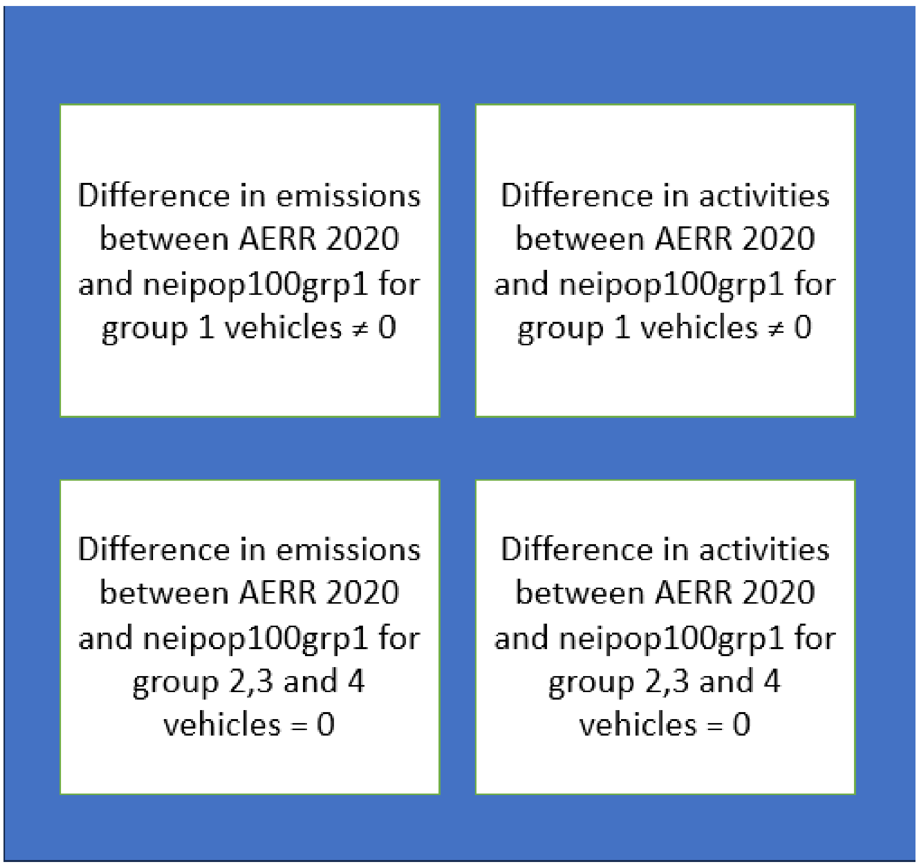


Figure 9. Verifying Results for Source Type Group 1 Relative to the Remaining Source Type Groups for the neipop100grp1 Scenario

4.1.3 Simple and Interaction Scenarios

This section shows the results of sensitivity analysis for the simple and interaction scenarios. For Figure 10 through Figure 18, the y-axis is the percent differences, and the x-axis shows the scenario's short names. The first four boxplots, or "ticks," show the simple population scenarios where the weightage of the NEI population is increased from 25% to 100%. Tick 5 to tick 8 show interaction scenarios of population and age distribution. Tick 9 to tick 12 show interaction scenarios of population and speed distribution. Tick 13 to Tick 16 shows the population, age, and speed distribution interaction scenario. The last four ticks show simple scenarios of age distribution, average age distribution, speed distribution, and road distribution, respectively.

4.1.3.1Group 1 (LDV) Scenarios

Figure 9 shows the results of the sensitivity analysis for Group 1 LDVs. Figure 10 shows the emission difference for CO2eq using AERR results as the baseline. As the NEI population's weight increased from 0 to 100 percent, the median percentage difference in emissions in simple population scenarios increased from 0 to 10 percent. A similar trend was observed for all three interaction scenarios. The percent difference for the simple speed distribution scenario was slightly greater than 5 percent. However, for the simple age distribution and road distribution scenarios, the percentage differences were negligible. This shows that population is a major factor that explains the emission difference between NEI and AERR estimates.

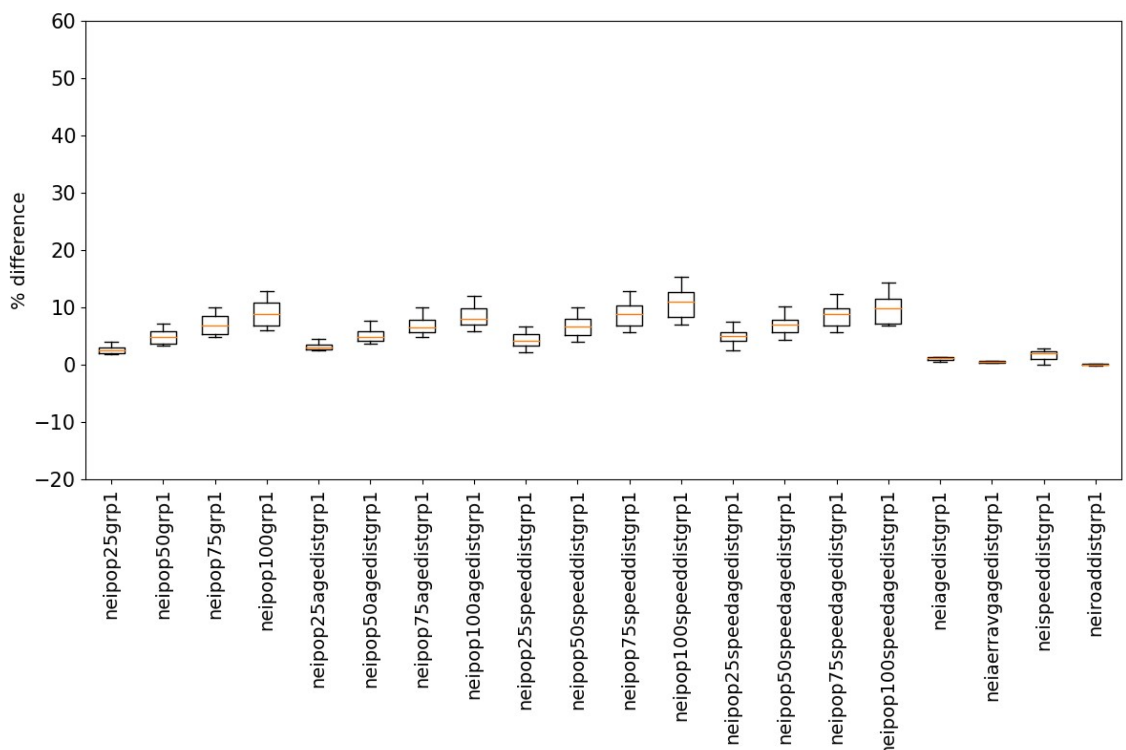


Figure 10. Percentage Differences of CO2eq for LDV (Group 1) Scenarios

Figure 11 shows the difference in NOx emissions using AERR results as the baseline. The same trend was observed as that of CO2eq. The percent difference increased from 0 to 20 percent as the weightage of the NEI population increased from 0 to 100 percent in simple population scenarios. However, the percent difference grew from 0 to 10 percent in the population and age interaction scenarios. For the age distribution scenario, the percentage difference was negative for the NEI age distribution. These two results show that the NEI age distribution negatively affected NOx emissions when compared to AERR NOx emissions. The percent difference was negligible for simple speed distribution and road distribution scenarios.

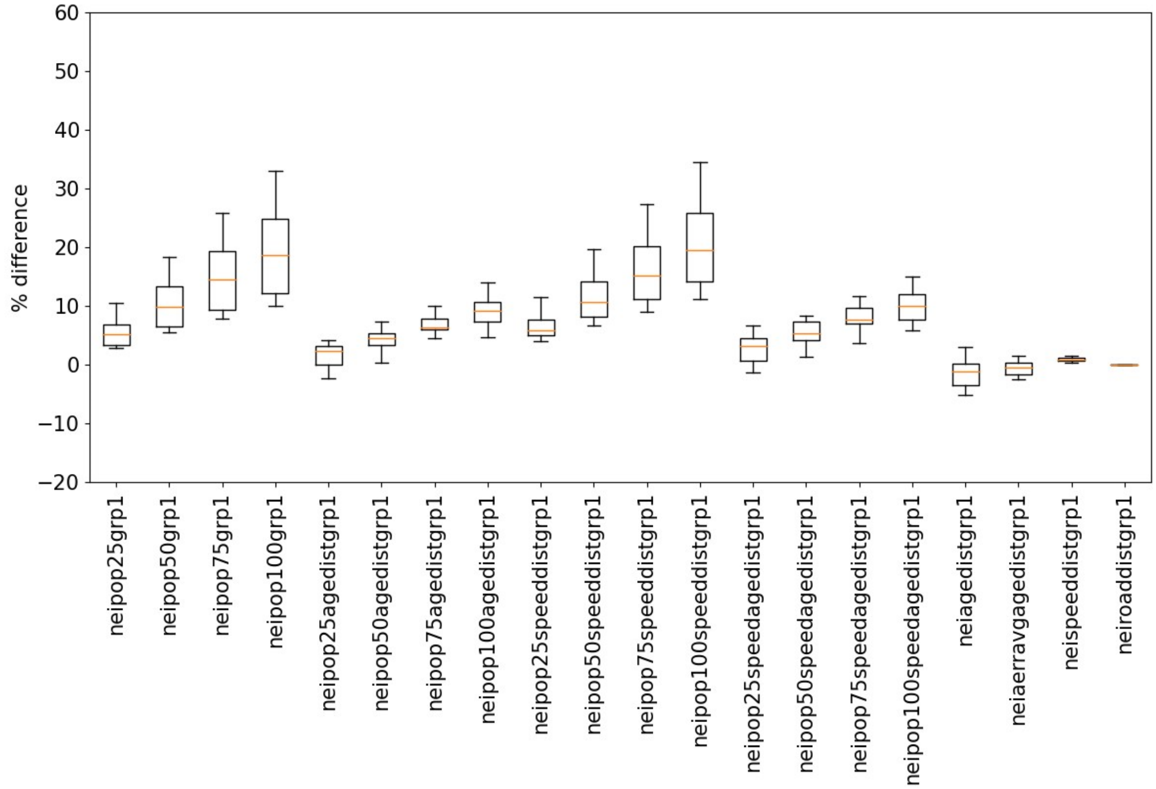


Figure 11. Percentage Differences of NOx for LDV (Group 1) Scenarios

Figure 12 shows the emission difference for CO using AERR results as the baseline. A similar trend was observed in this comparison, and the percentage difference increased from 0 to 20 percent as the NEI population's weight increased from 0 to 100 percent in simple population scenarios. The percent difference in speed distribution was greater than 5 percent in simple speed distribution scenarios. When combined, the difference increased from 0 to 30 percent for the population and speed interaction scenarios. The percent difference between age distribution and road distribution scenarios was nearly zero.

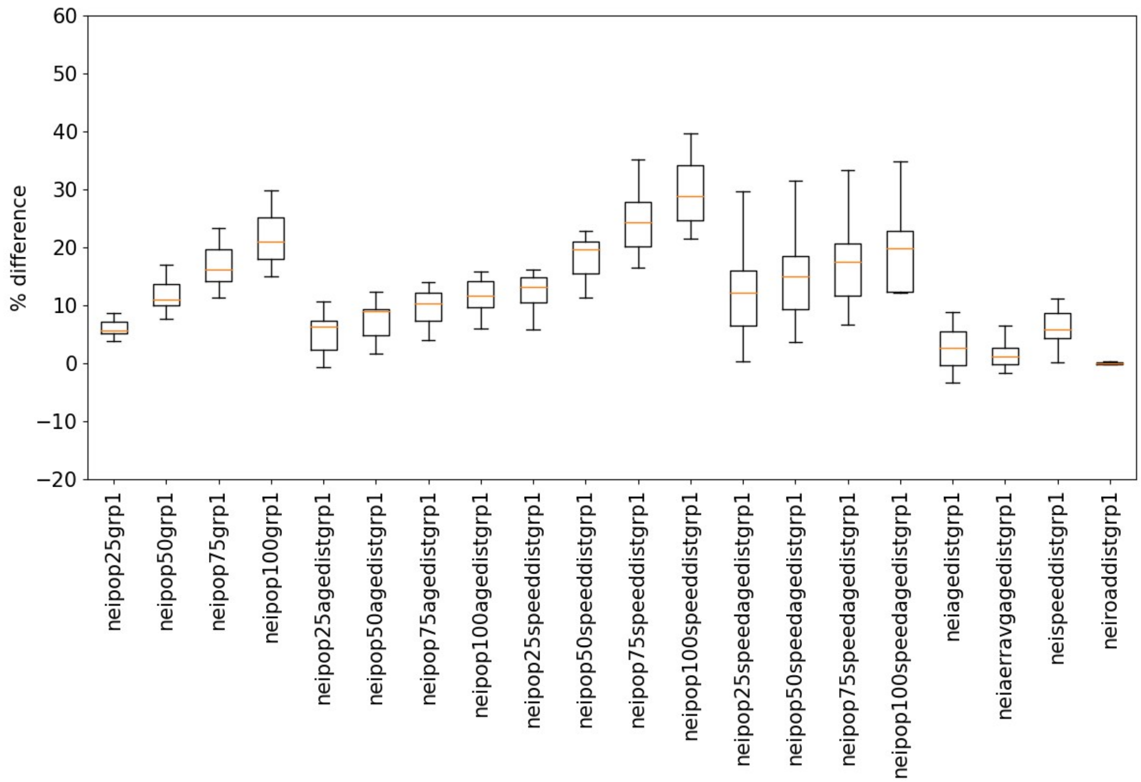


Figure 12. Percentage Differences of NOx for LDV (Group 1) Scenarios

Figure 13 and Figure 14 show the percent difference for $PM_{2.5}$ and PM_{10}, respectively. The same trend was observed for population and interaction scenarios in the case of $PM_{2.5}$ and PM_{10}. For $PM_{2.5}$, NEI age distribution had a negative effect on the percent difference in simple age distribution scenarios. In the case of PM_{10}, speed distribution had a negative effect on emissions, as shown in simple PM_{10} scenarios. Parts of the boxplots were below zero for the interaction scenarios of population and speed distribution. For both $PM_{2.5}$ and PM_{10}, the percent difference in emissions for age, road, and speed distribution was less than 5 percent.

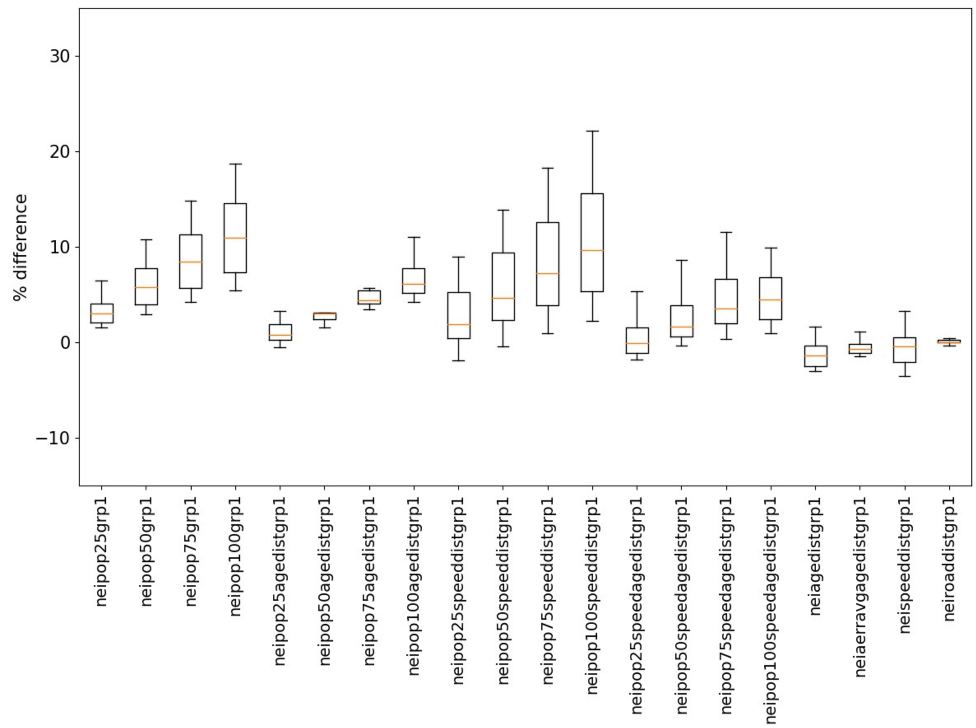


Figure 13. Percentage Differences of PM2.5 for LDV (Group 1) Scenarios

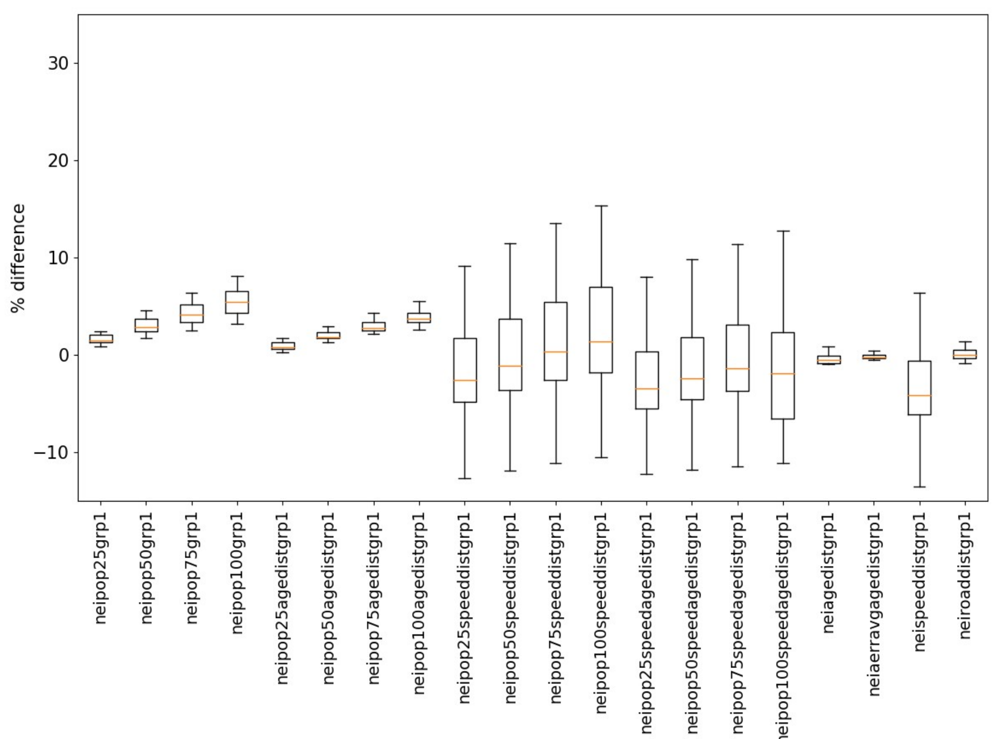


Figure 14. Percentage Differences of PM10 for LDV (Group 1) Scenarios

Figure 15 shows the percent difference for VOCs. It should be noted that variations in the difference in VOC emissions were greater than those of other pollutants, which is indicated by the range of the box plots. For a simple population scenario, the difference is up to 30 percent in emission differences. The same trends were observed for population as well as interaction scenarios. If simple population and interaction scenarios were compared, simple population scenarios had higher VOC emissions than interaction scenarios of population and age. This aligned with the effect of simple age distribution scenarios having negative effects on VOCs, which had a large part of the boxplot below zero. The percent difference in simple speed distribution and road distribution scenarios was less than 5 percent.

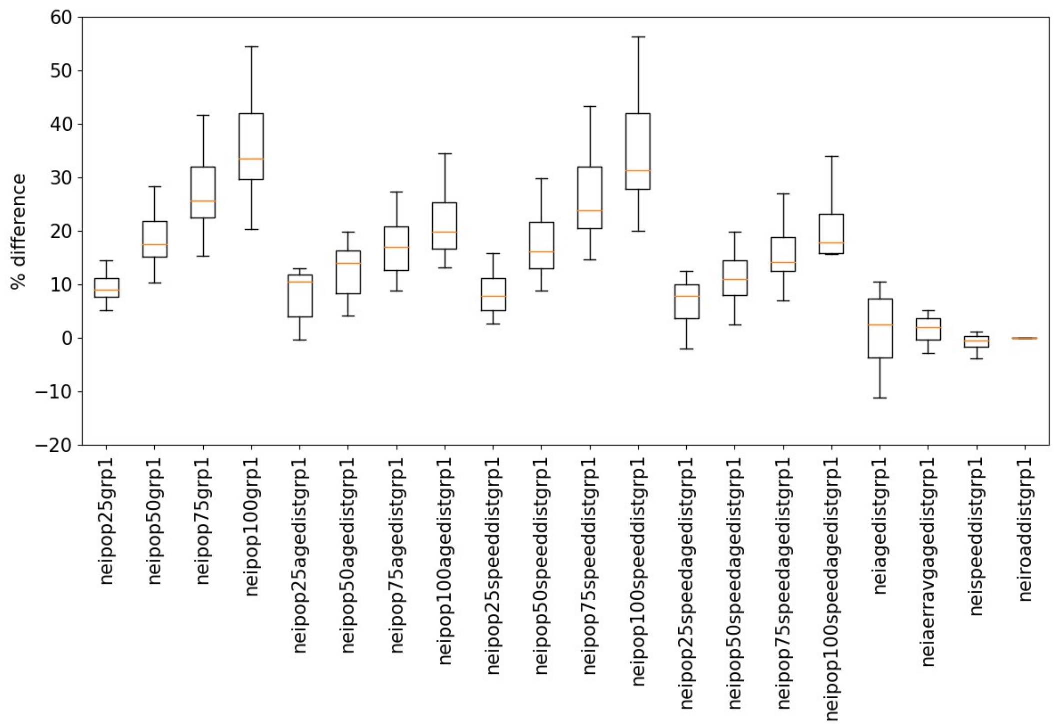


Figure 15. Percentage Differences of VOC for LDV (Group 1) Scenarios

For Group 1 (LDV), the TTI research team observed a trend in emission differences for all pollutants in terms of vehicle population change and age distribution change from AERR to NEI. As the population varied from AERR to NEI, emissions significantly increased and were highly sensitive to population changes. EPA utilized IHS registration and telematics data to develop CDBs related to source type population and age distribution, while AERR used TxDMV registration data. On the other hand, in the literature review, the TTI research team pointed out that the average car age in EPA data is almost 15% greater (i.e., older) than the average car age in AERR. A large variation was observed in the total population for Group 1 scenarios with source types 21 and 31 (passenger cars and passenger trucks) in NEI and AERR CDBs. EPA redistributed many of the passenger cars as passenger trucks. The average speed bin of NEI was, on average, 1.5% less (slower) than the average speed bin of AERR. EPA adopted various data sources like IHS market data and StreetLight telematics data to develop the VMT distribution factors. All in all, for LDVs, the higher number of passenger trucks, older age, and slower speeds explain the higher emissions of considered scenarios.

4.1.3.2Group 2 (BUS) and Group 3 (ST) Scenarios

For buses and STs, the TTI research team observed that the percent differences of all pollutants were less than 5 percent to negligible for every scenario. This is due to the variations in these two groups being too small in terms of the sensitivity analysis. The results from scenarios from these two groups will not be discussed in detail in this report.

4.1.3.3Group 4 (CT) Scenarios

Figure 16 to Figure 18 show the results of the sensitivity analysis for Group 4 scenarios with CT for selected pollutants. Figure 16 shows the difference in NOx emissions percentages. For the simple population scenario, a reverse trend was observed, similar to that of Group 1 LDVs. As the NEI population's weightage increases from 0 to 100 percent, the NOx percent difference decreases from 0 to 2.5 percent. The NOx percent difference in the age distribution scenario is 7 percent, and that in the speed distribution scenario is 2.5 percent. In the interaction scenario neipop25speedagedistgrp4, the percent difference was about 12 percent. The interaction effect of population, speed, and age distribution differences led to higher emissions than the simple scenarios where these three were changed individually.

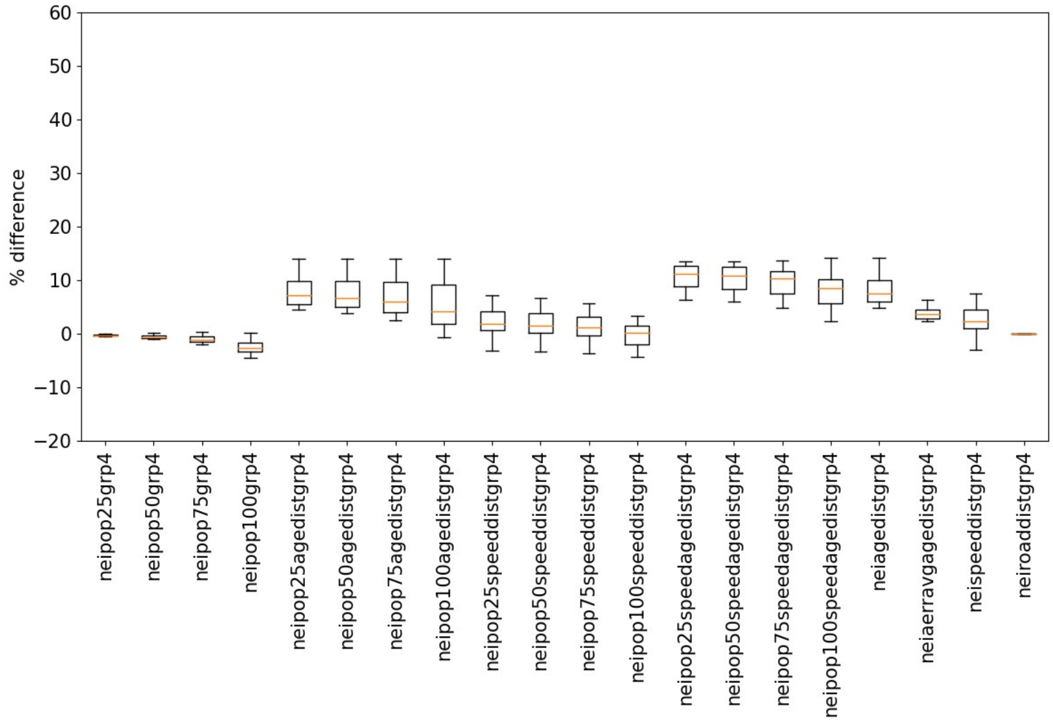


Figure 16. Percentage Differences of NOx for CT (Group 4) Scenarios

Figure 17 and Figure 18 shows the analysis for $PM_{2.5}$ and PM_{10} for Group 4 scenarios. Decreasing trends for the simple population scenarios and interaction scenarios were observed for both $PM_{2.5}$ and PM_{10}. The percentage difference of $PM_{2.5}$ in the age distribution scenario is around 13 percent. This observation is in line with the literature review that AERR has a higher number of long-haul combination trucks. The effect of speed distribution and road distribution were negligible for PM emissions.

For Group 4 scenarios, the trend is more complicated. The source registration data in the EPA-developed CDBs does not reliably distinguish between short-haul and long-haul activity. Source types 52 and 53 (single-unit trucks) have the same age distributions as those of source types 61 and 62 (combination-unit trucks). In addition, all age distributions for long-haul trucks (source types 53 and 62) are a national average because these vehicles are expected to travel long distances from the county where they are registered. This could be a possible explanation for the significant emission difference in the age-distribution scenario.

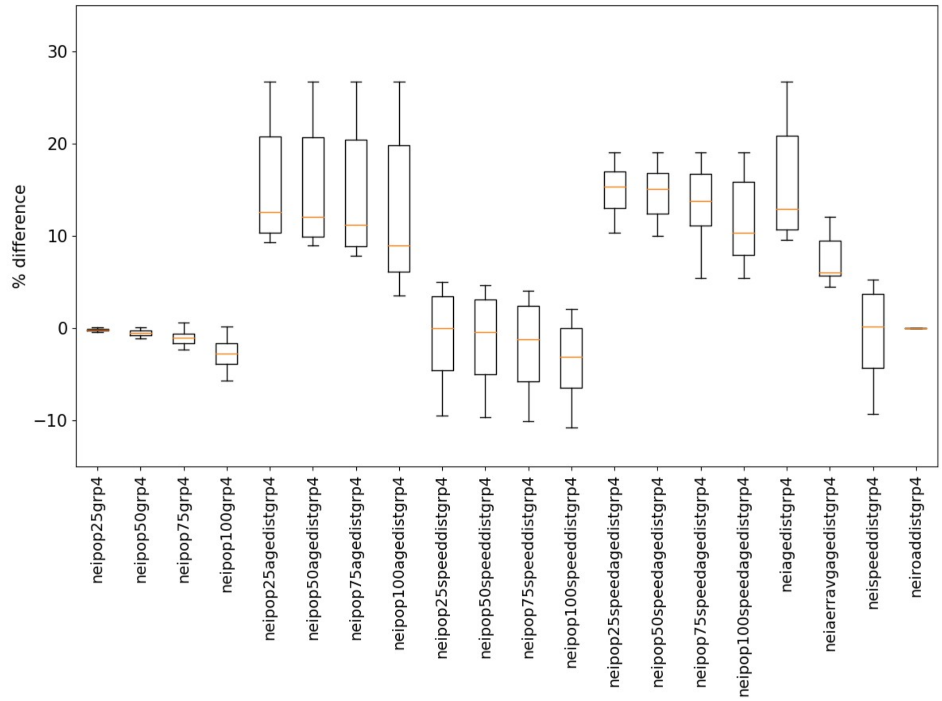


Figure 17. Percentage Differences of $PM_{2.5}$ for CT (Group 4) Scenarios

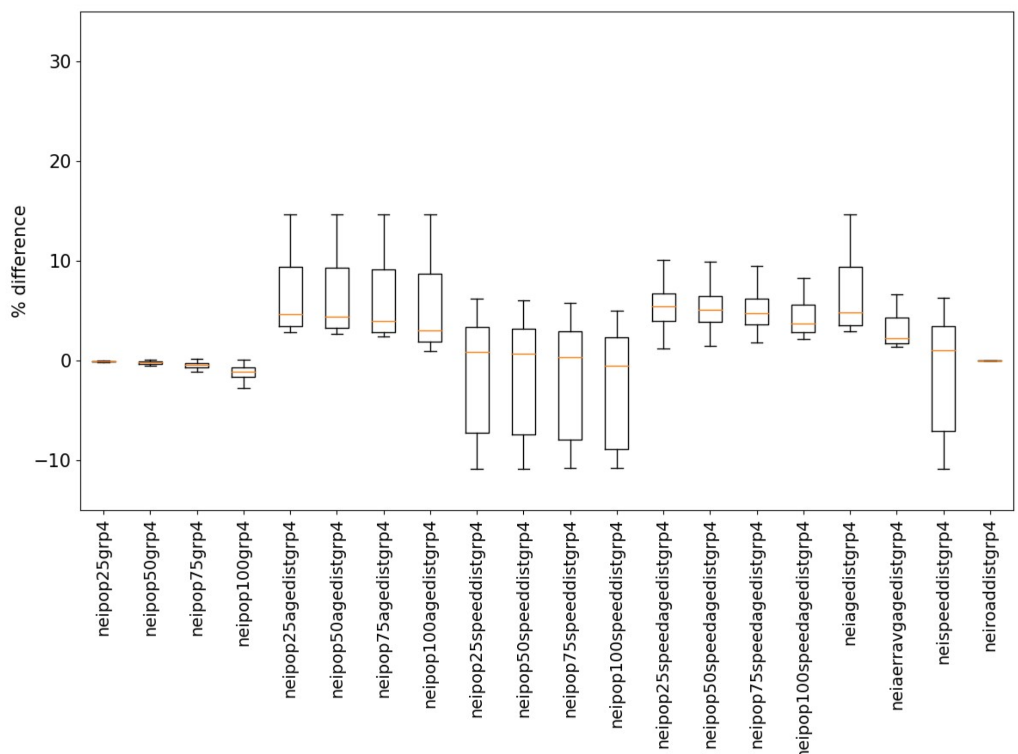


Figure 18. Percentage Differences of PM_{10} for CT (Group 4) Scenarios

4.1.3.4Scenario Sensitivity Levels

The sensitivity plots can help us analyze the impact of inputs on each variable. If the TTI research team want to improve the accuracy of a pollutant, more sensitive inputs should be given more attention. Table 15 and Table 16 gives an overview of sensitivity analysis. The following sensitivity levels were indicated in the tables for all the scenarios in terms of emission differences calculated based on AERR:

- **0-5%: Negligible (n)**
- **5-20%: Medium (m)**
- **>20%: Substantial (s)**

Table 15. Sensitivity Overview: Base Scenarios

Pollutants	CO_{2eq}	NOx	CO	$PM_{2.5}$	PM_{10}	VOC
NEI	m	s	m	s	m	s
NEI Meteorology	n	n	n	n	n	n
NEI Fuels	n	n	n	n	n	n
NEI Meteorology and Fuels	n	n	n	n	n	n
NEI VMT	m	m	m	n	n	m
NEI Hotelling	n	n	n	n	n	n
MOVES Defaults	s	s	s	s	s	s

In Table 16, source type groups Bus and ST were excluded because the sensitivity for those is negligible for all pollutants.

Table 16. Sensitivity Overview: Simple and Interaction Scenarios

Pollutants	CO2eq		NOx		CO		PM2.5		PM10		VOC	
Source type groups	LDV	CT	LDV	CT	LDV	CT	LDV	CT	LDV	CT	LDV	CT
NEI Population	m	n	m	n	s	n	m	n	m	n	s	n
NEI Age distribution	n	n	n	m	n	n	n	m	n	n	n	n
NEI Speed distribution	n	n	n	n	m	n	n	n	n	n	n	n
NEI Road distribution	n	n	n	n	n	n	n	n	n	n	n	n
NEI Population* Age distribution	m	n	m	n	m	n	m	m	n	n	s	n
NEI Population* Speed distribution	m	n	m	n	s	n	m	n	n	n	s	n
NEI Population* Age distribution * Speed distribution	m	n	m	m	m	n	n	m	n	n	m	n

For considered pollutants, NEI showed higher emissions than AERR. When MOVES default inputs were used, emissions were considerably higher. When meteorology and fuels data from NEI was used, the difference in emissions was less, which shows that for Texas, emissions differences between NEI and AERR are not very sensitive to these input differences. More attention should be given to developing input CDBs for which sensitivity is substantial (>20%) and medium (5-20%).

Emission estimates only using MOVES defaults were neither close to NEI nor AERR. Hence, using MOVES defaults is not recommended. It is crucial to develop county-specific inputs instead of using defaults representing national averages for count-level SIP emissions inventories.

Higher emissions were observed when VMT was distributed according to MOVES source types while keeping the total HPMS VMT the same. Hence, it is important to focus on how the VMT is distributed and input for various source types.

Among all four groups, the major difference was caused by the LDVs source type group (Group 1). For LDVs, the emissions gradually increased as the population shifted from AERR to NEI. NEI age distribution has a negative effect on NOx and PM2.5. NEI speed distribution has a positive effect on CO. NEI speed distribution has a negative effect on PM10. For buses and STs, the percent difference is close to zero for all the pollutants. This suggests that more focus should be given to improving the estimates for LDVs instead of Buses and STs. NOx, PM2.5, and PM10 emissions were sensitive for CTs in terms of age distribution.

Large variations in the emissions differences were observed across counties. This suggests that accurate input data for individual counties is crucial to get accurate estimates. However, a general trend is observed in the sensitivity analysis based on which conclusions were drawn. The next section will discuss the general regression trend on the sensitivity of emission estimates on input parameters.

4.2 Sensitivity with respect to Input Parameters: Regression Analysis

In this section, the sensitivity of pollutants with respect to each input variable is analyzed. Based on the results from Section 4.1, the vehicle population and age distribution are highly sensitive to the emission difference between NEI results and AERR results. In terms of the groups, the absolute difference in Group 1 and Group 4 is most significant compared to other groups. The TTI research team chose 1) vehicle population in Group 1 scenarios and 2) age distribution in Group 4 scenarios as the cases and studied their detailed impact on emissions.

A point to note here is that in NEI and AERR input datasets, the total VMT was the same for HPMS source types. For example, in the *neipop100grp1* scenario, if VMT allocated to source types 21, 31, and 32 were added, they were equal to the VMT allocated to the HPMS source type of LDVs.

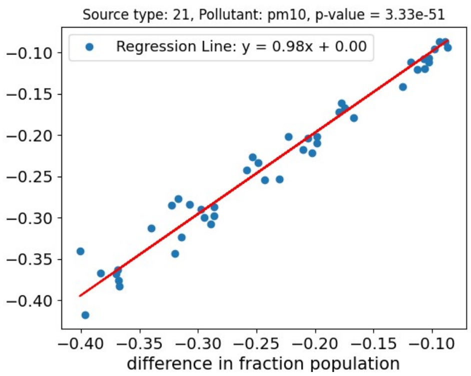
For each source type in the group, the emission sensitivity to the source type is calculated as follows:

$$sensitivity = \frac{Difference\ in\ emission\ percentage\ by\ source\ type}{Difference\ in\ population\ percentage\ by\ the\ source\ type}$$

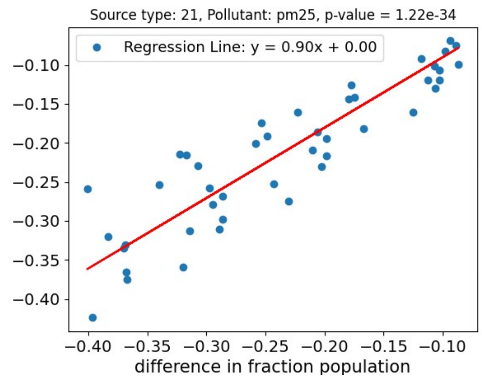
4.2.1.1 Vehicle Population in Group 1 (LDV) Scenarios

In Figure 19 through Figure 21, the x-axis is the difference in population fraction for source type 21, and the y-axis is the difference in emission fraction for the pollutant. The data of 11 representative counties were plotted and for each county from four simple population scenarios – neipop25grp1, neipop50grp1, neipop75grp1, neipop100grp1. Hence, in each of the plots, there are 44 data points. The regression line was passed through point (0,0) by setting the y-intercept parameter to zero. Figure 19 through Figure 21 show regression plots for source types 21, 31, and 32, respectively. For all the plots, the TTI research team observed that regression parameters were statistically significant and linear to the population changes. For source type 21, PM10, CO2eq, and CO sensitivities were higher than other pollutants. For source type 31, the sensitivities of all the pollutants were almost the same as source type 21. With the changes in source type 32 populations, the sensitivities of pollutants were much higher than those on changes in source types 21 and 31 populations.

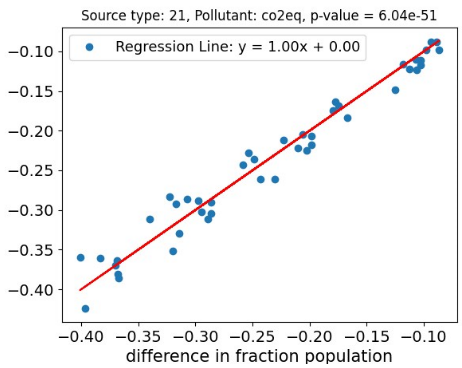
The data of various counties, including urban and rural areas, was plotted in the same regression model. The statistical significance showed that the same model can be applied to counties with different characteristics. Hence, this model can also be used as a predictive model, given that the total VMT for an HPMS source type group is the same in two scenarios.



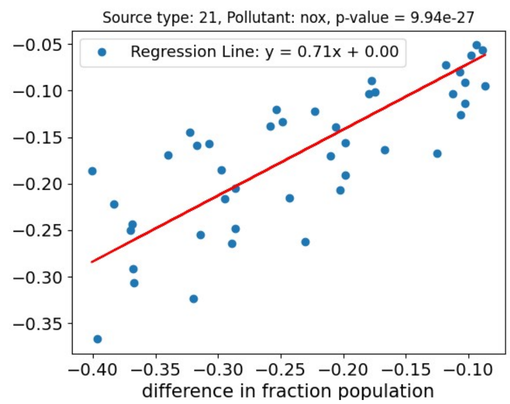
(a) PM10



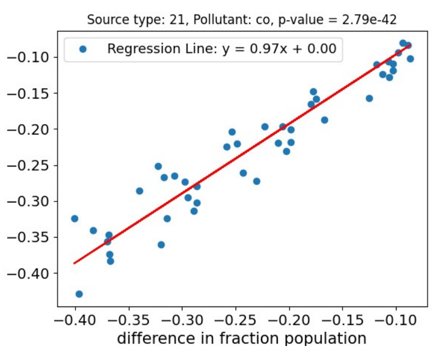
(b) PM2.5



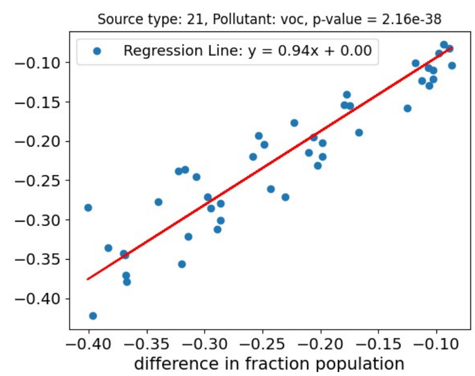
(c) CO2eq



(d) NOx



(e) CO

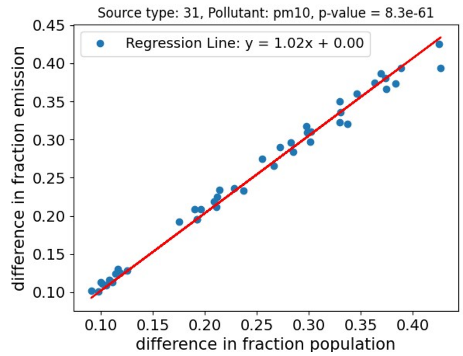


(f) VOC

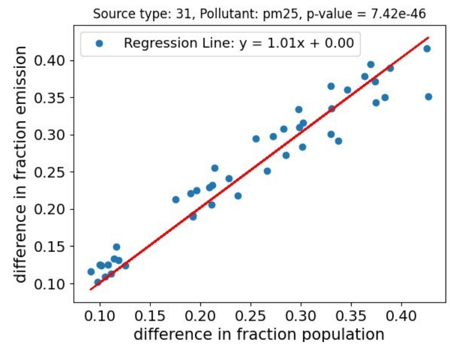
X = [(Population for Source type 21/ (Population for Source type 21 + 31 + 32) in NEI) – (Population for Source type 21/ (Population for Source type 21 + 31 + 32) in AERR)]

Y_pollutant = [(Emissions of pollutant due to source type 21/ Emissions of pollutant due to source types (21 + 31 + 32) in NEI) – (Emissions of pollutant due to source type 21/ Emissions of pollutant due to (21 + 31 + 32) in AERR]

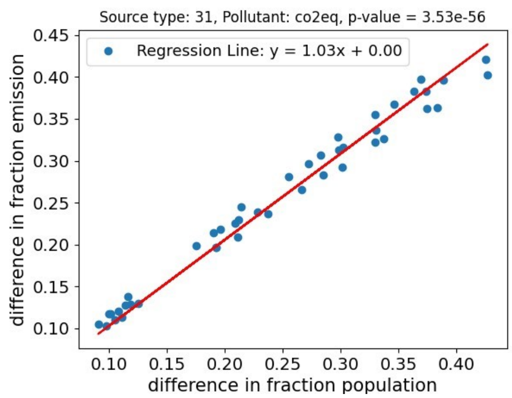
Figure 19. Regression Analysis on Sensitivities of Emission Difference by Population Difference in Simple Population Scenario for Source Type 21



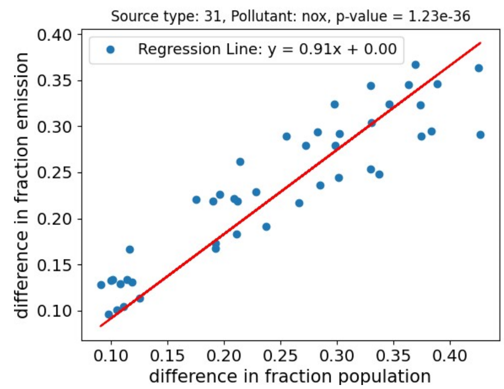
(a) PM10



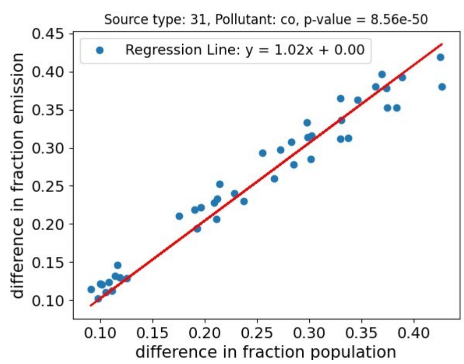
(b) PM2.5



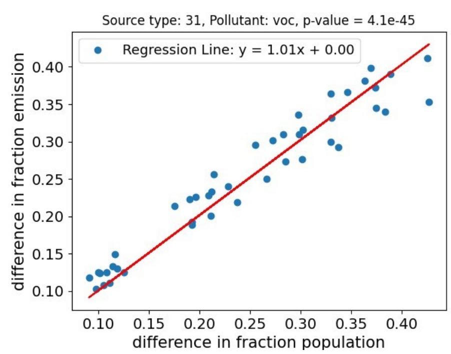
(c) CO2eq



(d) NOx



(e) CO

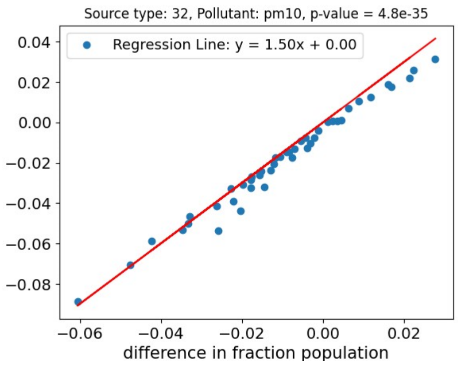


(f) VOC

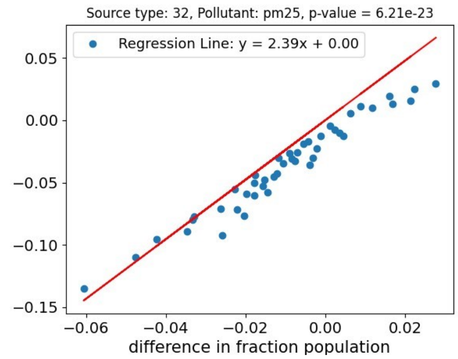
X = [(Population for Source type 31/ (Population for Source type 21 + 31 + 32) in NEI) – (Population for Source type 31/ (Population for Source type 21 + 31 + 32) in AERR)]

Y_pollutant = [(Emissions of pollutant due to source type 31/ Emissions of pollutant due to source types (21 + 31 + 32) in NEI) – (Emissions of pollutant due to source type 31/ Emissions of pollutant due to (21 + 31 + 32) in AERR]

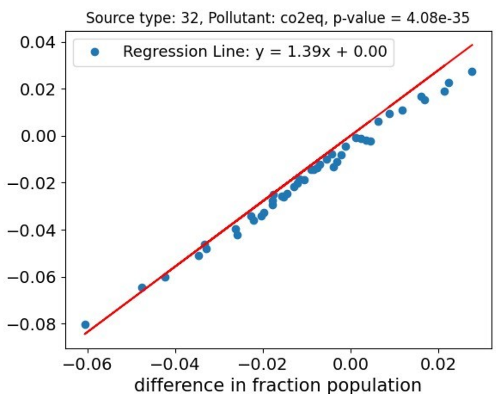
Figure 20 Regression Analysis on Sensitivities of Emission Difference by Population Difference in Simple Population Scenario for Source Type 31



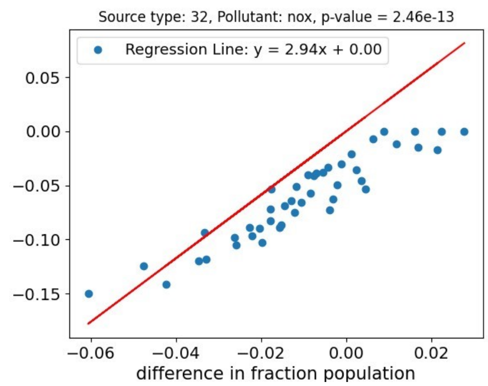
(a) PM10



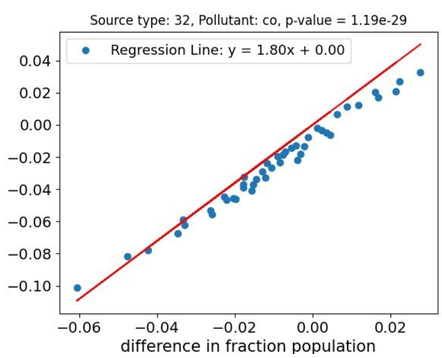
(b) PM2.5



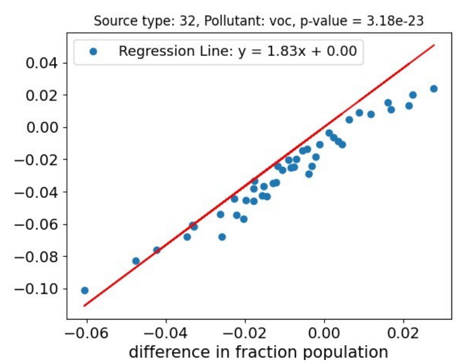
(c) CO2eq



(d) NOx



(e) CO



(f) VOC

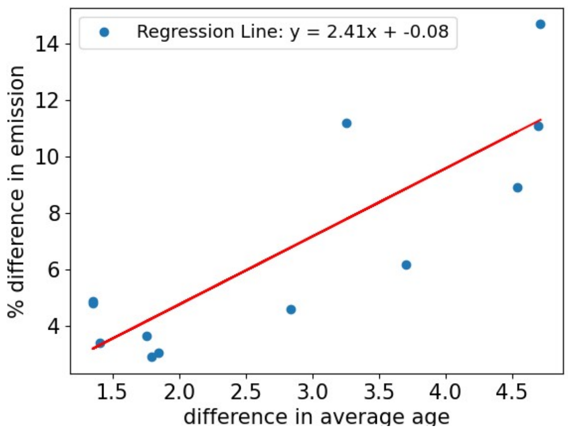
X = [(Population for Source type 32/ (Population for Source type 21 + 31 + 32) in NEI) – (Population for Source type 32/ (Population for Source type 21 + 31 + 32) in AERR)]

Y_pollutant = [(Emissions of pollutant due to source type 32/ Emissions of pollutant due to source types (21 + 31 + 32) in NEI) – (Emissions of pollutant due to source type 32/ Emissions of pollutant due to (21 + 31 + 32) in AERR]

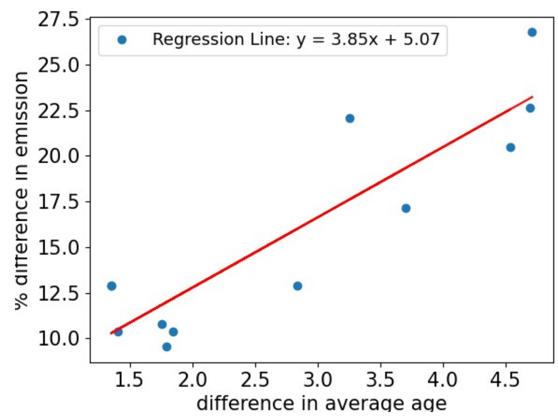
Figure 21. Regression Analysis on Sensitivities of Emission Difference by Population Difference in Simple Population Scenario for Source Type 32

4.2.1.2Age Distribution in Group 4 (CT) Scenarios

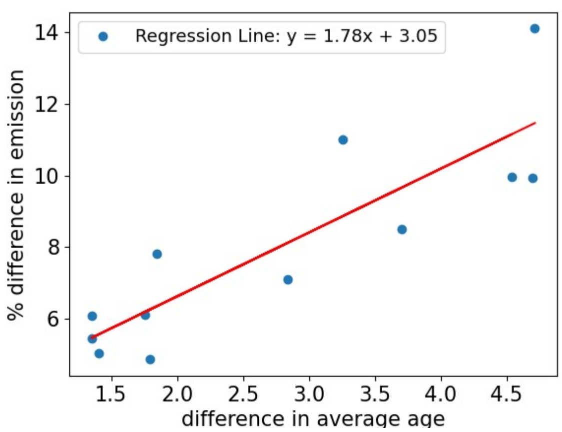
Age distribution, or average age, was found to be a sensitive variable for Group 4 Scenarios. Figure 22 shows the sensitivity of pollutants with respect to the average age of CTs. One plot has twelve data points corresponding to the *neiagedistgrp4* scenario. The slope of the regression line represents the expected increase in the percent difference of pollutants if the expected difference in average age increased by one year. $PM_{2.5}$ has the highest sensitivity. If the average age of CTs is older than one year, the expected increase in percent difference in PM2.5 emission is about 3.85 percent.



(a) PM10



(b) PM2.5



(c) NOx

the difference in average age = (average age for source types 61 and 62 in Scenario) - (average age for source types 61 and 62 in AERR)

Figure 22. Regression Analysis on Sensitivities of Emission Difference by Population Difference in Simple Population Scenario for Combination Trucks

4.3 Chapter Summary

This chapter discussed the results of the sensitivity with respect to scenarios and the input parameters. As a part of the base scenarios analysis, several input parameters, including vehicle population, vehicle speed distribution, age distribution, and road type distribution, were initially tested to assess the level of emissions differences in the AERR 2020 resulting from swapping, in particular, NEI input values (individually for fuels, meteorology, hotelling activity, and allocation of VMT by vehicle category; and for fuels and meteorology together) while keeping other AERR CDB inputs constant. In the simple and interaction scenarios, one or more inputs were changed in AERR 2020 CDBs, keeping the other inputs the same. The TTI research team categorized MOVES source types into four groups: source types 21, 31, and 32 as Group 1; source types 41, 42, and 43 as Group 2; source types 51, 52, 53, and 54 as Group 3; and source types 61 and 62 as group 4. The emission difference between AERR 2020 and the considered scenario was utilized to determine the sensitivity of that input. A key feature of the assessment plan, which is different from other related studies, is that inputs used in the sensitivity scenarios were directly derived from real data sources (NEI and AERR); hence, the sensitivity analysis is bounded within the range of realistic values. The team performed a total of 1056 simulation runs on a selection of twelve representative counties and various scenarios. Key findings from the study are as follows.

- For considered pollutants, NEI showed higher emissions than AERR. When MOVES default inputs were used, emissions were considerably higher.
- Emission estimates using MOVES defaults were neither close to NEI nor AERR. Hence, using MOVES defaults is not recommended. It is crucial to develop county-specific inputs instead of using defaults representing national averages.
- For all the source type groups, fuel inputs, meteorology inputs, and road distribution have negligible effects on pollutants.
- Higher emissions were observed when VMTs were distributed according to MOVES source types as opposed to HPMS source types while keeping the total HPMS VMT the same. Hence, it was important to focus on how the VMT was distributed, inputted, and entered for various source types.

- Emissions were more sensitive to the parameters of LDVs and CTs as compared to buses and source utility types single unit trucks; hence, more attention should be given to improving inputs for these source types.
- The population of LDVs was found to have a substantial (>20%) impact on emissions and hence was concluded to be a highly sensitivity parameter. The age and speed distribution, when only changed, had a low (0-5%) to medium (5-20%) impact on emissions. However, when age, population, and speed distribution were changed simultaneously, their interaction effect had a substantial impact (>20%) on emissions.
- For combination trucks, the age distribution was found to be the most important input with a medium (5-20%) impact on emissions.

5 DISCUSSIONS

This chapter dives deeper into results and how emissions processes are mapped to activity. The TTI research team grouped the emission processes into four rate categories. Then, the TTI research team answer why emissions of NEI are higher, though the total HPMS VMT of NEI and AERR are the same.

5.1 Rate Category Analysis

Emissions activity was divided into four categories, as shown in Table 17. The rate categories are convenient for understanding the causes of emissions for categories listed under the corresponding process labels. For each scenario, the difference in emissions was analyzed for each rate category.

Table 17. Rate Categories

Rate Categories	Processes
VMT-based (Road type !=1)	Running Exhaust Crankcase Running Exhaust Breakwear Tirewear Evap Permeation* Evap Fuel Vapor Venting* Evap Fuel Leaks*
Hotelling	Crankcase Extended Idle Exhaust Extended Idle Exhaust Auxiliary Power Unit Exhaust
Starts	Start Exhaust Crankcase Start Exhaust
Off-network (Road type ==1)	Running Exhaust Crankcase Running Exhaust Evap Permeation Evap Fuel Vapor Venting Evap Fuel Leaks

* Processes happening during all the activity

Table 18 shows the contribution of the VMT-based rate category for each pollutant. The TTI research team observe that except VOC, the VMT-based rate category is a major contributor for all other pollutants, causing ~70-90% of emissions.

Table 18. Percentage Contribution of VMT-Based Processes in Emission of Each Pollutant for *neipop100* Scenario

county	CO	NO_x	CO_{2eq}	VOC	PM_{10}	$PM_{2.5}$
48007	81.33	72.63	90.04	27.88	95.37	83.61
48013	84.87	81.71	91.66	32.89	94.21	83.92
48039	77.43	74.05	90.48	25.33	94.67	82.48
48041	80.47	76.47	90.16	30.63	94.78	82.83
48081	86.69	86.54	93.22	41.10	94.80	85.78
48141	79.10	79.60	90.11	30.77	93.81	80.07
48143	83.10	80.96	92.12	33.48	94.16	82.85
48201	80.56	76.73	90.90	30.66	96.30	86.28
48243	90.30	90.85	93.76	50.31	95.31	88.46
48439	75.01	74.47	89.63	26.39	94.25	80.55
48453	77.79	72.33	90.02	24.47	93.88	79.81

Knowing the main contributor of emissions is VMT-based activity, the TTI research team looked into emission differences. Table 19 shows an example of rate categories explaining emission differences. For each pollutant, the emission difference in the *aerr* scenario and *neipop100* scenario is calculated. The fraction of emission difference explained by each of the rate categories is reported. The number in one column sums up to 1. This exercise is done for all the counties. For CO, 0.687 in the VMT-based rate category shows that 68.7% of the emission difference between the aerr and neipop100 scenario is caused due to VMT-based activity.

Table 19. Emission difference for Aransas County (48007) for neipop100grp1 scenario

Rate Category	CO	NOx	CO_{2eq}	VOC	PM_{10}	$PM_{2.5}$
VMT Based	0.687	0.714	0.868	0.274	0.792	0.720
Hotelling	0.000	0.000	0.000	0.000	0.000	0.000
Starts	0.298	0.246	0.074	0.240	0.241	0.237
Off-network	0.016	0.040	0.058	0.486	0.058	0.060

The following conclusions can be drawn from this analysis:

- For CO, NOx, PM_{10}, and $PM_{2.5}$, 70-80% of the emission difference is caused by VMT-based activity. The remaining 20-30% difference is explained by starts. The difference in off-network activity is negligible.
- For CO_{2eq}, ~ 90% of the emission difference is explained by VMT-based activity and ~10% by starts.
- For VOCs, ~50% difference is explained by off-network activity. ~20% emission difference is explained by starts, and ~30% is explained by running.

5.2 Why Emissions are Higher Though the HPMS VMT is the Same

VMT distribution can cause differences in emissions, though the total HPMS VMT is the same. This was observed in the *neivmt* scenario where up to 15 % higher emissions were observed when VMT was input according to MOVES source types as compared to HPMS vehicle types.

To look deeper into this finding, we aggregated total VMT for source types 21, 31, and 32 from the *moves activity output* table. Table 20 shows VMT distributed to each of the source types. We can see that in her and neipop100 scenarios, the total VMT allocated to group 1 is the same. However, there is a considerable difference in VMT distributions. In the neipop100 scenario, more VMTs are distributed towards source type 31. This turned out to be a deciding factor for the higher emissions of NEI as compared to AERR.

Table 21 shows VMT distributed for source types 21, 31, and 32 in the neipop100age scenario. If the numbers are compared with

Table 21, a higher VMT is allocated to source type 31 in the neipop100age scenario than in the neipop100 scenario. This suggests that age distribution also affects the allocation of VMT. A more extensive explanation is given in the next section.

Table 20. VMT Distribution Comparison for LDVs in neipop100 Scenario

county	vmt_21 neipop100	vmt_31 neipop100	vmt_32 neipop100	vmt_21 aerr	vmt_31 aerr	vmt_32 aerr	Total VMT aerr	Total VMT neipop100
48007	52,885,832	139,250,581	5,507,086	128,467,410	55,132,895	14,043,213	197,643,517	197,643,499
48013	173,205,487	374,843,082	33,787,752	352,004,292	183,180,117	46,651,901	581,836,309	581,836,321
48039	895,810,526	1,643,803,849	103,204,157	1,870,511,924	615,532,931	156,773,757	2,642,818,613	2,642,818,532
48041	496,261,983	803,594,453	69,363,154	1,002,158,243	292,548,021	74,513,673	1,369,219,937	1,369,219,590
48081	11,456,203	43,333,279	2,031,547	29,776,208	21,555,091	5,489,739	56,821,038	56,821,030
48141	2,796,599,073	3,112,908,754	186,468,055	4,710,737,693	1,104,051,355	281,186,579	6,095,975,627	6,095,975,882
48143	100,878,188	281,627,954	19,402,830	237,359,055	131,143,793	33,406,138	401,908,986	401,908,972
48201	14,711,296,919	19,865,041,043	1,831,448,817	28,144,906,361	6,585,564,465	1,677,316,859	36,407,787,685	36,407,786,779
48243	9,796,437	38,454,294	2,913,204	30,337,165	16,599,194	4,227,585	51,163,944	51,163,935
48439	6,297,957,346	7,592,995,567	941,004,408	11,758,797,196	2,449,261,790	623,898,323	14,831,957,309	14,831,957,321
48453	3,582,162,182	4,182,173,132	514,182,880	6,876,271,786	1,117,566,479	284,680,210	8,278,518,474	8,278,518,195

Table 21. VMT Distribution for LDVs in neipop100age Scenario

county	vmt_21 neipop100age	vmt_31 neipop100age	vmt_32 neipop100age	vmt_21 aerr	vmt_31 aerr	vmt_32 aerr	Total VMT aerr	Total VMT neipop100age
48007	49,703,592	142,103,933	5,835,988	128,467,410	55,132,895	14,043,213	197,643,517	197,643,512
48013	163,346,137	379,059,685	39,430,550	352,004,292	183,180,117	46,651,901	581,836,309	581,836,371
48039	847,822,491	1,685,405,669	109,590,468	1,870,511,924	615,532,931	156,773,757	2,642,818,613	2,642,818,629
48041	462,796,405	832,961,105	73,462,055	1,002,158,243	292,548,021	74,513,673	1,369,219,937	1,369,219,565
48081	10,693,013	44,228,948	1,899,074	29,776,208	21,555,091	5,489,739	56,821,038	56,821,036
48141	2,700,871,429	3,194,709,754	200,394,074	4,710,737,693	1,104,051,355	281,186,579	6,095,975,627	6,095,975,258
48143	94,807,823	285,864,761	21,236,377	237,359,055	131,143,793	33,406,138	401,908,986	401,908,962
48201	13,962,875,603	20,371,859,619	2,073,056,506	28,144,906,361	6,585,564,465	1,677,316,859	36,407,787,685	36,407,791,729
48243	8,916,198	39,320,303	2,927,439	30,337,165	16,599,194	4,227,585	51,163,944	51,163,940
48439	5,955,023,358	7,768,970,396	1,107,963,712	11,758,797,196	2,449,261,790	623,898,323	14,831,957,309	14,831,957,467
48453	3,372,487,141	4,326,770,965	579,260,779	6,876,271,786	1,117,566,479	284,680,210	8,278,518,474	8,278,518,885

5.2.1 Relative Milage Accumulation Rate

For emission calculations, MOVES needs to estimate the miles traveled by each age and source type. MOVES uses a relative mileage accumulation rate (RMAR) in combination with source type populations and age distributions to distribute the total annual miles driven by each HPMS vehicle type to each source type and age group. Using this approach, the vehicle population and the total VMT can vary from calendar year to calendar year, but the proportional travel by an individual vehicle of each age will not vary. The RMAR is determined from the mileage accumulation rate (MAR) within each HPMS vehicle classification such that the annual mileage accumulation for a single vehicle of each age of a source type is relative to the mileage accumulation of all of the source types and ages within the HPMS vehicle classification. For example, passenger cars, passenger trucks, and light commercial trucks are all within the same HPMS vehicle classification (Light-duty vehicles, HPMSVTypeID 25). New (age 0) passenger trucks and light commercial trucks are defined to have an RMAR of one (1.0), and new passenger cars have an RMAR of 0.885. This means that when MOVES allocates the VMT assigned to the light-duty vehicle HPMS class to passenger cars, passenger trucks, and light commercial trucks, a passenger car of age 0 will be assigned only 88.5 percent of the annual VMT assigned to a passenger truck or light commercial truck of age 0. For the RMAR values used in MOVES3, the reader is referred to Figure 6-2 in Population and Activity of Onroad Vehicles (USEPA 2020). We showed here that population and age distribution have a considerable impact on emissions. Both inputs can account for differences in emissions, though the total VMT is the same.

5.3 Comparison with Other Studies on Sensitivity

We compared our results with other studies on the sensitivity of MOVES. The other important studies include NCHRP 25-38 (Porter et al., 2015) and FHWA 2023 (Chupp et al., 2023). The main difference between the above studies and this study is that we anchored scenario development with available datasets. For example, in the NCHRP 25-38 study, 20% of the passenger cars were shifted to passenger trucks in the source type population scenario. In this study, we start from the source type population in AERR and gradually move toward the NEI population. This makes our study more realistic. With the above scenario, the NCHRP 25-38 report found that the emission difference was 1% for NOx and PM and 2% for VOCs. However, we found that the source type population of LDVs has>5% impact on all the considered pollutants. This NCHRP 25-38 methodology of scenario development is more prevalent for inputs like temperature.

For example, in the VOLPE study (Noel et al., 2012), temperatures varied from -40 degrees to 120 degrees compared with the base case of 60 degrees. They concluded that emissions are highly sensitive to temperatures. However, temperature data is available with reasonable accuracy. Hence, considering temperature variation within two different datasets is more reasonable. We concluded that the sensitivity of meteorology inputs like temperature and relative humidity sourced from different datasets for the same areas/seasons is negligible.

FHWA 2022 report focuses on hotelling, off-network idle, average speed distribution, and age distribution. They concluded that total inventories show significantly less sensitivity to hotelling and off-network idle activity compared to average speed and age distribution. This finding aligns with our findings. FHWA mentions age and speed but does not mention population distribution. We found population distribution to have a significant impact on the sensitivity. Koupal et al. (2014) also found population distribution, age, and speed to be significant factors. We found that the emission changes for age were not just due to emission rates but also to activity changes. How inputs are provided (VMT by HPMS vehicle type or MOVES source use type) matters.

6 ACTIVITY DATA RANKING

Based on the results of this study, population, age distribution, source type VMT distribution, and average speed distribution impacted emission estimates. HPMS vehicle-type VMT is directly from TxDOT data sources and does not change. The MOVES source-type VMT distribution within the HPMS categories can differ and impact emissions.

This chapter presents a ranking system (task 5.1) and the resulting rank matrix (task 5.2) for data sources to improve sensitive activity inputs and emission estimates. For the ranking system, we used a decision matrix (American Society for Quality, 2023) with a weightage 1, 2, or 3 (low, medium, and high) for different attributes. The weights determine the importance of each attribute with respect to each other and the emission inventory development process. In addition to the weights, each attribute is assigned a score. The final score for each dataset is a weighted sum of the individual attribute score. The datasets (or types of datasets) with higher scores can be purchased or explored first. Following is a list of different types of datasets considered:

- Registration and survey data
- Raw telematic data
- Processed telematic data
- Modeled data
- Default Database: MOVES Default Database/ Previous NEIs

6.1 REGISTRATION AND SURVEY DATA

Vehicle registration data is used at the county and district level to:

- Determine the proportions of vehicles in each modeled Source Use Type category, which in turn is used to derive fleet-relevant emission rates.
- Estimate actual SUT 61 short-haul VMT.
- Estimate total vehicle population to determine off-network activity such as source hours parked, vehicle starts, and heavy-duty vehicle idling.

This chapter considered the TxDMV data, IHS Markit–Polk registration data (S&P Global, 2023; S&P Global Mobility, 2023), and 2021 Vehicle Inventory and Use Survey (VIUS) data (U.S. Department of Transportation, Bureau of Transportation Statistics & U.S. Department of Commerce, U.S. Census Bureau, 2023). TxDMV data has been in use in Texas for over 20 years. It is reliable for regional emission estimates. IHS Markit–Polk dataset can provide more information, such as vehicle make, model, GVWR, fuel type, vehicle type, and registration type. 2021 VIUS data can provide information for differentiating activity of some source use types, such as light commercial trucks and personal trucks. Following is an overview of these three datasets.

Texas Department of Motor Vehicle (TxDMV)

Historically, Texas emission inventories use TxDMV data on the number, age distribution, and category of vehicles registered in Texas. TTI and TxDOT developed modified data requests to exclude certain registration classes (notably trailers) and unregistered vehicles. TTI requests vehicle registration data from the TxDMV through the Transportation Planning and Programming (TPP) Division, TxDOT, every two or three years. In March 2022, TxDOT received a new dataset from TxDMV for year-end 2021.

IHS Markit – Polk

IHS Markit – Polk registration data provides comprehensive information on vehicles, such as vehicle make, model, GVWR, fuel type, vehicle type, and registration type. EPA uses this data to develop defaults and NEIs.

2021 VIUS

VIUS 2002 had been used by MOVES for developing the vehicle population and activity in the past. Note it is available at the state level and not by counties. "The 2021 VIUS was conducted for the Bureau of Transportation Statistics (BTS) by the U.S. Census Bureau with support from the FHWA and the Department of Energy (DOE) to understand better the characteristics and use of vehicles on our nation's roads." "The 2021 VIUS is a survey of 150,000 vehicle owners of class 1 through 8 trucks, which includes vehicle body types, such as pickups, SUVs, minivans, light vans, straight trucks, and truck tractors. Depending on the size of the selected vehicle, the vehicle owner received either the Heavy Vehicle Questionnaire or the Light Vehicle Questionnaire. The survey was conducted from February through October 2022. "(U.S. Department of Transportation, Bureau of Transportation Statistics & U.S. Department of Commerce, U.S. Census Bureau, 2023)

6.2 Raw telematic data

This data type provides raw trajectory points for trips at different frequencies and accuracy depending on the underlying technology for collecting this dataset. It can be used for developing estimates, including spatial and temporal VMT distribution, average speed distribution, and start distribution. See (Koupal et al., 2022; Singh et al., 2022) for more details. This study considers Wejo data (Wejo Limited, 2023). Similar data is available through other vendors, such as Otonomo, INRIX, Geotab, Verizon Connect, and Moonshadow Mobile.

6.3 Processed telematic data

Processed telematics data provides the information in the raw telematics data to the user at a lower resolution than the raw dataset. For instance, processed telematics data might provide the speeds and volumes at the link level using raw telematics data. This dataset requires less processing than the raw data, as the processing is conducted by the data vendors instead of the user. This study has considered Streetlight Data (StreetLight Data, Inc., 2023) to represent processed telematics data. Other data vendors include INRIX and Moonshadow Mobile.

Streetlight can provide link-level data from probe vehicles for OpenStreetMap links. It can provide average speeds and VMT distribution. The underlying data comprises CV, GPS, and other probe vehicle data sources.

6.4 Modeled data

"*Modeled data are modeled trip trace data that is enriched with other non-passive data sources such as demographic data, land use data, credit card data, etc. It is then processed to develop synthetic travel diaries to subsequently model a region's daily travel patterns analogous to an activity-based model and assigned to the regional transportation network*." (Singh et al., 2022).

The TTI research team considered Replica (Replica, 2023) data for this study. Replica can provide telematic and activity-based travel demand model data for OpenStreetMap links. This can be used to develop average speed distribution and VMT distribution. In addition, we have considered the travel demand model data from metropolitan planning organizations currently used for developing VMT distribution by road type and average speed distribution.

6.5 Default Database: MOVES Default Database/ Previous NEIs

MOVES provides population and age distribution from 2020 IHS Markit -Polk and average speed distribution data based on 2020 Streetlight Data.

6.6 Ranking Weightage

Following is a list of attributes and their proposed weightage:

- **sensitivity impact:** this parameter measures emission sensitivity to the data sources. The values of this parameter are based on the assessment conducted in this study. Thus, this parameter is an objective measure. The TTI research team propose giving this parameter a weightage of **3,** given that it is measured from experiments conducted during this study.
- **price:** This parameter measures the cost of the data source. We have collected this information by reaching out to various vendors. The quotes we have gotten are rough estimates of the actual cost. We need to buy the datasets for a study area for more precise estimates. Given that this parameter is based on price quotes and is extremely important when deciding whether to use a data source, We are using a weight of **3**.
- **uncertainty (error rate)/ reliability (accuracy):** this parameter measures the ability of data to assess the ground truth. The entire motivation of this study is to improve sensitive activity estimates. Thus, the parameter is pivotal when deciding to use an alternative dataset. Its weight is proposed to be **2**. The weight is lower than the previous two parameters, as price is often the limiting factor when considering alternatives. Moreover, sensitivity is important, as improving the activity estimate for a parameter to which emission estimates are less sensitive will not be judicious.
- **frequency of availability:** the Texas inventory development framework was developed in the 1990s by TTI and TCEQ. It is a complex framework with multiple data sources and processing methodologies. Significant work and time are needed to update the methods and tools to incorporate new data sources. It is thus important that any new data source that replaces existing data sources would be available for future inventory development work. Note that although vendors change, the underlying datasets are the same so that underlying data might be

available through other sources. This parameter is given a weightage of **2**; given that this parameter is important, price and sensitivity are the highest priority.

- **Processing level of effort (LOE): This parameter tries to capture the time and effort required** to incorporate the new data sources in the current emission inventory development methodology. We are proposing a rating of **1** for this parameter, as the processing effort would be substantially only for the first iteration. The subsequent work can leverage previous work to update.

The following table (Table 22) shows the potential data sources and their scores for developing MOVES activity inputs identified during task 4 for the above attributes. Improving these inputs will lead to better emission estimates. The study team developed a ranking matrix for these datasets based on the attributes in the table below. Each dimension is scored between 1 and 3, where 3 is the best-suited value for emission inventory development work. In the table, three is represented by a value "low"; 1 is represented by a value "high."

The following Table 23 presents the rank scores based on the above-proposed ranking system. As mentioned before, the score is a weighted sum of the individual attribute scores weighted by the attributes' weightage. For instance, for TxDMV data, the score is 31 (3*3 (*Sensitivity*)+ 3*3 (*Cost*) + 2*2 (*Uncertainty*) +2*3 (*Availability Issues*) + 1*3 (*Processing LOE*)). The weightage of each attribute has been discussed above. The attribute scores are based on the table above, with a higher value (3) implying a better score on that attribute for a dataset. A higher weightage magnifies the score. That is why the weighted score for TxDMV data for Sensitivity is 9.

Table 22. Potential Data Sources Examples and Target MOVES Activity Inputs

Type	Example Data Source	MOVES Activity Inputs	Cost	Uncertainty (error rate)	Availability Issues	Processing LOE
Registration and Survey data	TxDMV	• Population • Age distribution	None. TTI has access to this data through the TxDOT.	Medium. The resolution is low, as it only provides county-level aggregate information on GVWR (and fuel type for heavy-duty vehicles).	Low. It has been used in Texas for over 20 years.	Low. There are well-established procedures for converting this data into MOVES inputs.
	IHS Markit – Polk (S&P Global, 2023; S&P Global Mobility, 2023)	• Population • Age distribution	The initial quote is $80,000.	Low. Higher resolution data compared to the TxDMV dataset accessible to TTI and TCEQ. The data reliability reduces for older model years before 2012.	Low. The first vehicle registration report from this provider dates back to 1920.	Medium. The dataset likely has a structure similar to the data received from TxDMV but with additional fields.
	2021 VIUS (U.S. Department of Transportation, Bureau of Transportation Statistics & U.S. Department of Commerce, U.S. Census Bureau, 2023)	• Population • Age Distribution • Source Use Type VMT	None.	Low. Uses a well-documented and rigorous methodology.	High. The last survey before this one was published in 2002.	Low. Small tables that can be directly used for activity estimates.
Raw Telematic data	Wejo (Wejo Limited, 2023)	• Average speed distribution • Road-type VMT distribution	High. TTI or TCEQ has access to this data through the TxDOT TPP division. So, access cost is low, but the data processing needs cloud computing, which can be expensive.	Low. It is the raw connected vehicle data, thus allowing the analyst to develop appropriate processing methodologies.	High. Wejo is going through bankruptcy currently (September 2023)	High. Need cloud computing infrastructure to process this dataset.
Processed Telematic data	Streetlight Data (StreetLight Data, Inc, 2023)	• Average speed distribution	$700,846 for the entire Texas. $284,587 for three	Low. It is based on observed roadway conditions.	Low. Streetlight Data has been in	Medium. Streetlight Data map matches and processes the

Type	Example Data Source	MOVES Activity Inputs	Cost	Uncertainty (error rate)	Availability Issues	Processing LOE
		• Road-type VMT distribution	major metro areas (Houston, San Antonio, and Dallas regions). $89,066 for Harris County.		business since 2011.	raw data on their end. It needs programming expertise to query link-level data.
Modeled data	Replica (Replica, 2023)	• Average speed distribution • Road-type VMT distribution	$250,000. Some of their data products are available through TxDOT. The dataset needed for the emission inventory application is not included.	Unknown. It is modeled data from activity-based models. A case study is needed to determine this dataset's reliability.	Unknown. It has been in business since 2019.	Medium. Needs the expertise to query link-level data.
	TDM	• Average speed distribution • Road-type VMT distribution	None. The TDM is available through TxDOT.	Low. They are routinely updated by the Metropolitan Planning Organizations (MPOs)	Low. It has been used in Texas for over 20 years.	Low. There are well-established procedures for converting this data into MOVES inputs.
MOVES Defaults	MOVES Default Database/ Previous NEIs (US EPA, 2023; USEPA, 2020)	• Population • Age distribution • Average speed distribution • Road-type VMT distribution	None. Older datasets for 2020 from IHS Markit – Polk and Streetlight Data are available.	Medium. Not the most recent datasets.	Medium.	Medium

Note that *the weightage of the different parameters and the values for these parameters for different datasets are subjective. Unless a pilot study is conducted, it would be hard to objectively evaluate the potential of different datasets and tradeoffs for improving emission estimates. Thus, the rank score should be considered a first step toward exploring alternative datasets. Also, the ranking is only for the potential of the various datasets for improving emission estimates.*

Table 23. Rank Matrix

Example Data Source	Sensitivity	Cost	Uncertainty	Availability Issues	Processing LOE	Rank Score
Weightage	**3**	**3**	**2**	**2**	**1**	**N/A**
TxDMV	3	3	2	3	3	*31*
IHS Markit – Polk	3	2	3	3	2	*29*
VIUS 2023	3	3	3	1	3	*29*
Wejo	1	1	3	1	1	*15*
Streetlight Data	1	1	3	3	2	*20*
Replica	1	1	2	2	2	*16*
TDM	1	3	3	3	3	*27*
MOVES Default Database/ Previous NEIs	3	3	2	2	2	*28*

Note that *the weightage of the different parameters and the values for these parameters for different datasets are subjective. Unless a pilot study is conducted, it would be hard to objectively evaluate the potential of different datasets and tradeoffs for improving emission estimates. Thus, the rank score should be considered a first step toward exploring alternative datasets. Also, the ranking is only for the potential of the various datasets for improving emission estimates.*

Based on this matrix, 2021 VIUS should be evaluated to develop local MOVES activity estimates. IHS Markit – Polk dataset should be considered for initial purchasing. In terms of the individual activity inputs identified in task 4, the takeaways are as follows:

- for population and age input tables, the TxDMV dataset is good. Consider working with TxDOT and TxDMV to get a higher resolution for this dataset. TCEQ can also consider purchasing IHS Markit–Polk data and analyzing the 2021 VIUS dataset.
- for the MOVES source, use the type VMT distribution input table; TCEQ can consider using the 2021 VIUS dataset.
- for the average speed distribution input table, TDM, MOVES defaults, and NEI tables can be used. Alternatively, TCEQ can consider using the Wejo data available through TxDOT.

7 CONCLUSION

The main objective of this study was to analyze the sensitivity of MOVES input variables. An extensive literature review of AERR and NEI studies was done, and it was pointed out that there are several differences in the key input CDBs of these studies, which led to differences in the final emission estimates. An assessment plan was developed to analyze the sensitivity of emissions with respect to various inputs. The TTI Research team selected twelve representative counties, and 1056 MOVES simulation scenarios were developed. These scenarios were divided into three categories: base, simple, and interaction. Base scenarios comprised AERR 2020, NEI 2020, MOVES defaults, and several input parameter test benchmark scenarios. In simple scenarios, one of the important input CDB tables from AERR 2020 CDB was replaced by the NEI 2020 CDB table. Input population, age distribution, and speed distribution tables were tested for sensitivity analysis. In the interaction scenarios, multiple inputs were changed from input CBDs. The source types were classified into the source type groups of LDVs, STs, Buses, and CTs. Emissions sensitivity was tested for each of the parameters and groups.

For considered pollutants, NEI showed higher emissions than AERR. When MOVES default inputs were used, emissions were considerably higher. Emission estimates using MOVES defaults were neither close to NEI nor AERR. Hence, using MOVES defaults is not recommended. It is crucial to develop county-specific inputs instead of using defaults representing national averages.

When meteorology and fuels data from NEI were used, the difference in emissions was less, which showed that for Texas, emissions were not very sensitive for these inputs when they varied within a realistic range. For all the source type groups, we observed that fuel inputs, meteorology inputs, and road distribution have negligible effects on pollutants. Hence, less attention can be given to improving these inputs.

Higher emissions were observed when VMT was distributed according to MOVES source types while keeping the total HPMS VMT the same. Hence, it is important to focus on how the VMT is distributed for various source types.

The major difference in emissions was caused by the LDVs source type group. For LDVs, the emissions gradually increased as the population shifted from AERR to NEI. NEI age distribution has a negative effect on NOx and PM2.5. NEI speed distribution has a positive effect on CO. NEI speed distribution has a negative effect on PM_{10}. For buses and STs, the percent difference is close to zero for all the pollutants. This suggests that more focus should be given to improving the estimates for LDVs instead of Buses and STs. NOx, $PM_{2.5}$, and PM_{10} emissions were sensitive for CTs.

Rate category analysis showed that the VMT-based rate category is a major contributor to each pollutant. We observe that except VOC, the VMT-based rate category is a major contributor for all other pollutants, causing ~70-90% of emissions. LDV population difference in NEI and AERR is a major contributor to emission differences; hence, efforts should be made to get vehicle populations from other data sources more accurately.

Based on the findings, the TTI Research team developed a ranking system for potential data sources that could be used and explored to improve current emission inventories. For population and age input tables, the TxDMV dataset is good. TCEQ can consider purchasing IHS Markit-Polk data and analyzing the 2021 VIUS dataset. TDM, MOVES defaults, and NEI tables can be used for the average speed distribution input table. Alternatively, TCEQ can consider buying telematics data.

8 REFERENCES

American Society for Quality. (2023). What is a Decision Matrix? Pugh, Problem, or Selection Grid | ASQ. https://asq.org/quality-resources/decision-matrix

Chicago Metropolitan Agency for Planning. 2018. *On to 2050 Travel Demand Model Documentation*. Chicago Metropolitan Agency for Planning.

Chupp, W., Zubrow, A., Raymond, A., Liu, K., Yuen, A., Eilbert, A., Chang, S., Mittelman, A., & John A. Volpe National Transportation Systems Center (U.S.). (2023). MOVES3 Sensitivity Analysis (FHWA-HEP-23-020). https://rosap.ntl.bts.gov/view/dot/67328

Clarksville Urbanized Area MPO. 2019. *Air Quality Conformity Analysis (Kentucky Portion)*. Clarksville Urbanized Area MPO.

Koupal, J. W., DenBleyker, A., Manne, G., Batista, M. H., Schmitt, T., Kishan, S., & Richardson, C. (2022). Capabilities and Limitations of Telematics for Vehicle Emissions Inventories (CRC E-131; p. 55). Coordinating Research Council, Inc.

Koupal, J., DeFries, T., Palacios, C., Fincher, S., & Preusse, D. (2014). Motor Vehicle Emissions Simulator Input Data: Evaluation and Sensitivity Analysis of Data Submitted for 2011 National Emissions Inventory. Transportation Research Record: Journal of the Transportation Research Board, 2427(1), 63–72. https://doi.org/10.3141/2427-07

Noel, G. J., Wayson, R. L., & John A. Volpe National Transportation Systems Center (U.S.). (2012). MOVES2010a regional level sensitivity analysis (DOT-VNTSC-FHWA-12-05). https://rosap.ntl.bts.gov/view/dot/9706

Office of Transportation and Air Quality. 2020. *MOVES3 Technical Guidance: Using MOVES to Prepare Emission Inventories for State Implementation Plans and Transportation Conformity*. United States Environmental Protection Agency. https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P1010LY2.txt (April 18, 2022).

Pennsylvania Department of Transportation. 2018. *Air Quality Conformity Analysis Report*. Pennsylvania Department of Transportation and CCMPO.

Porter, C., Kall, D., Beagan, D., Margiotta, R., Koupal, J., Fincher, S., Stanard, A., National Cooperative Highway Research Program, Transportation Research Board, & National Academies of Sciences, Engineering, and Medicine. (2015). Input Guidelines for Motor

Vehicle Emissions Simulator Model, Volume 3: Final Report (p. 22212). Transportation Research Board. https://doi.org/10.17226/22212

Replica. (2023). Replica: Data to Drive Decisions about the Built Environment. https://www.replicahq.com/

S&P Global Mobility. (2023). Polk Automotive Solutions. S&P Global. https://www.spglobal.com/mobility/en/products/polk-automotive-solutions.html

S&P Global. (2023). Polk was acquired by IHS Markit, now part of S&P Global. S&P Global. https://www.spglobal.com/en/enterprise/btp/Polk.html

Singh, G., Sivaraman, V., & Hard, E. (2022). State of Emerging Mobility Big Data Sources and its Applications. Texas A&M Transportation Institute. https://static.tti.tamu.edu/tti.tamu.edu/documents/TTI-2022-8.pdf

StreetLight Data, Inc. (2023). StreetLight Data: Transportation Analytics On Demand. StreetLight Data. https://www.streetlightdata.com/

The Atlanta Region's Plan. 2015. *Conformity Determination Report*.

U.S. Department of Transportation, Bureau of Transportation Statistics, & U.S. Department of Commerce, U.S. Census Bureau. (2023). 2021 Vehicle Inventory and Use Survey [dataset]. https://www.census.gov/programs-surveys/vius/data/tables.html

US EPA, O. (2023, January 13). 2020 National Emissions Inventory (NEI) Technical Support Document (TSD) [Other Policies and Guidance]. https://www.epa.gov/air-emissions-inventories/2020-national-emissions-inventory-nei-technical-support-document-tsd

USEPA. (2020). Population and activity of on-road vehicles in MOVES3 (EPA-420-R-20-023; p. 232). Office of Transportation and Air Quality, US Environmental Protection Agency. https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P1011TF8.pdf

USEPA. 2023. *2020 National Emissions Inventory Technical Support Document: Onroad Mobile Sources*. U.S. Environmental Protection Agency.

Venugopal, M., Boardman, M., Gu, C., Bibeka, A., & Li, T. (2022). 2020 On-Road Mobile Source Annual, Summer Weekday, and Winter Weekday Emissions Inventories. Texas Commission on Environmental Quality (TCEQ).

https://www.tceq.texas.gov/downloads/air-quality/2020-on-road-mobile-source-annual-summer-weekday-and-winter-weekday-emissions-inventories

Wejo Limited. (2023). Wejo | Connected Vehicle Data Solutions for Organizations. https://www.wejo.com/

APPENDIX A: MONTGOMERY COUNTY INPUT COMPARISON

Appendix A shows the CDB comparison for Montgomery County. Multiple parameters can vary between the 2020 NEI CDBs and 2020 AERR CDBs. Thus, this section aims to showcase the differences between the two studies using one Texas county.

Figure 23 compares total VMT according to HPMS source type. It should be noted that TTI used the HPMS source type, and EPA used the MOVES source type for vehicle classification. MOVES source type VMT was aggregated for the analysis as HPMS source type. From Figure 23, we can say that EPA distributed VMT into MOVES source types, keeping the total VMT constant.

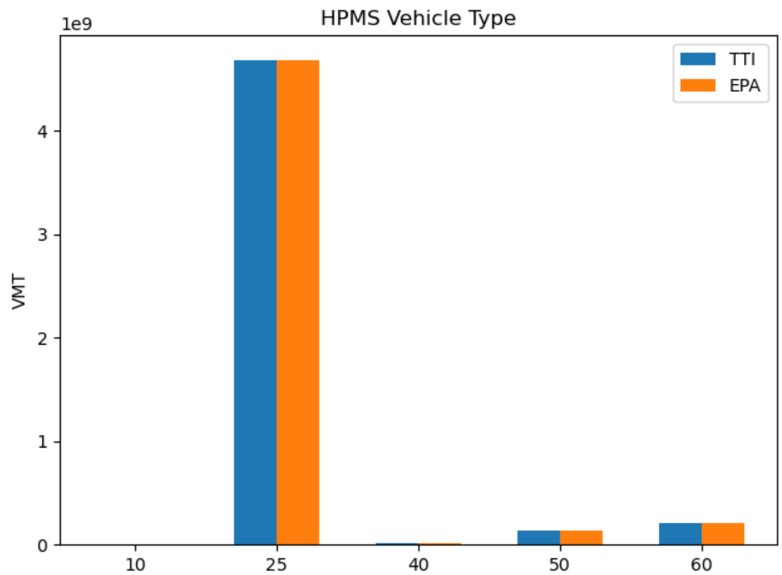
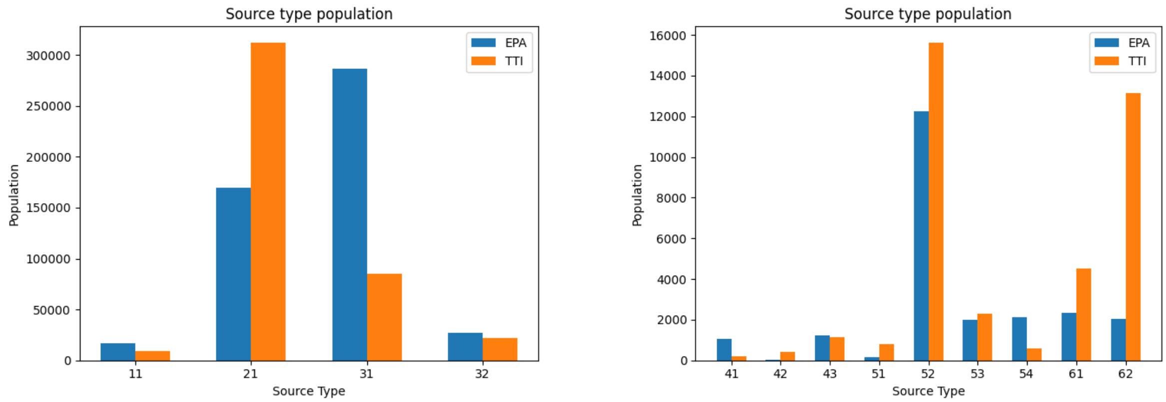


Figure 23. VMT Comparison

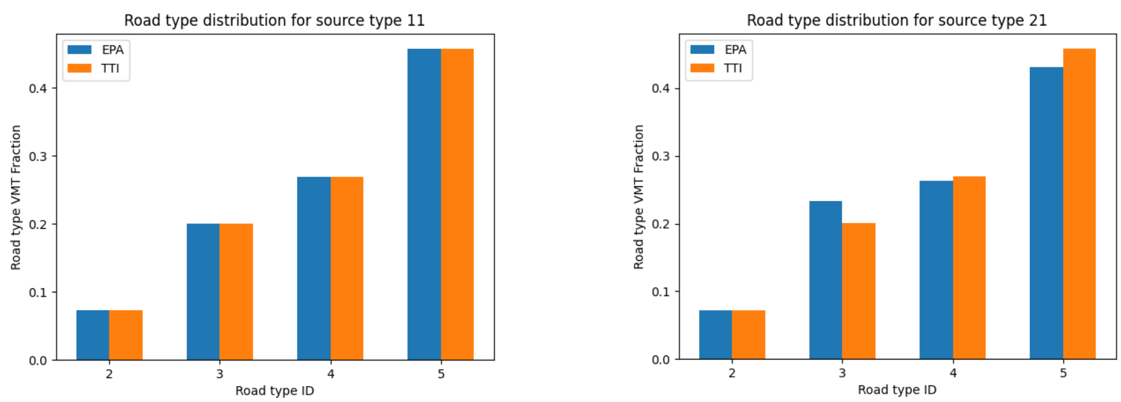
Figure 24 shows a comparison of the source type population estimated. It can be seen that the number of passenger cars is higher for TTI, but EPA estimated more passenger trucks. TTI estimated a larger number of short-haul and long-haul combination trucks.



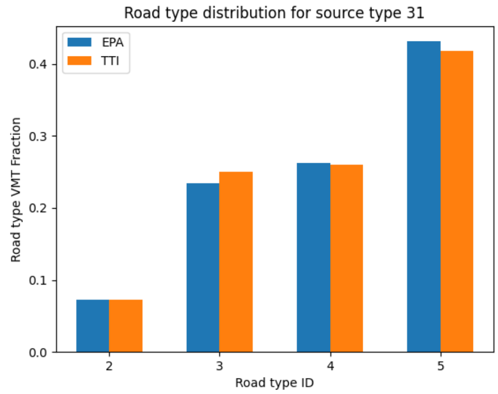
(a) (b)

Figure 24. Source Type Population

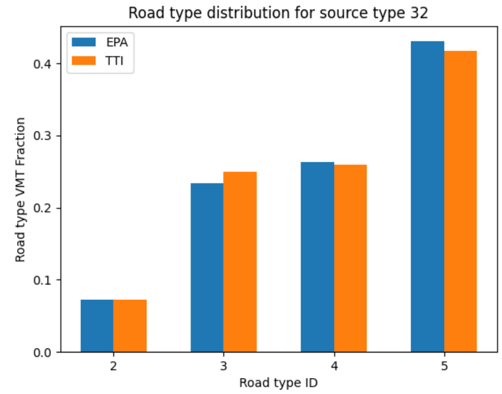
Figure 25 shows road-type VMT distribution for various source types. There are minor differences in VMT fraction for source types 21, 31, and 32. For other source types, the road-type VMT fractions are almost comparable.



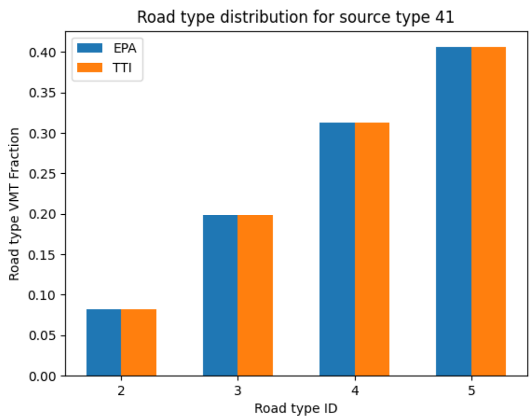
(a) (b)


Road type distribution for source type 31
EPA
TTI
Road type VMT Fraction
0.0
0.1
0.2
0.3
0.4
2
3
4
5
Road type ID

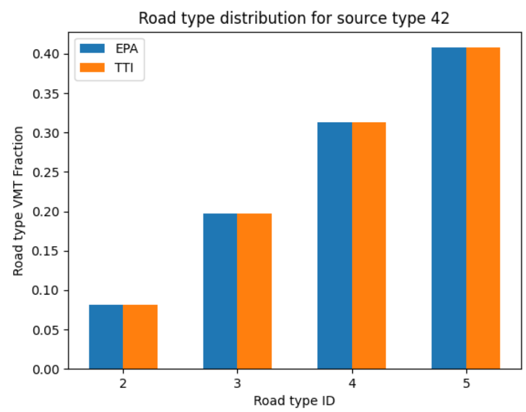
(c)


Road type distribution for source type 32
EPA
TTI
Road type VMT Fraction
0.0
0.1
0.2
0.3
0.4
2
3
4
5
Road type ID

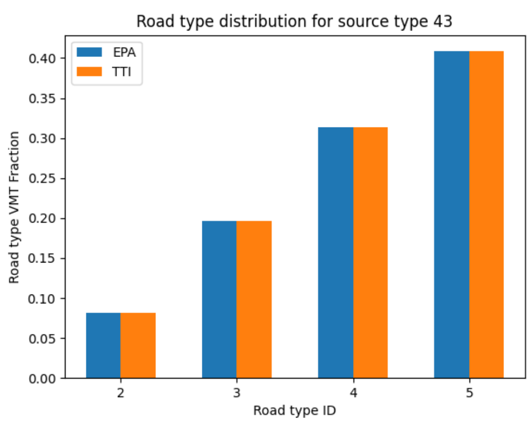
(d)


Road type distribution for source type 41
EPA
TTI
Road type VMT Fraction
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
2
3
4
5
Road type ID

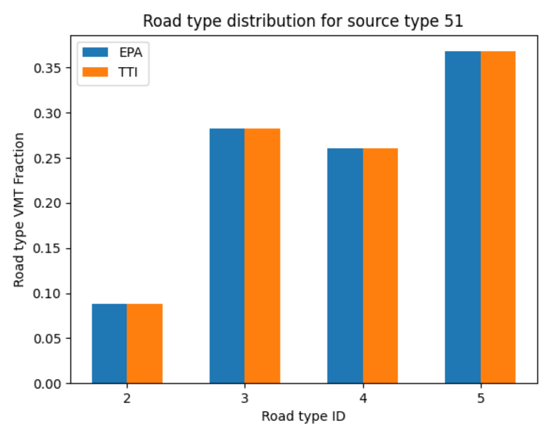
(e)


Road type distribution for source type 42
EPA
TTI
Road type VMT Fraction
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
2
3
4
5
Road type ID

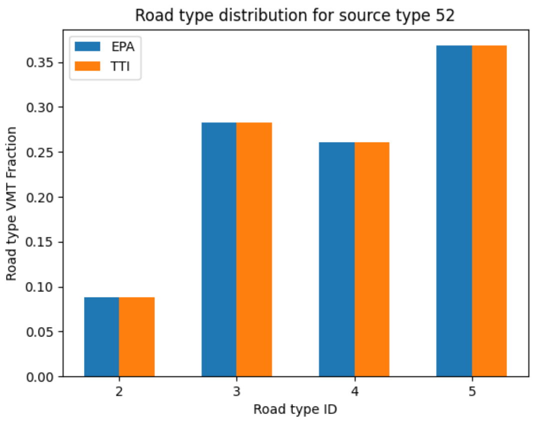
(f)


Road type distribution for source type 43
EPA
TTI
Road type VMT Fraction
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
2
3
4
5
Road type ID

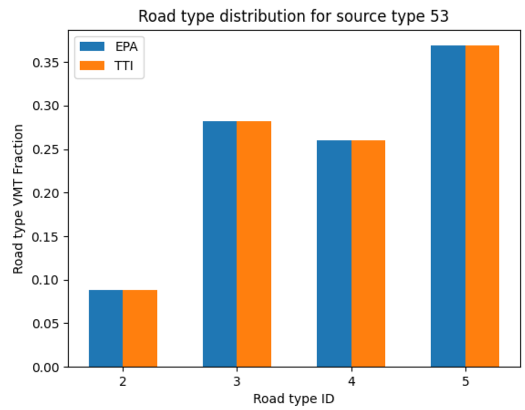
(g)


Road type distribution for source type 51
EPA
TTI
Road type VMT Fraction
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
2
3
4
5
Road type ID

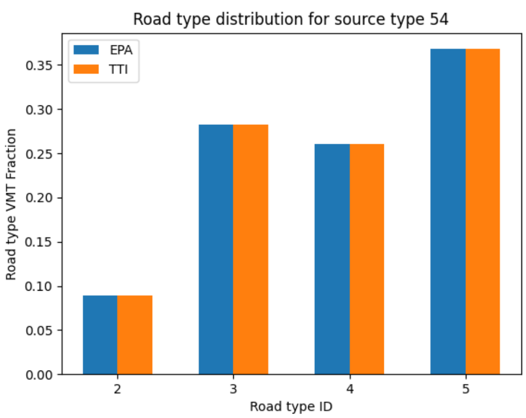
(h)



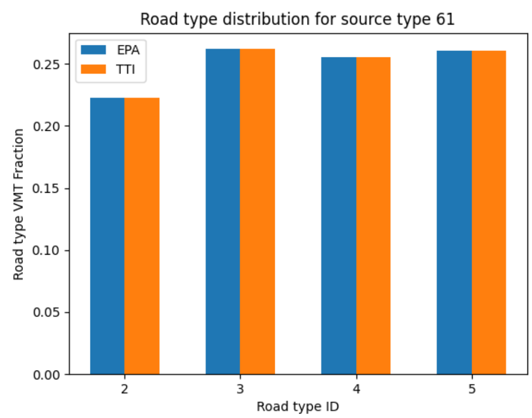
(i)



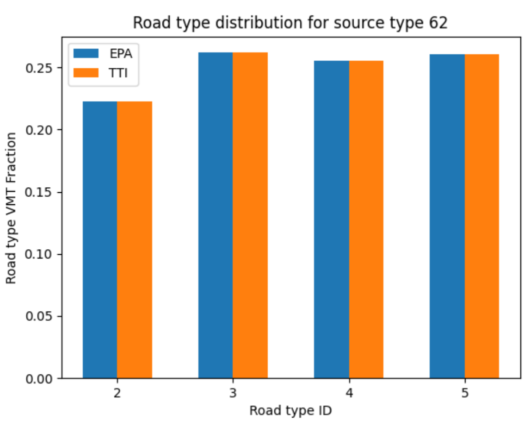
(j)



(k)



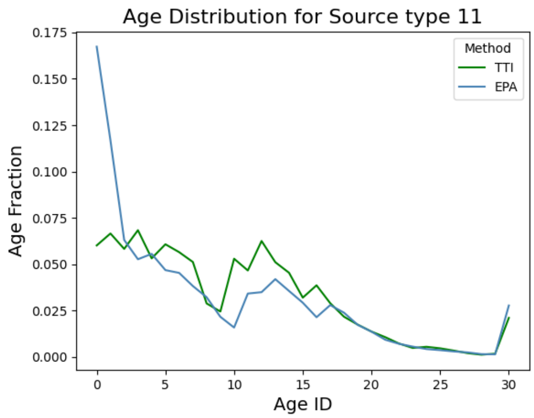
(l)



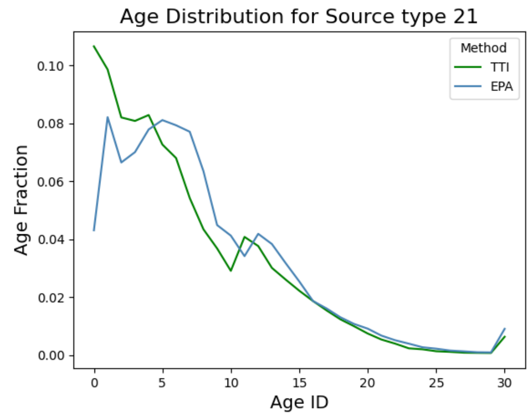
(m)

Figure 25. Road Type Distribution

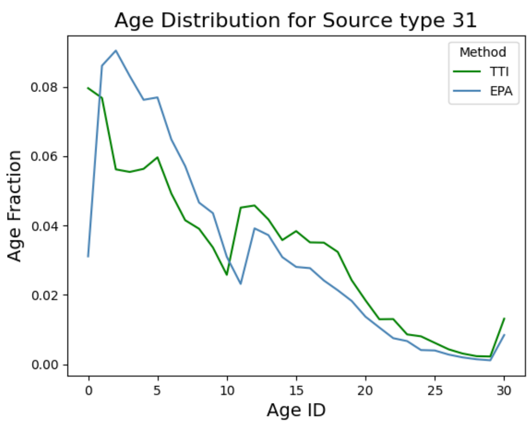
Age distributions are shown in Figure 26. For passenger cars (source type 21), TTI estimated larger fractions for newer vehicles. On the other hand, the EPA estimated a larger fraction for vehicles 5 to 10 years older. For passenger trucks, EPA estimated a larger fraction of newer vehicles, and TTI estimated a larger fraction of older vehicles.



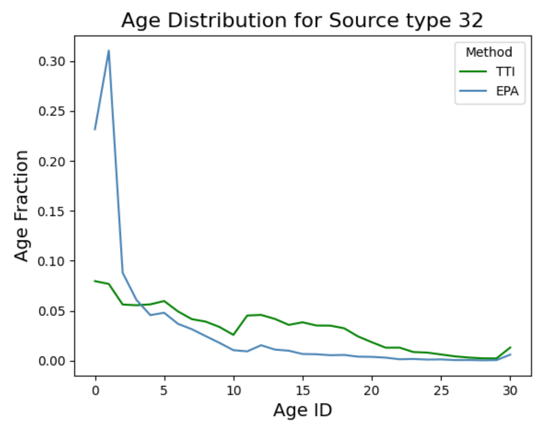
(a)



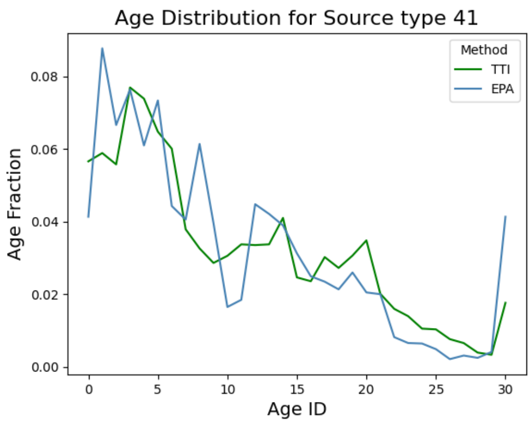
(b)



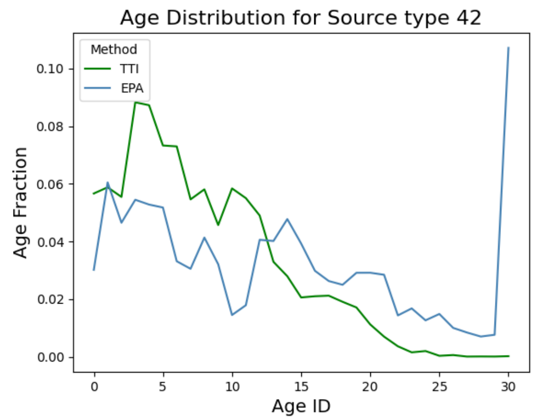
(c)



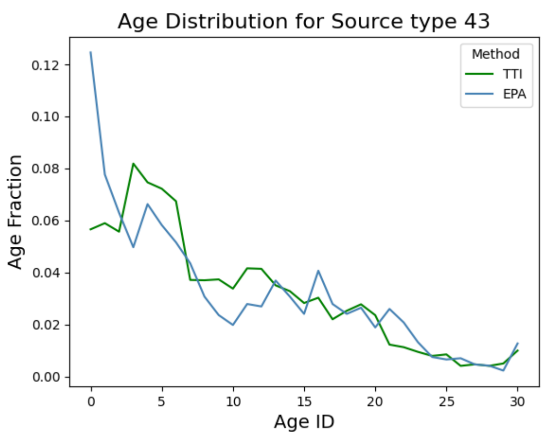
(d)


Age Distribution for Source type 41
Method
TTI
EPA
Age Fraction
Age ID

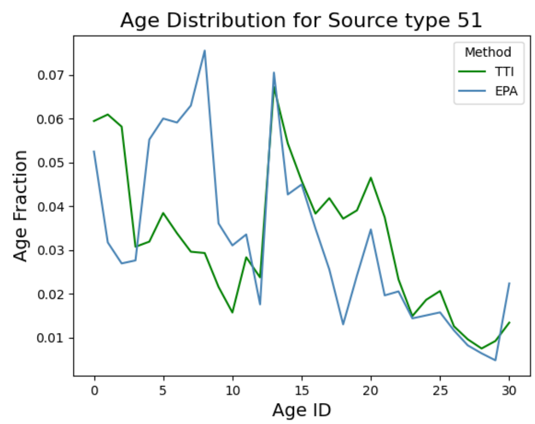
(e)


Age Distribution for Source type 42
Method
TTI
EPA
Age Fraction
Age ID

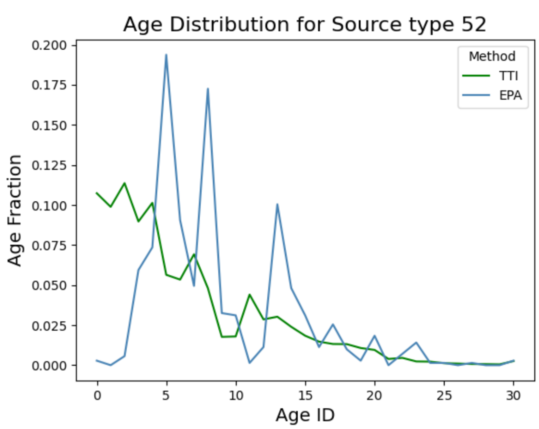
(f)


Age Distribution for Source type 43
Method
TTI
EPA
Age Fraction
Age ID

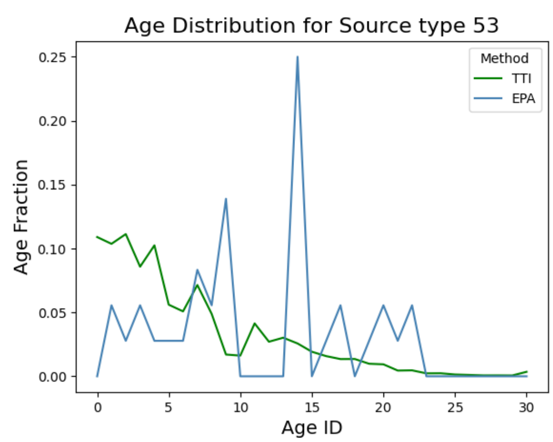
(g)


Age Distribution for Source type 51
Method
TTI
EPA
Age Fraction
Age ID

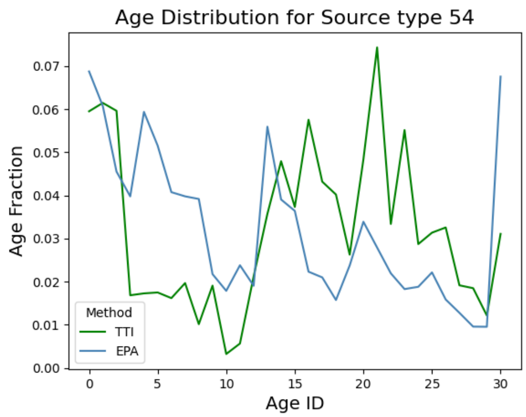
(h)


Age Distribution for Source type 52
Method
TTI
EPA
Age Fraction
Age ID

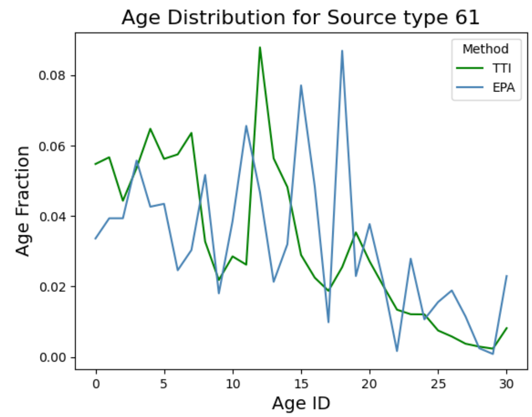
(i)


Age Distribution for Source type 53
Method
TTI
EPA
Age Fraction
Age ID

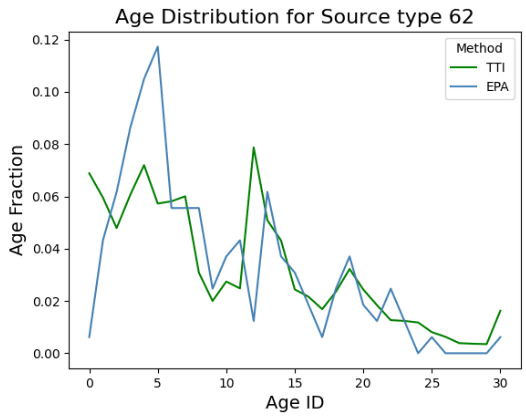
(j)



(k)



(l)



(m)

Figure 26. Source Type Age Distribution

Figure 27 shows speed distribution for passenger cars and road type 2. It can be observed that the speed bin distribution is almost comparable.

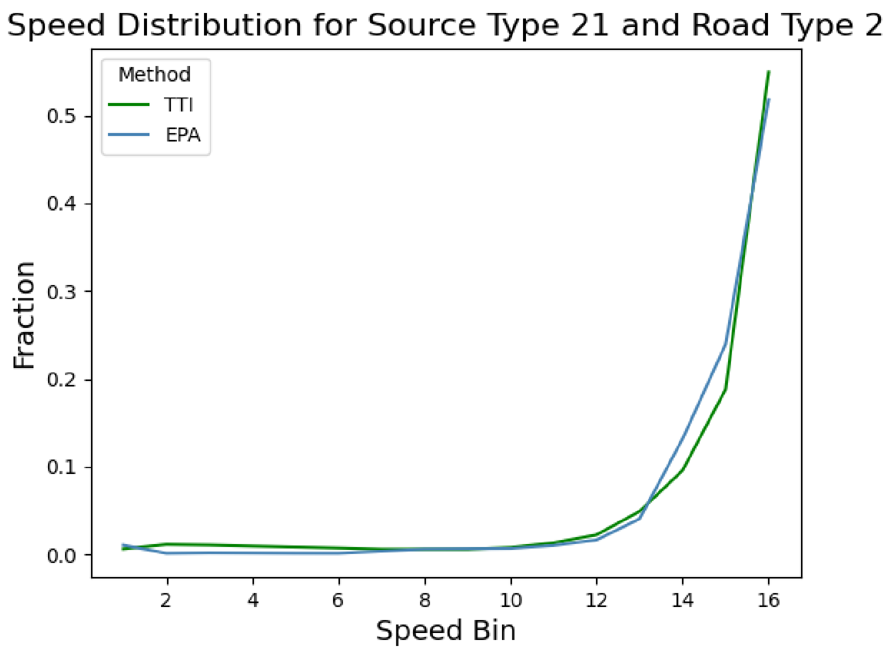
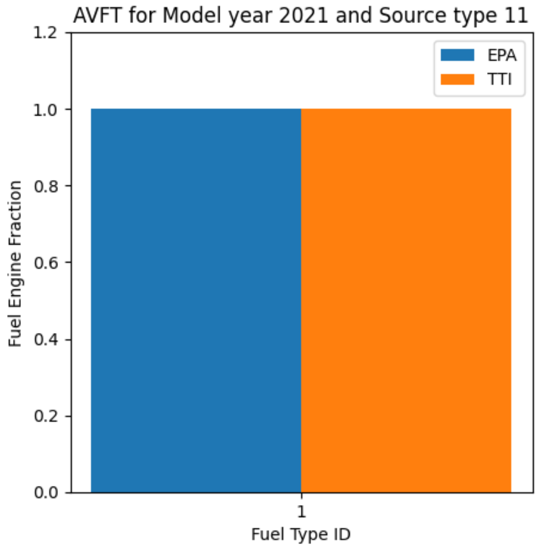
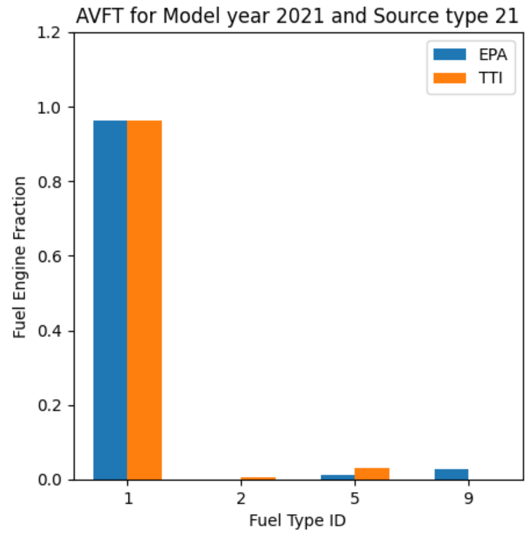


Figure 27. Speed Distribution

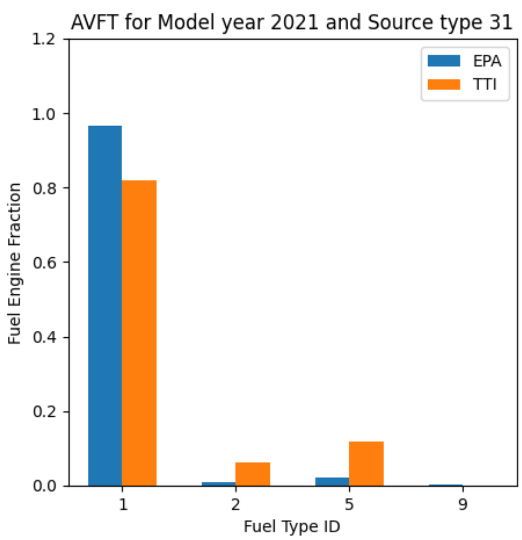
Figure 28 compares the fuel type distribution for a specific model year. The fuel type distribution is almost comparable for most of the source types. For example, source type 11 has 100% of vehicles on gas. Source types 41, 42, and 62 have the same fuel types. Some interesting differences include the inclusion of electric vehicles for passenger cars and passenger trucks EPA.


AVFT for Model year 2021 and Source type 11
Fuel Engine Fraction
Fuel Type ID
EPA
TTI
1.2
1.0
0.8
0.6
0.4
0.2
0.0
1

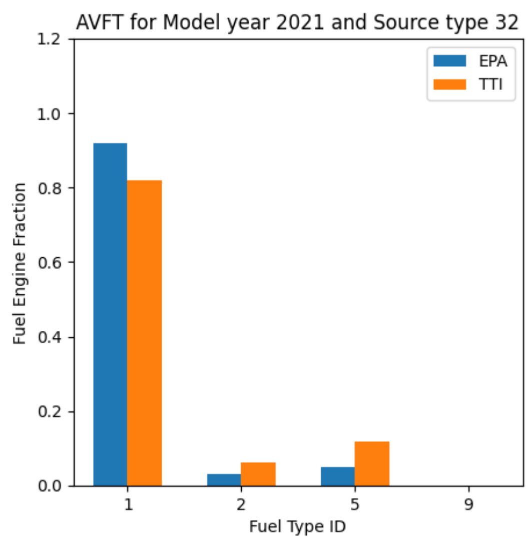
(a)


AVFT for Model year 2021 and Source type 21
Fuel Engine Fraction
Fuel Type ID
EPA
TTI
1.2
1.0
0.8
0.6
0.4
0.2
0.0
1
2
5
9

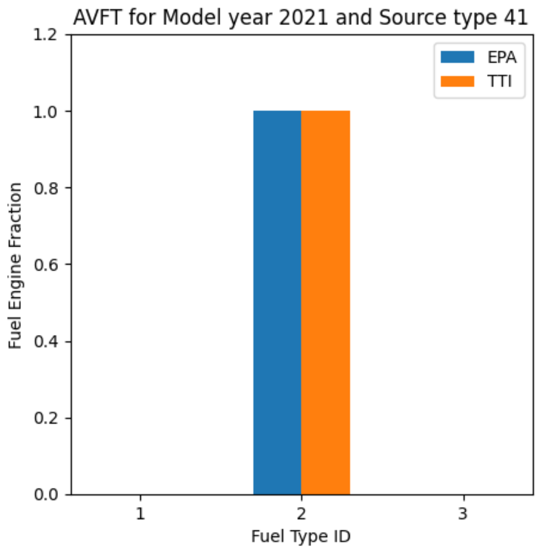
(b)


AVFT for Model year 2021 and Source type 31
Fuel Engine Fraction
Fuel Type ID
EPA
TTI
1.2
1.0
0.8
0.6
0.4
0.2
0.0
1
2
5
9

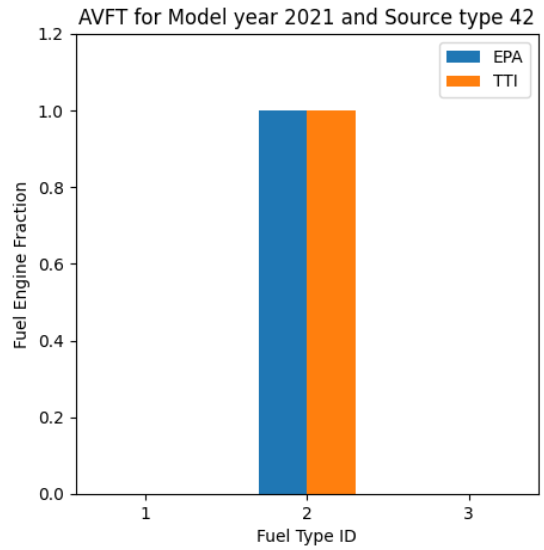
(c)


AVFT for Model year 2021 and Source type 32
Fuel Engine Fraction
Fuel Type ID
EPA
TTI
1.2
1.0
0.8
0.6
0.4
0.2
0.0
1
2
5
9

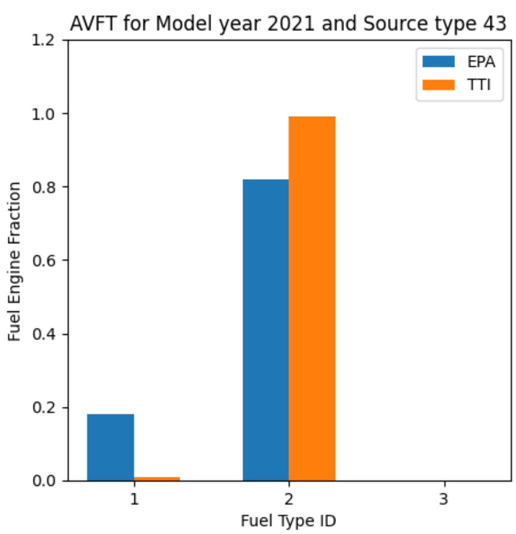
(d)


AVFT for Model year 2021 and Source type 41
EPA
TTI
Fuel Engine Fraction
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1
2
3
Fuel Type ID

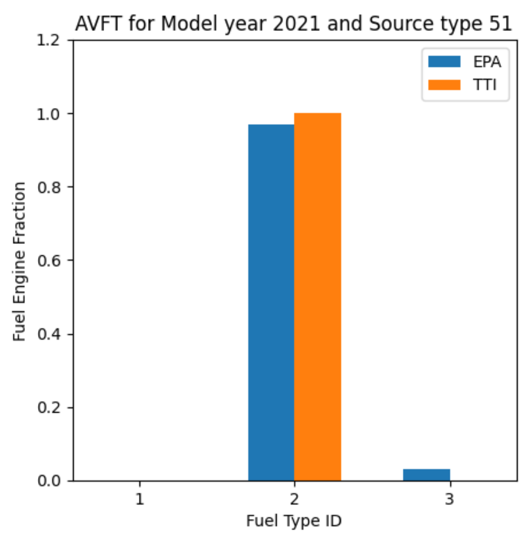
(e)


AVFT for Model year 2021 and Source type 42
EPA
TTI
Fuel Engine Fraction
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1
2
3
Fuel Type ID

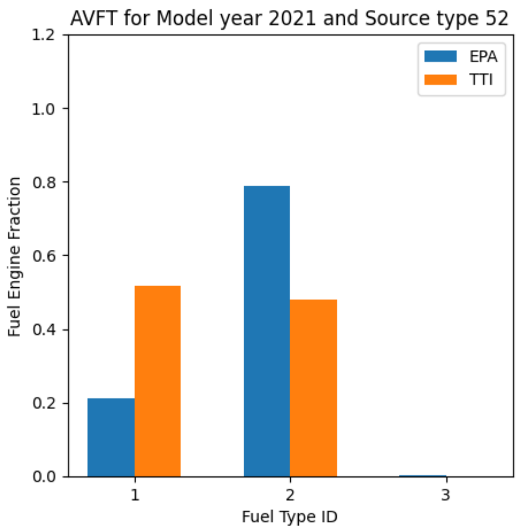
(f)


AVFT for Model year 2021 and Source type 43
EPA
TTI
Fuel Engine Fraction
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1
2
3
Fuel Type ID

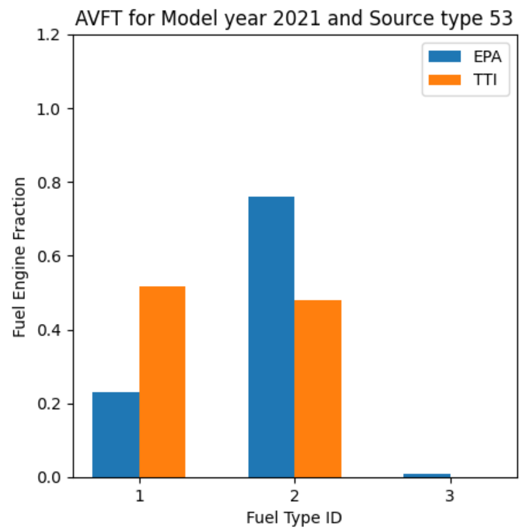
(g)


AVFT for Model year 2021 and Source type 51
EPA
TTI
Fuel Engine Fraction
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1
2
3
Fuel Type ID

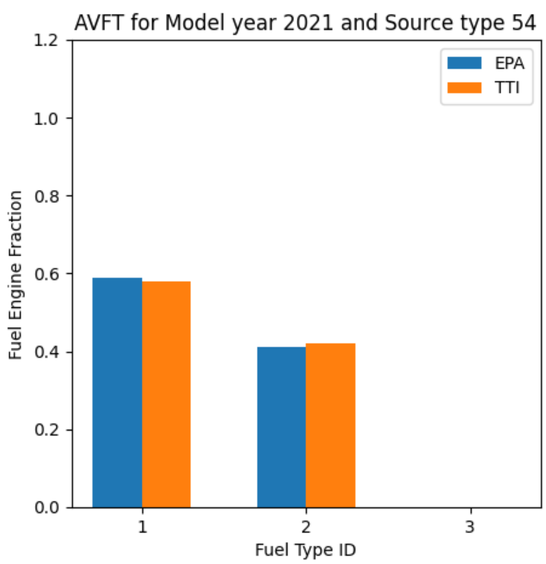
(h)


AVFT for Model year 2021 and Source type 52
EPA
TTI
Fuel Engine Fraction
1.2
1.0
0.8
0.6
0.4
0.2
0.0
1
2
3
Fuel Type ID

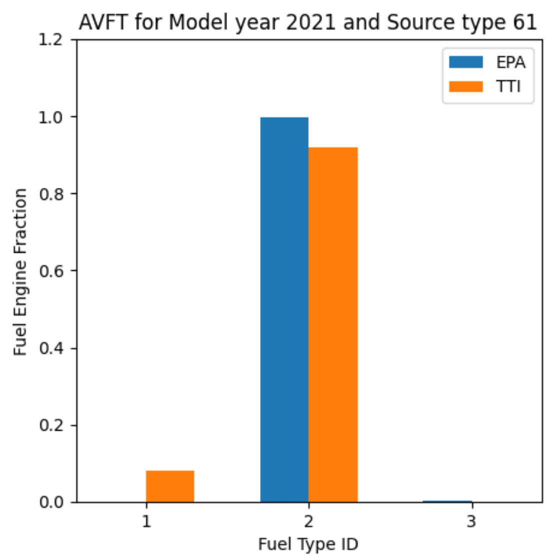
(i)


AVFT for Model year 2021 and Source type 53
EPA
TTI
Fuel Engine Fraction
1.2
1.0
0.8
0.6
0.4
0.2
0.0
1
2
3
Fuel Type ID

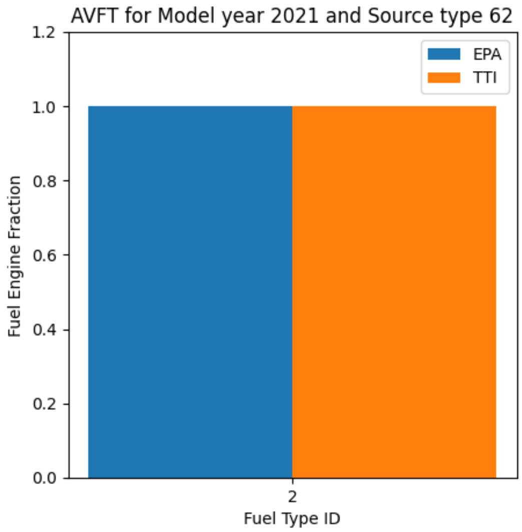
(j)


AVFT for Model year 2021 and Source type 54
EPA
TTI
Fuel Engine Fraction
1.2
1.0
0.8
0.6
0.4
0.2
0.0
1
2
3
Fuel Type ID

(k)


AVFT for Model year 2021 and Source type 61
EPA
TTI
Fuel Engine Fraction
1.2
1.0
0.8
0.6
0.4
0.2
0.0
1
2
3
Fuel Type ID

(l)



(m)

Figure 28. Fuel Type Distribution

APPENDIX B: SENSITIVITY ANALYSIS SCENARIO MOVES CDBS AND MRS FILES

This appendix is available separately in an electronic format (e.g., .docx, .xlsx, .pdf, .txt, .zip, or other format.) and can be provided upon request.

APPENDIX C: SENSITIVITY RESULTS FIGURES

This appendix is available separately in an electronic format (e.g., .docx, .xlsx, .pdf, .txt, .zip, or other format.) and can be provided upon request.