CRC Report No. A-84

STUDY OF MOVES INFORMATION FOR THE NATIONAL EMISSION INVENTORY

Final Report

October 2013

COORDINATING RESEARCH COUNCIL, INC.
5755 NORTH POINT PARKWAY·SUITE 265·ALPHARETTA, GA 30022

The Coordinating Research Council, Inc. (CRC) is a non-profit corporation supported by the petroleum and automotive equipment industries. CRC operates through the committees made up of technical experts from industry and government who voluntarily participate. The four main areas of research within CRC are: air pollution (atmospheric and engineering studies); aviation fuels, lubricants, and equipment performance, heavy-duty vehicle fuels, lubricants, and equipment performance (e.g., diesel trucks); and light-duty vehicle fuels, lubricants, and equipment performance (e.g., passenger cars). CRC's function is to provide the mechanism for joint research conducted by the two industries that will help in determining the optimum combination of petroleum products and automotive equipment. CRC's work is limited to research that is mutually beneficial to the two industries involved, and all information is available to the public.

NOTICE:

Information contained in this report was submitted to the Transportation Research Board for publication in a journal article and was accepted as indicated below. Permission has been granted to the primary author, John Koupal, and to CRC for publication of the original CRC Final Report on the CRC website:

Dear Mr. Koupal:

The Transportation Research Board grants permission to the Coordinating Research Council, to post directly on its website (www.crcao.org) and to distribute via e-mail portions of the paper, "Evaluation and Sensitivity Analysis of MOVES Input Data Submitted for the 2011 National Emissions Inventory," by J. Koupal, T. DeFries, C. Palacios, S. Fincher, and D. Preusse, as identified in your request of December 3, 2013, subject to the following conditions:

1. **Please cite the paper as presented at the 93rd Annual Meeting of the Transportation Research Board, January 2014, Washington, D.C., and accepted for publication in the 2014 series of the Transportation Research Record: Journal of the Transportation Research Board (forthcoming). Copyright, National Academy of Sciences.**
2. **Please acknowledge that the material from your paper is reproduced with permission of the Transportation Research Board.**
3. **None of this material may be presented to imply endorsement by TRB of a product, method, practice, or policy.**

Thank you for your request. Please let me know if you have any questions.

Sincerely,

Javy Awan
Director of Publications
Transportation Research Board

Eastern Research Group, Inc.

Study of MOVES Information for the National Emission Inventory: CRC Project A-84

Final Report

Prepared for:

Coordinating Research Council

Prepared by:

Eastern Research Group, Inc.

October 21, 2013

(this page left intentionally blank)

ERG Project No.: 3881.00.001.001

Study of MOVES Information for the National Emission Inventory: CRC Project A-84

Final Report

Prepared for:

Coordinating Research Council, Inc.
3650 Mansell Road, Suite 140
Alpharetta, GA 30022

Prepared by:

J. Koupal
T.H. DeFries
C.F. Palacios
S.W. Fincher
Eastern Research Group, Inc.
3508 Far West Blvd., Suite 210
Austin, TX 78731

October 21, 2013

(this page left intentionally blank)

Table of Contents

List of Tables

List of Figures

(this page left intentionally blank)

Executive Summary

Introduction

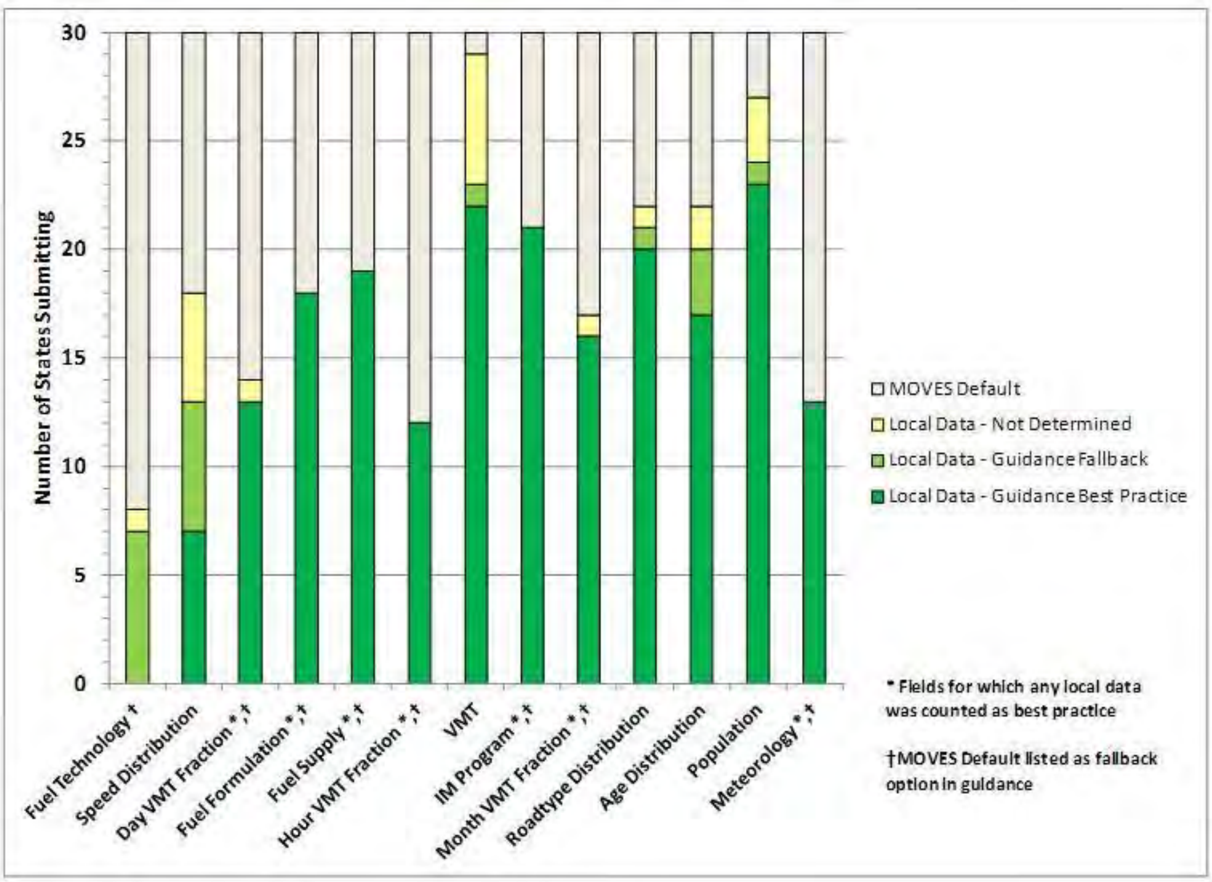
The U.S. EPA compiles the National Emissions Inventory (NEI) to provide a comprehensive nation-wide estimate of annual air emissions of criteria and hazardous pollutants from all sectors. The NEI is developed on a three-year cycle, reporting annual emissions every third year; the analysis year for the current cycle is 2011. During the development cycle EPA works closely with state, local and tribal environmental agencies to compile emissions inventories for each county in the U.S., down to very detailed subsector levels. The resulting compilation provides the official U.S. emissions inventory and serves as the basis for numerous efforts including trends analysis, air quality planning, regulation development and health exposure analyses.

The 2011 NEI is the first to rely solely on EPA's Motor Vehicle Emission Simulator (MOVES) to estimate on-road emissions, replacing the MOBILE6 model used for the last several NEI versions (1999, 2002, 2005, and 2008). State/local/tribal air agencies were encouraged to submit county-level inputs for MOVES, to be used by EPA in developing the on-road inventory for their state. In conjunction with the switch to MOVES, EPA developed a standardized framework for agencies to submit MOVES input data culled from local sources, taking advantage of MOVES features to facilitate the processing and input of local data. As a result, local vehicle fleet and activity data were made available from 30 states, covering over 1,400 counties, representing one of the largest compilations of local MOVES data to date. Under contract with the Coordinating Research Council's (CRC) Atmospheric Impacts Committee, Eastern Research Group, Inc. (ERG) analyzed the submitted MOVES input data, focusing on three questions: 1) how do the methods used to generate the submitted data compare to best practice? 2) What is the range of data submitted for MOVES inputs, and how does it compare to MOVES defaults? And 3) how large is the change in MOVES emission predictions based on the range of submitted data?

Evaluation of State MOVES Submissions vs. Best Practice

Under Task 1a, ERG reviewed all of the documentation provided by state agencies along with CDB submittals to assess, for each CDB input, how the methods compared to EPA's best practice guidance. The results of this assessment are summarized in Figure ES-1, and provide insight into what states are doing to gather data for submission. While several states generated local data for the majority of inputs, no state provided custom data for all inputs. Where custom data wasn't provided, MOVES defaults were submitted (as noted, for some fields MOVES defaults are mentioned as a "fallback" option in the EPA guidance). The likelihood of states submitting custom data depends on the input field, and is an indication of how readily available local data is for that field. State transportation departments collect detailed activity information, providing a ready source for VMT and allocations by road type, month, day and hour, while state vehicle registration databases are a ready source for vehicle population and age distribution inputs; the prevalence of local data for these inputs is therefore higher. Average speed distribution data required by MOVES is more challenging to obtain, particularly for rural areas and urban areas not employing travel demand models; VMT by fuel technology also appears more difficult to obtain. As a result, these inputs have a lower prevalence of local data submission.

Figure ES-1. Breakdown of CDB Inputs for Counties Submitting at Least 1 Local Input



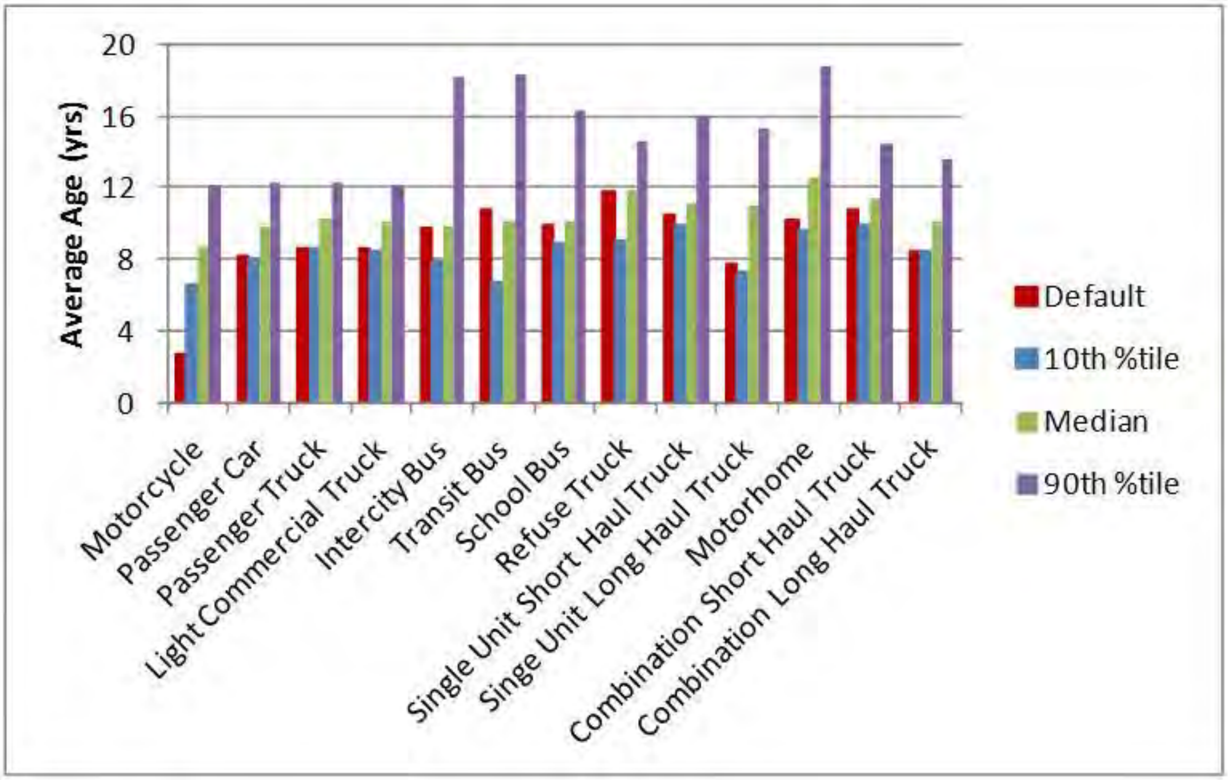
Analysis of State MOVES Submissions

Under Task 1b, detailed analysis was performed on the state-submitted data for five primary CDB inputs: age distribution, vehicle population, VMT, road type distribution, and average speed distribution. For each, county data were selected to represent the 10th and 90th percentile distributions, as well as the median. For age distribution, population and VMT this analysis was done for each source type. For average speed and road type distribution, this was done for hour "clusters". These were compared to MOVES defaults were applicable. The results for each of the five inputs are summarized below.

Age Distribution

Levels for age distribution were chosen based on average age. The median, 10th and 90th percentile average ages for the submitted data, by source type, are shown in Figure ES-2, along with the averages of the default MOVES distributions used in the NEI if states didn't submit data. The data show that for most source types, the average age of the MOVES default is at the lower end of the submitted data distribution. Many states submitted age distribution data with an older average age than MOVES defaults, particularly for buses and heavy trucks. Analysis by county showed that more populous urban areas tended to have younger fleets while rural areas had older fleets, so the population-weighted average age is generally below the median.

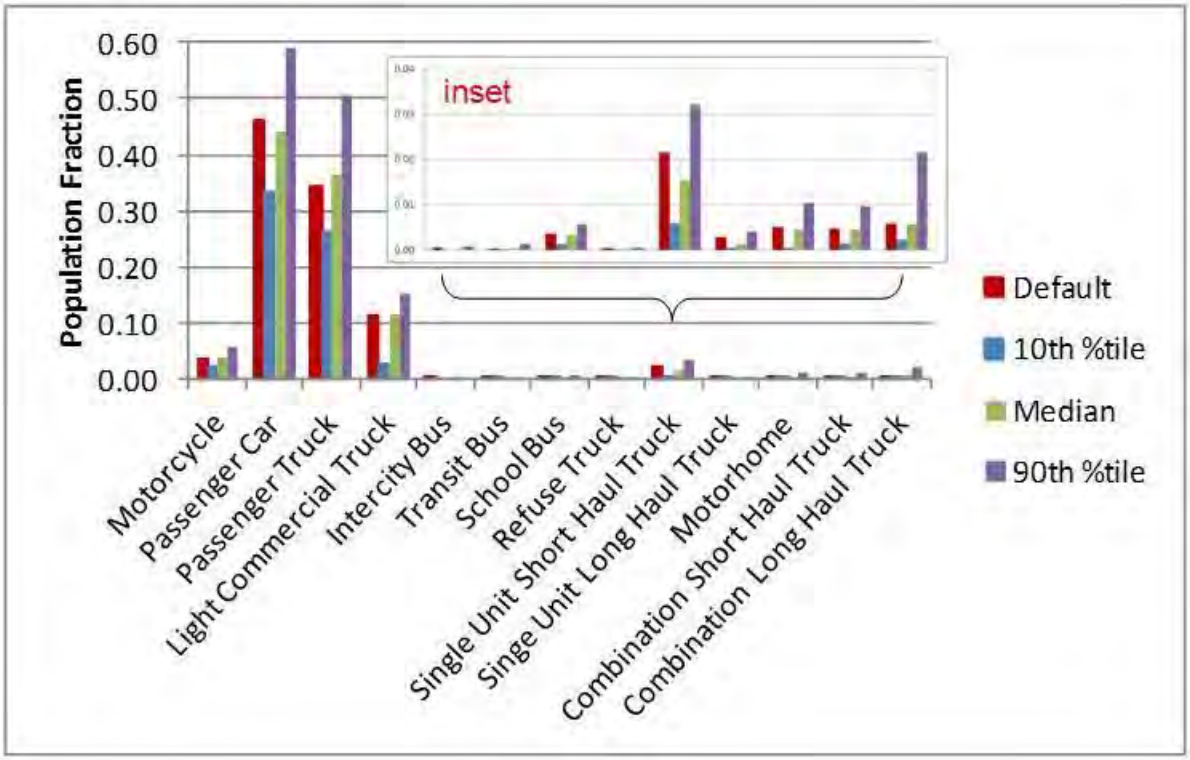
Figure ES-2. Median, 10th and 90th Percentile Average Age, vs. MOVES Defaults



Vehicle Population

Levels for vehicle (source type) population were based on population fraction, i.e. the fraction of total population made up by a given source type. Figure ES-3 shows the median, 10^{th} and 90^{th} percentile population fraction for each source type, along with the MOVES national default; the inset is a magnification of source types with relatively low fractions.

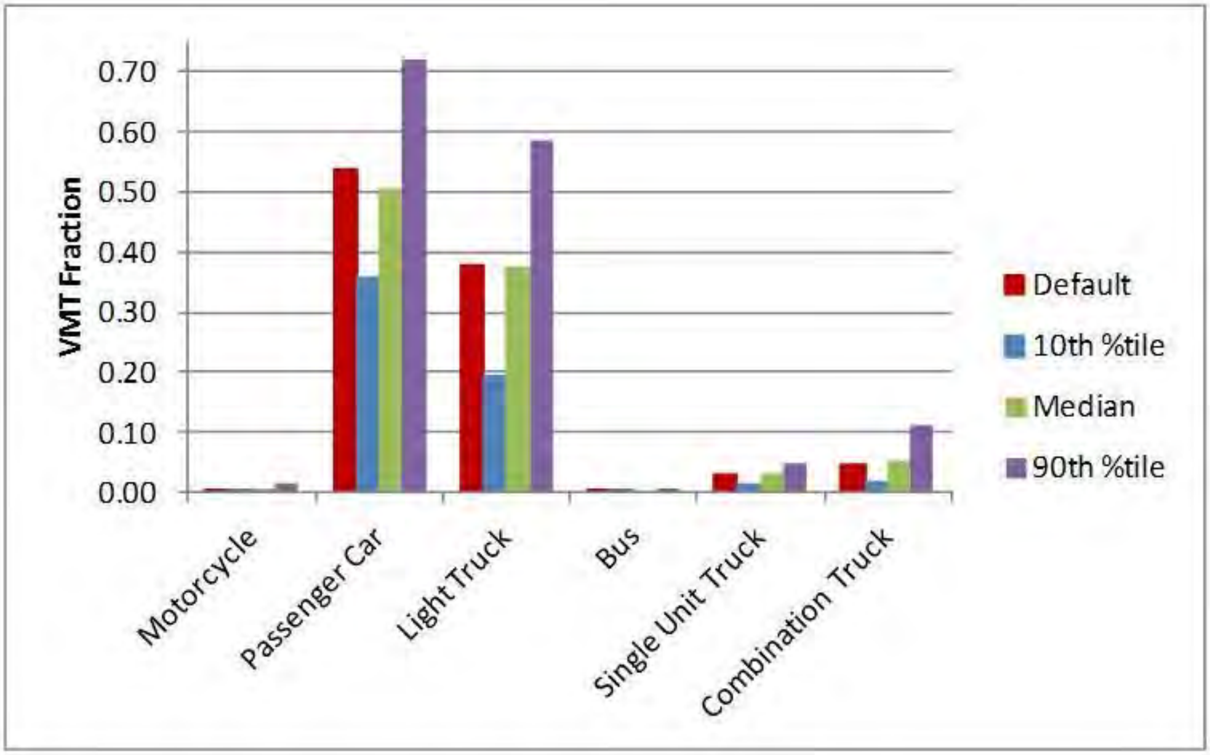
Figure ES-3. Median, 10th and 90th Percentile Population Fractions vs. MOVES Defaults



Vehicle Miles Travelled

Levels for VMT were based on VMT fraction by Highway Performance Monitoring System (HPMS) vehicle class, the classification MOVES uses for VMT. Figure ES-4 shows the median, 10^{th} and 90^{th} percentile fraction for each HPMS vehicle class, along with the MOVES national defaults. As shown, the defaults fall within the range of data, and track the median values well. The data show a large difference for some vehicle classes between the 10^{th} and 90^{th} percentile values, reflecting large variation across the country in how VMT is split. For passenger cars, the 90^{th} percentile fraction is around 2 times that of the 10^{th} percentile fraction; for light trucks, 3 times; for combination trucks, the 90^{th} percentile fraction is nearly 7 times higher than the 10^{th} percentile fraction.

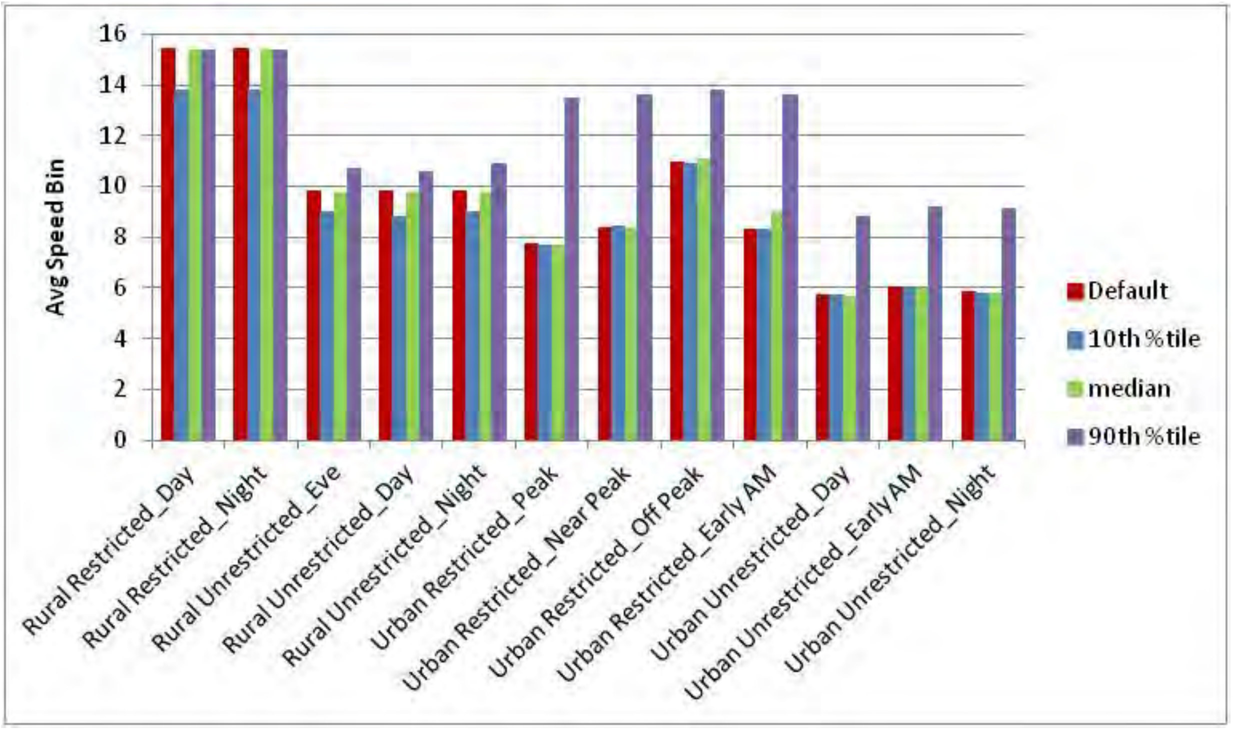
Figure ES-4. Median, 10th and 90th Percentile VMT Fractions, vs. MOVES Defaults



Average Speed Distribution & Road Type Distribution

Average speed distribution and road type distribution were analyzed in conjunction, since the effects of the two inputs are so closely related; assessing the impact of changes in average speed distribution depends on the mix of road types in a particular county. MOVES allows up to 4,992 unique average speed distributions, accounting for each source type, hour, day type (weekend/weekday), and four road types. To distill this down to a manageable size, ERG assumed that average speed on a given road type would be uniform across source types, then analyzed real-world activity data across day and hour to define hour "clusters" of similar average speed distributions that could be analyzed as a block. ERG performed this analysis on a dataset compiled from commercial GPS devices on private vehicles, purchased by EPA to support update of MOVES activity defaults. From this analysis, 13 unique clusters were defined across the four road types, 24 hours and two day types. The submitted data were analyzed to determine the median, 10^{th} and 90^{th} percentile average speeds for each cluster, with results shown in Figure ES-5 (these are shown by MOVES average speed bin – roughly, each bin represents a 5 mph increase in average speed – e.g., Bin 8 is an average speed of 35 mph, Bin 14 is an average speed of 65 mph). For many clusters, the median and 10^{th} percentile, or 90^{th} percentile, values were the same or nearly so; this is another indication that many states submitted MOVES defaults for this input field. The range of road type distributions are not shown reflected in this chart, but were assessed as part of the MOVES sensitivity work performed under Task 2.

Figure ES-5. Median, 10th and 90th Percentile average speed bins, vs. MOVES defaults

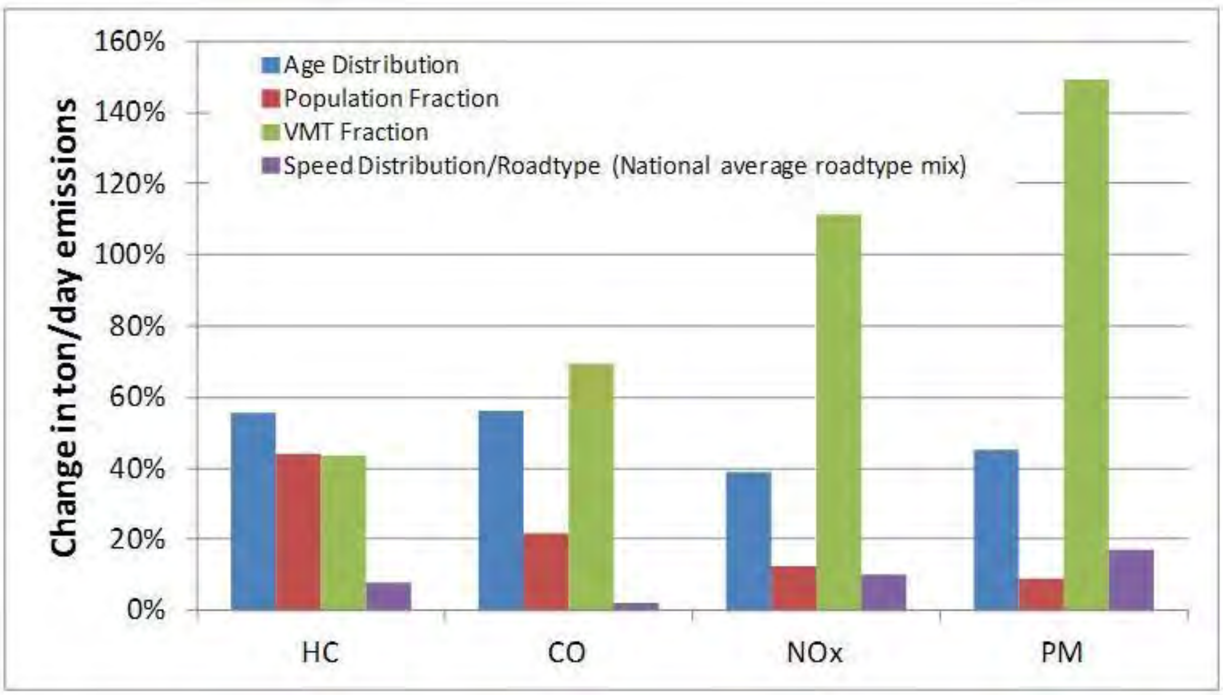


Emissions Sensitivity Analysis

Task 2 evaluated the sensitivity of MOVES emission results to the range of inputs observed in the submissions. This evaluation was performed at the source type level, so that the importance of variations in each source type on overall MOVES emission predictions could be evaluated. The impact of changing each source type's input from the 10^{th} to 90^{th} percentile level on total daily HC, CO, NOx and PM emission predictions was analyzed for a typical July day in Montgomery County, Texas (outside Houston). When one input was varied, all other inputs were held constant at median levels (for the remaining primary inputs) or default levels (all other inputs), so the magnitude of change for each input / source type combination could be evaluated and compared. Evaluating the change in total daily emissions, across all source types and emission processes, provided the proper weighting of inventory contribution from each source type and emission processes in the sensitivity analysis. The overall goal was to establish which source type / input combinations are most influential on total emissions, to help guide MOVES users towards data collection efforts that will yield the most improvement in local inventories.

The Task 2 sensitivity analysis was constructed to allow comparison across the primary inputs; this allows an assessment of how the contribution to total daily emissions of each input compares at the source type (or cluster) level, based on data submitted by states for the NEI. As an overview, Figure ES-6 shows the total increase in daily emissions if all source types went from the 10^{th} to 90^{th} percentile level for each of the inputs (90^{th} to 10^{th} for average speed). These are the compiled totals from the results for each input, presented in the body of the report.

Figure ES-6. Total Change in Ton/Day Emissions Going from 10^{th} to 90^{th} Percentile Input for all Source Types/Hour Clusters, by Input



These results show that the importance of each input varies depending on pollutant. Overall, VMT changes contributed the highest change in emissions, with particularly large increases for NOx and PM. Age distribution is also very influential, particularly for HC and CO. Population is most influential for HC among the pollutants, which reflects the importance of start and evaporative emissions in total HC. The influence of average speed is the lowest, which may seem a surprising result – however, it should be noted that this analysis looked at total daily emissions, where on the whole variability in speed is relatively small and generally restricted to a few hours of the day. An analysis focused on hourly emissions at the project level would likely show a larger influence from speed.

Table ES-1 presents the top five most influential inputs (i.e. largest change in total daily emissions when the input is varied from 10^{th} to 90^{th} percentile) at the source type / cluster level, by pollutant. Note that for this analysis 100 percent road type VMT was used for the average speed clusters, representing an upper bound for this input. As shown, inputs for passenger cars and trucks tend to be towards the top of the list, as would be expected; combination trucks are very prominent for NOx and PM.

Table ES-1. Most Influential Inputs by Source Type / Cluster

	HC	
Source Type/Cluster	**Input Varied**	**Increase in Total Ton/Day Emissions**
Passenger Car	Age Distribution	23.5%
Passenger Truck	Age Distribution	22.3%
Passenger Truck	Population Fraction	15.7%
Passenger Truck	VMT Fraction	13.9%
Passenger Car	Population Fraction	12.4%
	CO	
Source Type/Cluster	**Input Varied**	**Increase in Total Ton/Day Emissions**
Passenger Truck	VMT Fraction	29.8%
Passenger Truck	Age Distribution	24.9%
Passenger Car	VMT Fraction	21.8%
Passenger Car	Age Distribution	21.3%
Light Commercial Truck	Population Fraction	10.8%
	NOx	
Source Type/Cluster	**Input Varied**	**Increase in Total Ton/Day Emissions**
Combination Long Haul Truck	VMT Fraction	39.0%
Passenger Truck	VMT Fraction	21.5%
Combination Short Haul Truck	VMT Fraction	19.9%
Urban Unrestricted_Day	Average Speed Distribution	18.0%
Passenger Car	VMT Fraction	12.4%
	PM	
Source Type/Cluster	**Input Varied**	**Increase in Total Ton/Day Emissions**
Combination Long Haul Truck	VMT Fraction	78.6%
Combination Short Haul Truck	VMT Fraction	34.7%
Urban Unrestricted_Day	Average Speed Distribution	18.3%
Rural Unrestricted_Day	Average Speed Distribution	14.0%
Combination Long Haul Truck	Age Distribution	12.7%

Task 1 & 2 Conclusions and Import for Practitioners

Task 1a found that the frequency of local data and adherence to best practice guidance varies depending on input; VMT inputs follow guidance to the highest degree, while fewer agencies provide average speed for their area. Detailed analysis of age distribution, population fraction, VMT fraction, road type distribution and average speed distribution under Task 1b showed large variation in these inputs. In particular, submitted age distributions were generally older than assumed by MOVES defaults. The amount of data submitted allowed a better assessment of the spread of data across states. This spread in general was larger than expected, which has important implications for improving regional and national onroad emission inventories. The focus by EPA, state air and transportation agencies, and research consortiums such as CRC and TRB on collecting better local data to replace national defaults is justified.

Under Task 2, a MOVES sensitivity analysis was performed based on levels determined from the spread of the submitted data. Sensitivity analysis on MOVES emissions confirms a strong influence on this spread of inputs on emissions predictions. Overall, MOVES *total daily* emissions varied up to 56 percent for HC, 70 percent for CO, 111 percent for NOx and 149 percent for PM based on changes in a single input within the range of state-submitted data. The importance of each input and source type varied depending on pollutant; for HC and CO, dominated by light-duty gasoline sources, passenger car and truck age distribution and population were highly influential. For NOx and PM, a large variation in combination truck VMT fraction led to very large emission differences.

Task 1 and 2 presents the first comparison of the relative influence of emissions by input and source type based on state-submitted data, which will help MOVES users target areas where focused data collection will lead to the most improvement in emission inventory estimates. Overall, this work underscores the critical need for good local data in developing regional emission inventories.

Recommendations for Improvement

The purpose of Task 3 was to provide recommendations for improving data submissions to the NEI, for the ultimate objective of improving NEI emission estimates and its use as a repository for the best local data for use in MOVES. To this end, ERG's recommendations center on two primary objectives: increasing the number of states submitting best practice MOVES inputs, and taking advantage of emerging data sources to improving the scope and quality of local MOVES inputs. The author's recommendations related to these two objectives are provided below, and discussed in detail in Section 6 of the report:

- Conduct outreach to states that did not provide data.
- Use emission sensitivity results to establish priorities for outreach.
- Provide assistance in compiling and converting SIP/Conformity MOVES inputs.
- Take advantage of emerging data sources to broaden and improve MOVES inputs.
- Use national databases to supplement local data.

Overall, MOVES provides a tremendous amount of flexibility in customizing the model to local areas. The air quality community can continue to refine regional and national emission inventories to take advantage of new data and new approaches brought on by more sophisticated data collection technologies, and broader compilation of data sources. The 2011 NEI was an excellent first step in compiling data from many states; the A-84 project showed the variety in approaches and data, and the importance of these data on improving emission estimates. The recommendations discussed above are some ways that practitioners can continue to build on this success for subsequent NEIs.

1.0 Introduction

The U.S. EPA compiles the National Emissions Inventory (NEI) to provide a comprehensive nation-wide estimate of annual air emissions of criteria and hazardous pollutants from all sectors.[1] The NEI is developed on a three-year cycle, reporting annual emissions every third year; the analysis year for the current cycle is 2011. During the development cycle EPA works closely with state, local and tribal environmental agencies to compile emissions inventories for each county in the U.S., down to very detailed subsector levels. The resulting compilation provides the official U.S. emissions inventory and serves as the basis for numerous efforts including trends analysis, air quality planning, regulation development and health exposure analyses.

One important use of the NEI is to estimate the contribution of each sector towards total air emissions in the U.S. The 2008 NEI estimated that on-road vehicles contributed approximately 30 percent of total NOx emissions, 20 percent of VOC emissions and 50 percent of CO emissions nationwide; within urban areas, contributions are higher due to the concentration of vehicle miles travelled (VMT).[2] The significant contribution of the transportation sector to criteria and hazardous air emissions, and the complexity involved in compiling a detailed on-road emissions inventory, requires that significant effort and attention be placed towards the on-road component of the NEI. This effort puts a premium on obtaining county-level data for vehicle fleet and activity patterns.

The 2011 NEI is the first to rely solely on EPA's Motor Vehicle Emission Simulator (MOVES) to estimate on-road emissions, replacing the MOBILE6 model used for the last several NEI versions (1999, 2002, 2005, and 2008). State/local/tribal air agencies were encouraged to submit county-level inputs for MOVES, to be used by EPA in developing the on-road inventory for their state. In conjunction with the switch to MOVES, EPA developed a standardized framework for agencies to submit MOVES input data culled from local sources, taking advantage of MOVES features to facilitate the processing and input of local data.[3] As a result, local vehicle fleet and activity data were made available from 30 states, covering over 1,400 counties, representing one of the largest compilations of local MOVES data to date. Under contract with the Coordinating Research Council's (CRC) Atmospheric Impacts Committee, Eastern Research Group, Inc. (ERG) analyzed the submitted MOVES input data, focusing on

---

[1] U.S. EPA National Emission Inventory Air Pollutant Emission Trends Data http://www.epa.gov/ttnchie1/trends/
[2] U.S. EPA, *Our Nations Air: Status and Trends Through 2010,* Report No. EPA-454-R-12-001, February 2012
[3] U.S. EPA, "Instructions for Submitting MOVES County Database (CDB) Files", September 2012 http://www.epa.gov/ttn/chief/eis/2011nei/submit_moves_inputs.pdf

three questions: 1) how do the methods used to generate the submitted data compare to best practice? 2) What is the range of data submitted for MOVES inputs, and how does it compare to MOVES defaults? And 3) how large is the change in MOVES emission predictions based on the range of submitted data?

2.0 Overview of NEI On-Road Process

The framework for submitting local inputs to the on-road NEI process was the MOVES County Data Manager (CDM) interface, which facilitates data entry to a subset of MOVES data tables most relevant for constructing local on-road emission inventories. The result of this process is a unique MOVES County Database (CDB) for each county submitting data. The primary data inputs in CDBs include:

- Vehicle miles travelled (VMT)
- Vehicle population
- Age distribution
- Average speed distribution
- Road type VMT distribution
- Month/Day/Hour VMT distributions
- Fuel properties & market share
- Inspection/Maintenance (I/M) program parameters
- Meteorology data

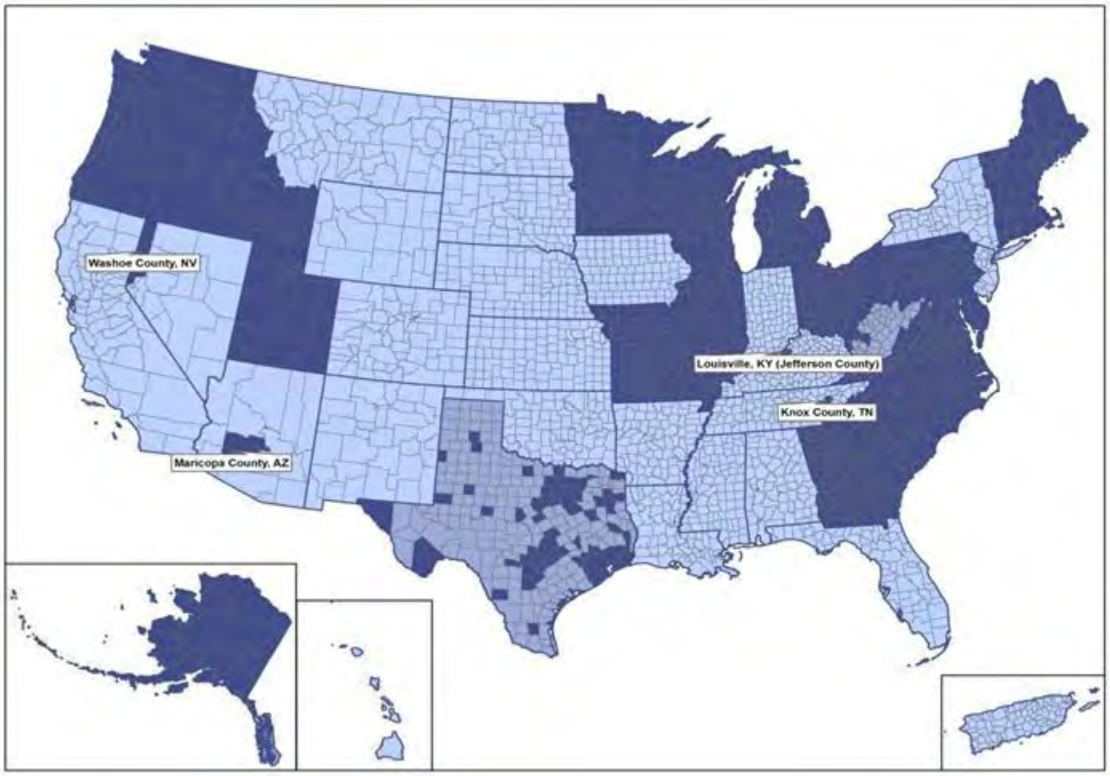
The complete list of data tables that make up a CDB are shown in Table 2-1. The list includes the MOVES data tables that house the above listed data, as well as “information” tables that provide MOVES with meta data necessary for a particular run (county, state, year etc.).

Table 2-1. MOVES CDB Tables

CDB TABLE	DESCRIPTION OF CONTENT
auditlog	Information about the creation of the database
avft	Diesel sales fractions
avgspeeddistribution	Average speed distributions
county	Description of the county
dayvmtfraction	VMT distribution across the type of day
fuelformulation	Fuel properties
fuelsupply	Fuel differences by month of the year
fuelsupplyyear	Year for the fuel properties
hourvmtfraction	VMT distribution across the hours of the day
hpmsvtypeyear	Total annual VMT by HPMS vehicle type
imcoverage	Description of the Inspection and Maintenance program
monthvmtfraction	VMT distribution across the months of the year
roadtype	Description of the road types
roadtypedistribution	VMT distribution across the road types
sourcetypeagedistribution	Distribution of vehicle ages
sourcetypeyear	Vehicle populations
state	Description of the state
year	Year of the database
zone	Allocations of starts, extended idle and vehicle hours parked to the county
zonemonthhour	Temperature and relative humidity values
zoneroadtype	Allocation of road types to the county
countyyear	Description of the Stage 2 program
emissionratebyage*	Implementation of California standards [not part of CDB but included for NEI since state-specific data is applicable]
sccroadtypedistribution	Allocation of results to SCC categories

CDBs were submitted to EPA through their Emission Inventory System (EIS), with an initial submission round ending in February 2013. In the initial round 1,342 CDBs were submitted by states. Figure 2-1 shows in dark blue the counties for which CDBs containing at least one local MOVES input were submitted; for some states, data were only supplied for a subset of counties. California (which does not use MOVES), Texas and several Tribal governments submitted completed emission inventories to EIS rather than modeling inputs; however, 69 CDBs used by Texas in their inventory preparation were also obtained for this analysis. A second round of NEI submissions later added inputs for New Jersey and Colorado, but these were not included in this analysis.

Figure 2-1. Counties with MOVES CDBs used in analysis (dark blue)



(this page left intentionally blank)

3.0 Task 1a: Qualitative Evaluation vs. Best Practice for State Submissions

Along with CDBs, agencies provided checklists detailing which data tables were provided (partial CDBs were accepted), and documentation for how local data were developed. A summary of the checklists, showing the number of counties submitting data tables, is shown in Table 3-1. The summary shows that the number of counties with submitted inputs vary greatly depending on inputs. The most populous submissions were for VMT (HPMSVtypeYear) and vehicle population (SourceTypeYear), both over 1,300 counties. The number of counties submitting data can be misleading, however, as a submission does not necessarily represent local data. Many states submitted MOVES default data as part of their CDB submission.

As part of the 2011 NEI plan, EPA provided guidance to agencies on the process and format of submitting local CDB data through EIS. This guidance did not include detailed technical specifications on how local data should be developed. However, for the development of State Implementation Plan (SIP) inventories, EPA has written a technical guidance document that discusses preferred data sources for each of these inputs, and fallback options where data is not available.[4] Responsible agencies are not required to follow this guidance for the NEI, but because the guidance encompasses best practice and identifies the most readily available data sources for CDB inputs, many states submitting data follow EPA's technical guidance. A summary of EPA guidance for data-related CDB inputs is shown in Table 3-2.

Under Task 1a, ERG reviewed all of the documentation provided by state agencies along with CDB submittals to assess, for each CDB input, how the methods compared to EPA's best practice guidance. The results of this assessment are summarized in Figure 3-1, and provide insight into what states are doing to gather data for submission. While several states generated local data for the majority of inputs, no state provided custom data for all inputs. Where custom data wasn't provided, MOVES defaults were submitted (as noted, for some fields MOVES defaults are mentioned as a "fallback" option in the EPA guidance). The likelihood of states submitting custom data depends on the input field, and is an indication of how readily available local data is for that field. State transportation departments collect detailed activity information, providing a ready source for VMT and allocations by road type, month, day and hour, while state vehicle registration databases are a ready source for vehicle population and age distribution inputs; the prevalence of local data for these inputs is therefore higher. Average speed distribution data required by MOVES is more challenging to obtain, particularly for rural areas

[4] *Using MOVES to Prepare Emission Inventories in State Implementation Plans and Transportation Conformity: Technical Guidance for MOVES2010, 2010a and 2010b,* U.S EPA, Report No. EPA-420-B-012-028, April 2012

and urban areas not employing travel demand models; VMT by fuel technology also appears more difficult to obtain. As a result, these inputs have a lower prevalence of local data submission.

Table 3-1. Number of Counties with Submitted Data, by State and MOVES CDB Input Table (Informational tables not included)

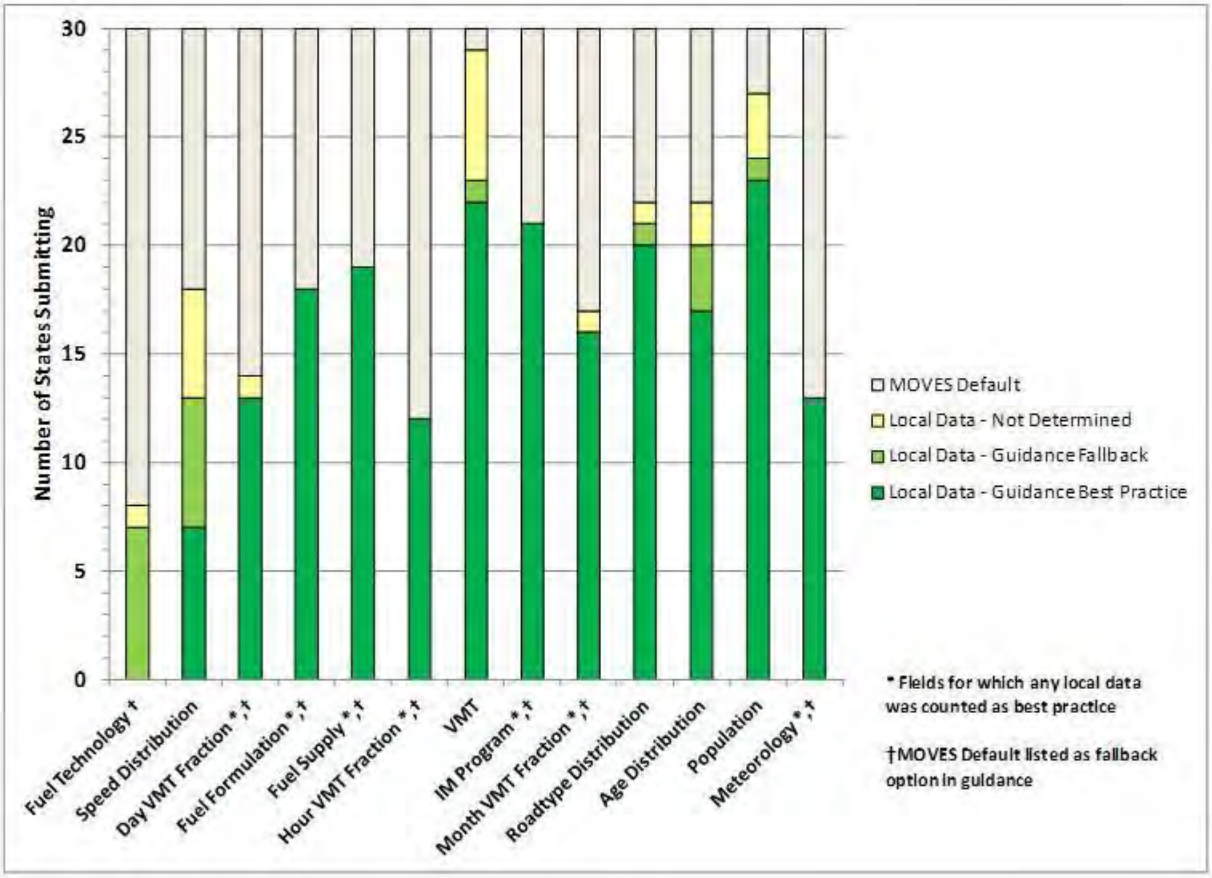
	CDB Table													
State/County	**AVFT**	**AvgSpeedDistribution**	**DayVMTFraction**	**FuelFormulation**	**FuelSupply**	**HourVMTFraction**	**HPMSVtypeYear**	**IMCoverage**	**MonthVMTFraction**	**RoadType**	**RoadTypeDistribution**	**SourceTypeAgeDistribution**	**SourceTypeYear**	**EmissionRateByAge**
Alaska	29	29	29	29	29	29	29	2	29		29	29	29	
Arizona (Maricopa County)	1	1	1	1	1	1	1	1	1	1	1	1	1	
Colorado								11**						
Connecticut		8	8	8	8	8	8	8		8	8	8	8	8
Delaware*		3		3	3		3	3	3	3	3	3	3	
District of Columbia	1	1	1	1	1	1	1	1	1	1	1	1	1	
Georgia		21	159			21	159	13	159	21	159	159	159	
Idaho	44	44	44		44	44	44	2	44	44	44	44	44	
Illinois		102	102	102	102	102	102	11	102	102	102	102	102	
Kentucky (Jefferson County)	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Maine		16	16	16	16	16	16	1	16	16	16		16	
Maryland	24	24	24	24	24	24	24	24	24		24	24	24	24
Massachusetts*		14	14		14	14	14	14	14	14	14	14	14	14
Michigan		83	83	83	83	83	83		83	83	83	83	83	76
Minnesota				87	87		87				87	4**	87	
Missouri					110		115	5			115	115	115	
Nevada (Washoe County)				1	1		1							
New Hampshire				10	10		10	10				10	10	
New Jersey		21	21	21	21	21	21	21	21		21	21	21	
North Carolina		19		100	100	100	100	100				100	100	
Ohio	88	88	88	1	1	88	88	14	88	23	88	88	88	
Oregon				36				6						
Pennsylvania		67	67	67	67	67	67	67	67		67	67	67	67
Rhode Island							5						5	
South Carolina			46			46	46		46				46	
Tennessee (Knox County)			1				1		1		1			
Utah	29	29	29	29	29	29	29	29	29	29	29	29	29	29
Vermont							14						14	
Virginia		134	40	34	34		134	10	40		134	134	134	
Washington	1		39	39	39	39	39	5	39		39	39	39	
West Virginia							13		13		13	13	13	
Wisconsin		7		6			72	7			72	72	72	
Total	**218**	**712**	**813**	**699**	**825**	**734**	**1327**	**355**	**821**	**346**	**1151**	**1157**	**1325**	**219**

**EIS checklist submitted blank, determined from documentation*
***Submitted directly to EPA staff, not through EIS*
**** This table includes submissions in 2nd round not included in the A-84 analysis, as discussed in Section 2.0*

Table 3-2. Summary of EPA's MOVES Technical Guidance for CDB Inputs

INPUT	BEST PRACTICE	SUGGESTED FALLBACKS
Source Type Population	Direct data from state registration data, local transit agencies, school districts, bus companies, refuse haulers	Ratio to VMT based on MOVES default VMT/Population ratios; MOVES default sourceType split; MOBILE6 inputs
Age Distribution	Unique registration data for each source type	Applying the same distribution to multiple sourcetypes; MOVES defaults; MOBILE6 inputs
VMT	State DOT data and/or Travel Demand Models calibrated to HPMS; EPA convertor to adjust to annual	Count-based programs, MOBILE6 inputs
Average Speed Distribution	Post-processing of Travel Demand Model output into each MOVES speed bin, by hour	Applying the same distribution to multiple source types (expected); single average speed; peak/off-peak or daily; MOBILE6 inputs
Speeds on Local Roads	*Include local road speeds as part of unrestricted distributions*	
Highways	*Reflect speeds on the highway only, not ramps (handled separately)*	
Road Type Distribution	"Consistent with transportation planning" (same sources as VMT)	Applying the same distribution to multiple source types within same HPMS class (expected), or even across HPMS classes; MOBILE6 inputs
Ramp Fraction	Optional input, no specific guidance on data sources given	MOVES defaults
Fuel Type & Technology (gas/diesel/CNG mix)	VMT by fuel type	Population (expected); MOVES defaults; MOBILE6 inputs
Fuel Formulation (fuel properties)	RFG fuel property info for RFG areas; regulatory RVP level in RVP control areas, accounting for 1 psi waiver; fuel survey data for non-RFG areas (e.g. NIPER, AAM)	Modify fuel properties where data available, use defaults for others; using fuel with desired properties in same PADD/year; MOVES defaults
Fuel Supply (market share)	Volume data	MOVES defaults
Inspection/Maintenance (I/M) Program	Modify defaults to make consistent with actual program	MOVES defaults
Compliance Rate	*Operating program data: sticker surveys, license plate surveys, no. of tests vs. potential tests*	*Should not use 100%; automatic registration denial program can assume 96%, but should update based on operating data when available*
Waiver Rate	*Historical waiver rates*	
Meteorology	Local data via sources such as National Climatic Data Center	MOVES default
Month, Day & Hour VMT Fractions	No explicit guidance	MOVES defaults

Figure 3-1. Breakdown of CDB Inputs for Counties Submitting at Least 1 Local Input



It is important to note that EPA did not intend to use all of the CDB input fields provided by states for the NEI. In particular, fuel formulation (properties), fuel supply (market share) and meteorology data were supplied by EPA even if states provided data, to ensure uniform methods and data sources across the entire U.S. State-supplied data characterizing the vehicle fleet (e.g., age distribution, population) and activity (e.g., VMT, speed distribution) were used by EPA in the NEI and therefore were a higher priority to assess for this analysis.

4.0 Task 1b: Analysis of Primary5 Tables for State Submissions

4.1 Selection of Tables to Investigate

The purpose of Task 1b was to perform an in-depth quantitative analysis of state submissions for a subset of CDB inputs, to assess the variability in inputs, and how they compared to MOVES defaults used in the NEI in the absence of state data. This was not possible for every data field contained in CDB, so ERG narrowed the focus of the task to data input which a) would be used in the NEI if provided by states, and b) were likely to have the largest impact on MOVES emission predictions when varied (to be confirmed in Task 2). Table 4-1 shows ERG's initial assessment of these, under the heading "Applicability to A-84".

Table 4-1. Initial Assessment of Input Data Applicability to A-84

Table	General Description	Applicability to A-84
auditlog	Meta-information about database creation	Low – informational only
avft	Fuel technology fractions (gas/diesel/CNG)	Medium – large variations in state-submitted data not expected
avgspeeddistribution	Distribution of average speeds	High – state-submitted data likely, significant variation from default likely, MOVES results highly sensitive to changes
county	County being modeled	Low - informational only
dayvmtfraction	VMT distribution by weekday/weekend	Medium – not significant factor in annual inventories
fuelformulation	List of possible fuels in area	Low – do not expect states to submit new formulations
fuelsupply	Market share of fuel formulations	Medium – do not expect states to have comprehensive information on fuel market share
fuelsupplyyear	Year market share applies to	Low – informational only
hourvmtfraction	Distribution of VMT by hour of the day	Medium – not a significant factor in annual inventories
hpmsvtypeyear	Total Vehicle Miles Travelled (VMT) by vehicle class	High – state-submitted data likely, significant variation from default likely, MOVES results highly sensitive to changes
imcoverage	I/M program parameters	Medium - states not expected to have significant changes
monthvmtfraction	Distribution of VMT by month	Medium - while more important for annual inventories than day/hour fractions, large variations from default not expected
roadtype	Description of road types	Low – informational only

Table	General Description	Applicability to A-84
roadtypedistribution	Distribution of VMT across road types	High - state-submitted data likely, significant variation from default likely, MOVES results highly sensitive to changes
sourcetypeagedistribution	Fleet age distribution	High - state-submitted data likely, significant variation from default likely, MOVES results highly sensitive to changes
sourcetypeyear	Vehicle populations	High - state-submitted data likely, significant variation from default likely, MOVES results highly sensitive to changes
state	State the county is in	Low - informational
year	Year being modeled	Low - informational
zone	Activity allocation within a county	Low – feature not expected to be applied for NEI
zonemonthhour	Meteorology	Medium – state-submitted data likely, but do not expect significant variations from defaults
zoneroadtype	Activity allocation within a county	Low – feature not expected to be applied for NEI
countyyear	Stage 2 refueling program	Low – feature not expected to be applied for NEI
emissionratebyage	To reflect California standard implementation for Section 177 states	Low – this is included in NEI submissions for states to confirm program that should be modeled – i.e. not a "sensitivity"
sccroadtypedistribution	Allocations to SCCs	Low – state-submitted data not expected

Based on the initial assessment above, ERG proposed to focus the evaluation of Task 1b on five "primary" inputs with high applicability to the objectives of A-84, that were used in the NEI if states provided the information: Average Speed Distribution, Vehicle Miles Travelled (VMT), Road Type Distribution, Age Distributions, and Vehicle Populations. In this investigation we call these the Primary5 tables.

4.2 Structure of the Primary5 MOVES Tables

The Primary5 MOVES tables under investigation in this study are made up of variables that describe the values in the tables. Tables 4-2 through 4-6 define several of the key variables (SourceTypeID, HPMSVtypeID, AgeID, RoadTypeID, HourDayID, and AvgSpeedBinID) used by the Primary5 MOVES tables. The variable definitions presented here are intended to match the definitions that are discussed in MOVES guidance reports.[5]

Table 4-2 shows the two variables (HPMSVtypeID, SourceTypeID) that are used to describe the 13 different vehicle types used in MOVES. Table 4-3 defines the AgeID variable in terms of the age of the vehicle in years. Table 4-4 defines five RoadTypeIDs. MOVES currently includes RoadTypeID=1 for Off Network roads, however, calculations are not made for that RoadTypeID. MOVES calculations are made for RoadTypeIDs 2, 3, 4, and 5. Table 4-5 shows the definitions for the HourDayID variable, which distinguishes the 24 hours in a day but which distinguishes only weekdays from weekend days. Table 4-6 defines AvgSpeedBinIDs in terms of speed midpoints at 0, 5, 10, 15, …, 70, and greater than 75 miles per hour.

Table 4-2. Definitions of MOVES SourceTypeIDs and HPMSVtypeIDs

HPMSVtypeID	SourceTypeID	Description of Vehicle Type
10	11	Motorcycle
20	21	Passenger Car
30	31	Passenger Truck
	32	Light Commercial Truck
40	41	Intercity Bus
	42	Transit Bus
	43	School Bus
50	51	Refuse Truck
	52	Single Unit Shorthaul Truck
	53	Single Unit Longhaul Truck
	54	Motorhome
60	61	Combination Shorthaul Truck
	62	Combination Longhaul Truck

[5] "MOVES2010 Highway Vehicle Population and Activity Data," EPA-420-R-10-026, November 2010.

Table 4-3. Definitions of MOVES Vehicle AgeIDs

AgeID	Description of Age
0	0 ≤ age in years < 1
1	1 ≤ age in years < 2
2	2 ≤ age in years < 3
3	3 ≤ age in years < 4
4	4 ≤ age in years < 5
5	5 ≤ age in years < 6
6	6 ≤ age in years < 7
7	7 ≤ age in years < 8
8	8 ≤ age in years < 9
9	9 ≤ age in years < 10
10	10 ≤ age in years < 11
11	11 ≤ age in years < 12
12	12 ≤ age in years < 13
13	13 ≤ age in years < 14
14	14 ≤ age in years < 15
15	15 ≤ age in years < 16
16	16 ≤ age in years < 17
17	17 ≤ age in years < 18
18	18 ≤ age in years < 19
19	19 ≤ age in years < 20
20	20 ≤ age in years < 21
21	21 ≤ age in years < 22
22	22 ≤ age in years < 23
23	23 ≤ age in years < 24
24	24 ≤ age in years < 25
25	25 ≤ age in years < 26
26	26 ≤ age in years < 27
27	27 ≤ age in years < 28
28	28 ≤ age in years < 29
29	29 ≤ age in years < 30
30	30 ≤ age in years

Table 4-4. Definitions of MOVES RoadTypeIDs

RoadTypeID	Description of Road Type
1	Off Network
2	Rural Restricted
3	Rural Unrestricted
4	Urban Restricted
5	Urban Unrestricted

Table 4-5. Definitions of MOVES HourDayIDs

HourDayID	Description of Day	Description of Hour
12	Weekend (SAT, SUN)	00:00:00 ≤ local time < 01:00:00
22	Weekend (SAT, SUN)	01:00:00 ≤ local time < 02:00:00
32	Weekend (SAT, SUN)	02:00:00 ≤ local time < 03:00:00
42	Weekend (SAT, SUN)	03:00:00 ≤ local time < 04:00:00
52	Weekend (SAT, SUN)	04:00:00 ≤ local time < 05:00:00
62	Weekend (SAT, SUN)	05:00:00 ≤ local time < 06:00:00
72	Weekend (SAT, SUN)	06:00:00 ≤ local time < 07:00:00
82	Weekend (SAT, SUN)	07:00:00 ≤ local time < 08:00:00
92	Weekend (SAT, SUN)	08:00:00 ≤ local time < 09:00:00
102	Weekend (SAT, SUN)	09:00:00 ≤ local time < 10:00:00
112	Weekend (SAT, SUN)	10:00:00 ≤ local time < 11:00:00
122	Weekend (SAT, SUN)	11:00:00 ≤ local time < 12:00:00
132	Weekend (SAT, SUN)	12:00:00 ≤ local time < 13:00:00
142	Weekend (SAT, SUN)	13:00:00 ≤ local time < 14:00:00
152	Weekend (SAT, SUN)	14:00:00 ≤ local time < 15:00:00
162	Weekend (SAT, SUN)	15:00:00 ≤ local time < 16:00:00
172	Weekend (SAT, SUN)	16:00:00 ≤ local time < 17:00:00
182	Weekend (SAT, SUN)	17:00:00 ≤ local time < 18:00:00
192	Weekend (SAT, SUN)	18:00:00 ≤ local time < 19:00:00
202	Weekend (SAT, SUN)	19:00:00 ≤ local time < 20:00:00
212	Weekend (SAT, SUN)	20:00:00 ≤ local time < 21:00:00
222	Weekend (SAT, SUN)	21:00:00 ≤ local time < 22:00:00
232	Weekend (SAT, SUN)	22:00:00 ≤ local time < 23:00:00
242	Weekend (SAT, SUN)	23:00:00 ≤ local time < 00:00:00
15	Weekday (MON-FRI)	00:00:00 ≤ local time < 01:00:00
25	Weekday (MON-FRI)	01:00:00 ≤ local time < 02:00:00
35	Weekday (MON-FRI)	02:00:00 ≤ local time < 03:00:00
45	Weekday (MON-FRI)	03:00:00 ≤ local time < 04:00:00
55	Weekday (MON-FRI)	04:00:00 ≤ local time < 05:00:00
65	Weekday (MON-FRI)	05:00:00 ≤ local time < 06:00:00
75	Weekday (MON-FRI)	06:00:00 ≤ local time < 07:00:00
85	Weekday (MON-FRI)	07:00:00 ≤ local time < 08:00:00
95	Weekday (MON-FRI)	08:00:00 ≤ local time < 09:00:00
105	Weekday (MON-FRI)	09:00:00 ≤ local time < 10:00:00
115	Weekday (MON-FRI)	10:00:00 ≤ local time < 11:00:00
125	Weekday (MON-FRI)	11:00:00 ≤ local time < 12:00:00
135	Weekday (MON-FRI)	12:00:00 ≤ local time < 13:00:00
145	Weekday (MON-FRI)	13:00:00 ≤ local time < 14:00:00
155	Weekday (MON-FRI)	14:00:00 ≤ local time < 15:00:00
165	Weekday (MON-FRI)	15:00:00 ≤ local time < 16:00:00
175	Weekday (MON-FRI)	16:00:00 ≤ local time < 17:00:00
185	Weekday (MON-FRI)	17:00:00 ≤ local time < 18:00:00
195	Weekday (MON-FRI)	18:00:00 ≤ local time < 19:00:00
205	Weekday (MON-FRI)	19:00:00 ≤ local time < 20:00:00
215	Weekday (MON-FRI)	20:00:00 ≤ local time < 21:00:00
225	Weekday (MON-FRI)	21:00:00 ≤ local time < 22:00:00
235	Weekday (MON-FRI)	22:00:00 ≤ local time < 23:00:00
245	Weekday (MON-FRI)	23:00:00 ≤ local time < 00:00:00

Table 4-6. Definitions of MOVES AvgSpeedBinIDs

AvgSpeedBinID	Description of Speed Bin
1	0 ≤ speed in mph < 2.5
2	2.5 ≤ speed in mph < 7.5
3	7.5 ≤ speed in mph < 12.5
4	12.5 ≤ speed in mph < 17.5
5	17.5 ≤ speed in mph < 22.5
6	22.5 ≤ speed in mph < 27.5
7	27.5 ≤ speed in mph < 32.5
8	32.5 ≤ speed in mph < 37.5
9	37.5 ≤ speed in mph < 42.5
10	42.5 ≤ speed in mph < 47.5
11	47.5 ≤ speed in mph < 52.5
12	52.5 ≤ speed in mph < 57.5
13	57.5 ≤ speed in mph < 62.5
14	62.5 ≤ speed in mph < 67.5
15	67.5 ≤ speed in mph < 72.5
16	72.5 ≤ speed in mph

The structures of the Primary5 tables are summarized in Table 4-7. The third column "Independent Variables that define each Combination" show the variables for which the "Response Variable" in the fifth column is provided. The tables for SourceTypeAgeDistribution, RoadTypeDistribution, and AvgSpeedDistribution have response variables that characterize the relative distribution of a quantity. For example, for the SourceTypeAgeDistribution table, the age distribution for each SourceTypeID is defined by a set of 31 AgeFractions – one fraction for each AgeID – that sum to one for each SourceTypeID. The tables for SourceTypeYear and HPMSVtypeYear are also distributions, but the response variables are absolute values rather than relative fractions that sum to one. For the analysis the absolute values that were submitted to counties for the SourceTypeYear and HPMSVtypeYear tables will be converted to relative values. This will be discussed later.

A few other variables appear in the Primary5 MOVES tables, but they are not being perturbed in this study. ZoneID, which appears in all five Primary5 tables, is basically a label for county. The values submitted for different counties will be examined to determine the range of submitted values to be perturbed for the effects of each of the Primary5 tables. YearID, which appears in the first three of the Primary5 tables, designates the calendar year for a MOVES run. In this study YearID is being held constant at 2011. The SalesGrowth factor and MigrationRate in the SourceTypeYear table and VMTGrowthFactor in the HPMSVtypeYear table are not being used since the runs in this study will be performed only for the 2011 calendar year.

BaseYearOffNetVMT in the HPMSVtypeYear table is not being used since off-network operation is not supported for this version of MOVES.

Table 4-7. Structure of Selected (Primary5) MOVES Tables

MOVES Table	Data Description (Data Uses)	Independent Variables that define each Combination (levels)	Fractionating Variable for each Combination (levels)	Response Variables	Response Variable Notes
SourceTypeAge Distribution	Fraction of vehicles in each age bin.	SourceTypeID (13)	AgeID (31)	AgeFraction	Sums to 1 for each SourceTypeID.
SourceTypeYear	Number of vehicles for each source type.	SourceTypeID (13)	None	SourceTypePopulation	
HPMSVtypeYear	Amount of VMT for each HPMS vehicle type.	HPMSVtypeID (6)	None	HPMSBaseYearVMT	
RoadType Distribution	Fraction of VMT driven on each road type.	SourceTypeID (13)	RoadTypeID (4*)	RoadTypeVMTFraction	Sums to 1 for each SourceTypeID.
AvgSpeed Distribution	Fraction of driving time spent in each speed bin.	SourceTypeID (13) RoadTypeID (4*) HourDayID (24weekend, 24weekday)	AvgSpeedBinID (16)	AvgSpeedFraction	Sums to 1 for each combination of SourceTypeID, RoadTypeID, and HourDayID.

The following specifically defines the variables contained in each of the Primary5 tables:

- **SourceTypeAgeDistribution Table** – Columns: ZoneID (county), SourceTypeID, YearID (not used for this study since all runs are for the same year, 2011), AgeID, AgeFraction.

- **SourceTypeYear Table** – Columns: ZoneID (county), YearID (not used for our single-year analysis), SourceTypeID, SalesGrowthFactor (not used for our single-year analysis), SourceTypePopulation, MigrationRate (not used for our single year analysis). For each county, the table contains the number of vehicles in the county's fleet for each SourceTypeID.

- **HPMSVtypeYear Table** – Columns: ZoneID (county), HPMSVtypeID, YearID (not used for our single-year analysis), VMTGrowthFactor (not used for our single-year analysis), HPMSBaseYearVMT, BaseYearOffNetVMT (not used, since off-network roads are not considered in this study). For each county, the table contains the VMT (miles per year) for each of the six HPMSVtypeIDs.

- **RoadTypeDistribution Table** – Columns: ZoneID (county), SourceTypeID, RoadTypeID, RoadTypeVMTFraction. The table is stratified by the 13 SourceTypeIDs. The RoadTypeVMTFraction field contains the fractional portion of the VMT on each of 4 RoadTypeIDs.

- **AvgSpeedDistribution Table** – Columns: ZoneID (county), SourceTypeID, RoadTypeID, HourDayID, AvgSpeedBinID, AvgSpeedFraction. The table is stratified by the 13 SourceTypeIDs, 4 RoadTypeIDs, and 48 HourDayIDs. The AvgSpeedFraction field contains the fractional portion of the driving time in each of 16 AvgSpeedBinIDs.

4.3 Analysis Approach for Task 1b

The goal of this task was to compare the state-submitted data for the Primary5 tables by comparison of inputs across states, and by comparison with the MOVES defaults that would be applied in states that did not submit data. The primary tool for this approach was a set of charts to give a graphical representation of variation in the state submittals.

For the Primary5 tables, there were three basic types of dataset structures. Each of the three required a somewhat different approach to the analysis. The SourceTypeYear and HPMSVtypeYear datasets each contain absolute numbers for vehicle population and VMT, by SourceTypeID, respectively. The primary analysis for these two tables was a comparison of the submitted numbers to the default numbers. The AvgSpeedDistribution and SourceTypeAgeDistribution tables each contain a continuous distribution – the fractional distribution of driving into each of 16 SpeedBinIDs, or the fractional distribution of the population at each vehicle AgeID with values from 0 to 30. For these two tables, a mean SpeedBinID value and a mean AgeID value were calculated and compared. However, the mean does not always include enough information to compare distributions – for example, two distributions may have the same mean, but one may have a large number of very high and a large number of very low values, while another may have all values near the mean. Even though the means are the same, the differences between the distributions could have an effect on emissions. Since a measure to compare the distributions as well as the means is useful, the vector angle between each county's submittal and the defaults was calculated (see below for a discussion of vector angles). Finally, the RoadTypeDistribution table contains a distribution of the fractional driving on each of four RoadTypeIDs, by SourceTypeID. For these non-continuous distributions, where there is no actual numeric relationship between the different road types, calculating the mean RoadTypeID is not very meaningful, and the calculation of the vector angles becomes an important tool.

In the sub-sections below, the concept of calculating vector angles is discussed. The result of comparing submittals for each of the Primary5 tables follows.

4.4 Cluster Analysis of Speed Distribution Data

As shown in Table 4-7, the AvgSpeedDistribution table requires a distribution of speeds for each combination of SourceTypeID, RoadTypeID, and HourDayID. Since there are 13 SourceTypeIDs, 4 RoadTypeIDs, and 48 HourDayIDs, 2,496 distributions need to be submitted for each county. More important for this project, comparisons of 2,496 speed distributions among the counties for Task 1b and the quantification of changes in 2,496 speed distributions on emissions for Task 2 would be huge tasks. Clearly, some simplifying assumptions need to be made to make the analysis of the submitted AvgSpeedDistribution tables more reasonable.

We begin by assuming that the speed distributions of different SourceTypeIDs on the same RoadTypeIDs at the same HourDayIDs will be the same. All vehicles will tend to have the same speeds under similar conditions because all are subject to the same speed limits and generally follow each other on roadways. Next, the speed distributions for adjacent HourDayIDs will tend to be similar for each RoadTypeID. For a given RoadTypeID, speed distributions will change smoothly over time. Therefore, it is likely that for each RoadTypeID the 48 HourDayIDs can be grouped into clusters that have similar speed distributions. Overall, by assuming that the 13 SourceTypeIDs have the same speed distributions and by clustering combinations of RoadTypeIDs and HourDayIDs, it may be possible to produce a relatively small number of speed distribution clusters (SpeedDistClusters) that have similar speed distributions within a cluster and different speed distributions between clusters.

To discover how to cluster the combinations of RoadTypeIDs and HourDayIDs, a set of speed data, that represents speeds driven across the U.S. and that is independent of the distributions submitted for the counties, is needed. The TomTom dataset, which was purchased from TomTom, is available from a recent EPA project. The following[6] summarizes the data:

> Vehicle in-use data is collected by TomTom, which manufactures and sells portable GPS units as well as an iPhone application. Some users of TomTom units give permission to TomTom to collect and store users' personal (anonymous) data on TomTom servers. The data is collected while the GPS unit is on, either in map

[6] "MOVES Activity, VMT, and Population Update," Work Plan, Version 3, prepared for U.S. Environmental Protection Agency, prepared by Eastern Research Group, EPA-121019, October 19, 2012.

or navigation mode. As long as the device is turned on, it is gathering data. A unit's GPS tracks are delivered to TomTom servers when data is collected either over the cell network as a "live" feed or as a "non-live" stream of data when the user connects to receive software or map updates.

Data collection began in January 2008. The data has been collected continuously, and the database currently has over 1 trillion data points. Since all U.S. drivers do not use a TomTom GPS unit or app and users who "opt in" are self-selective, biases could exist in the data that is collected. Anecdotally, drivers that own GPS units are less likely to use them when they drive in familiar areas in comparison with unfamiliar areas. TomTom data is obtained from units on all road types but at this time the data does not distinguish source types. Since these are portable devices, they are not able to capture vehicle information. TomTom suspects that "virtually all" of their vehicles are light-duty cars, trucks, and vans. TomTom data is obtained from all areas of North America where vehicles with TomTom GPS units drive. There are some areas where their data counts are low, but for any reasonably sized city they have an "excellent quantity of data."

Because TomTom units are not in all vehicles or even in a random fraction of all vehicles, the data cannot be used to determine absolute VMT. However, the TomTom data can be used to estimate light-duty vehicle speed distributions since some vehicles in traffic will be sending their speed data to TomTom.

The MOVES RoadTypeIDs were determined from the TomTom roadway segment classification. The determination of urban vs. rural for each roadway segment was made using census information so that the categorization is consistent with the definition used by MOVES. TomTom queried its database of historic traffic probes to produce a table of total distance and total time as a function of road type, month, weekday/weekend, hour of the day, and average speed bin for the calendar year 2011 for the 50 U.S. states and the District of Columbia.

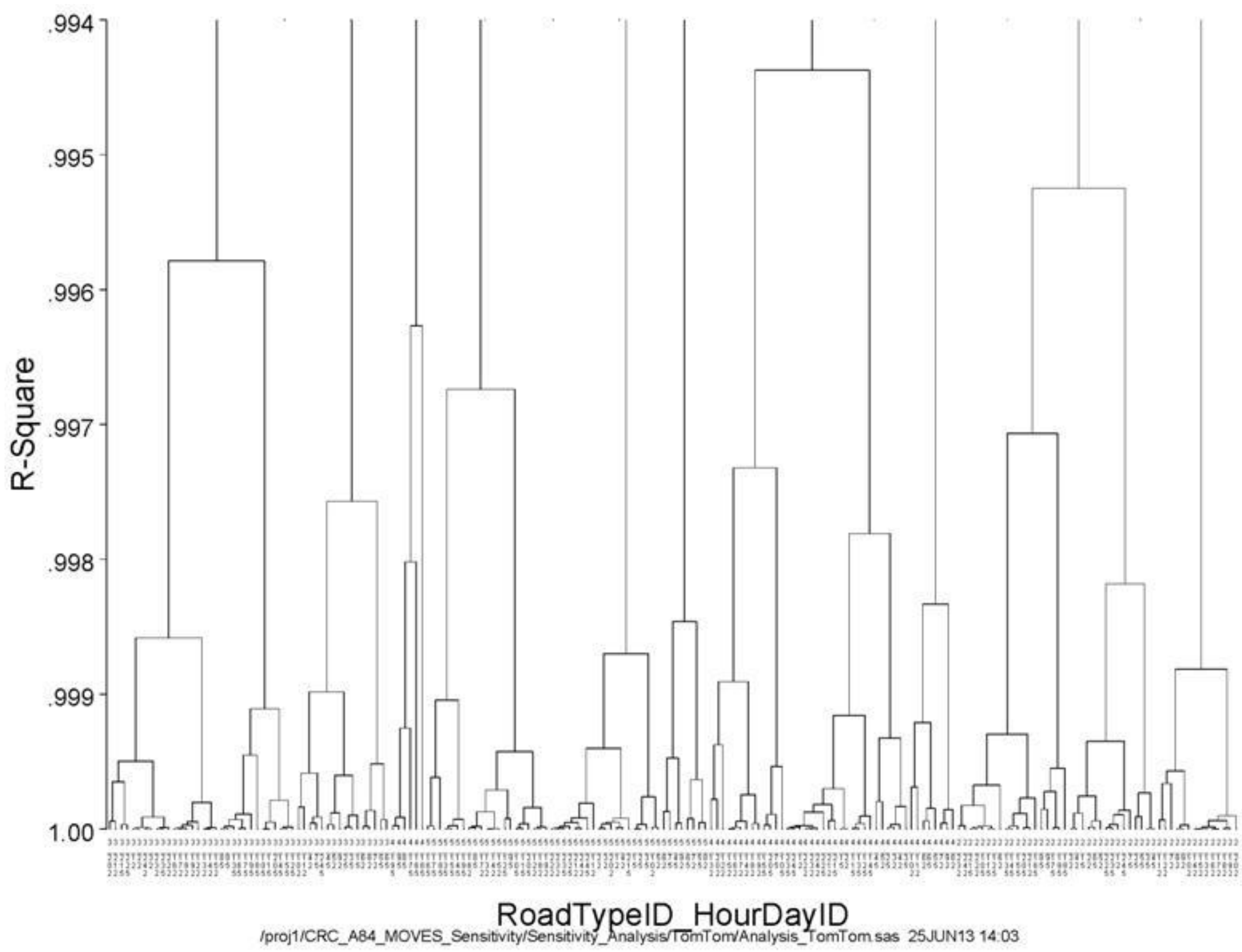
In that study the analysis of the TomTom data produced cumulative speed distributions as a function of AvgSpeedBinID for each of the 192 combinations of RoadTypeID and HourDayID. For this study, while making the assumption that all SourceTypeIDs have the same speed distributions, we want to find a small number of SpeedDistClusters where the speed distributions within each cluster are similar and the speed distributions between clusters are different.

For this analysis[7] we used the CLUSTER procedure in SAS using the Ward method. The cluster analysis produced the hierarchical tree shown in Figure 4-1. The cluster analysis begins with each of the 192 combinations of RoadTypeID and HourDayID in its own cluster. The procedure then combines the two combinations that are "closest" to each other, which results in 191 clusters. The closeness is calculated as the distance in the 16-dimensional space defined by

[7] /proj1/CRC_A84_MOVES_Sensitivity/Sensitivity_Analysis/TomTom/Analysis_TomTom.sas

the cumulative distribution of AvgSpeedFractions in terms of AvgSpeedBinIDs. The clustering is then iterated until all of the combinations are in one cluster. Along the way, several statistics are calculated, including the r-square. At the beginning the r-square is 1 since the clusters describe the variability of the 192 combinations exactly. As the clustering proceeds, the r-square drops. Based on an examination of the statistics, we chose to use 13 clusters to describe the 192 combinations of RoadTypeID and HourDayID. This clustering had an r-square of 0.996 associated with it. The 13 clusters are seen in Figure 4-1 by the RoadTypeID_HourDayID values at the bottom of the figure for the 13 clusters defined by the vertical lines that cross at an imaginary horizontal line where $r^2 = 0.996$.

Figure 4-1. Hierarchical Tree for Speed Distribution Cluster Analysis



When the combinations of RoadTypeIDs and HourDayIDs for each of the 13 SpeedDistClusters occur during the week are indicated by the colored bars in Figures 4-2 and 4-3 for weekend days and weekday days, respectively. The figures show that combinations within each cluster tend to be contiguous. For example, all combinations for SpeedDistCluster 13 occur from Hour 7 through Hour 20 on the weekends for rural restricted roadways. SpeedDistCluster 13 occurs exclusively on weekends, SpeedDistClusters 2 and 4 occur exclusively on weekdays, and the remaining 10 SpeedDistClusters occur on both weekends and weekdays.

A weekday vs. weekend comparison of the SpeedDistClusters in Figures 4-2 and 4-3 is instructive. For example, in Figure 4-3, SpeedDistClusters 9, 10, 8, and 4 for Urban Restricted roadways on weekdays clearly show the progression from the off-peak speeds of SpeedDistCluster 9 to the congested speeds of SpeedDistCluster 4. In contrast, Figure 4-2 for Urban Restricted roadways on weekends shows a more broad progression to daytime traffic, but the congested speeds of SpeedDistCluster 4 are not generally observed on weekends.

The average distribution of speeds for the 13 SpeedDistClusters are observed in the TomTom data and as submitted for the counties are compared in Figures 4-4 and 4-5, respectively. Overall, the speed distributions obtained from the TomTom data are more widely distributed across the AvgSpeedBinIDs than the corresponding speed distributions submitted for counties are. The rural speed distributions are indicated with dashed lines in these figures.

Both figures show four groups of SpeedDistClusters that have similar speed distributions. The Urban Unrestricted speed distributions (SpeedDistClusters 5, 6, 7) are close to each other and overall have the lowest speeds. The Rural Unrestricted speed distributions (SpeedDistClusters 1, 2, 3) are close to each other and overall have the next higher speeds. In Figure 4-5 these distributions are almost on top of each other. The Rural Restricted speed distributions (SpeedDistClusters 11, 12, 13) are close to each other and overall have the highest speeds. In Figure 4-5 these distributions are almost on top of each other. The remaining four clusters are for the Urban Restricted speed distributions (SpeedDistClusters 4, 8, 9, 10). In Figure 4-5 for the county-submitted data, the Urban Restricted speed distributions are relatively close to each other with SpeedDistCluster 4 having the lowest speed and SpeedDistCluster 9 having the highest speeds. On the other hand, the TomTom data in Figure 4-4 indicates that SpeedDistCluster 4, which is for weekday rush hour periods, has substantially lower speeds than for the other three Urban Restricted speed distributions (SpeedDistClusters 8, 9, 10), and unlike in Figure 4-5, the highest speed distribution is for SpeedDistCluster 10 among the Urban Restricted speed distributions.

Figure 4-2. RoadTypeID and HourDayID of the 13 SpeedDistClusters for Weekend Days

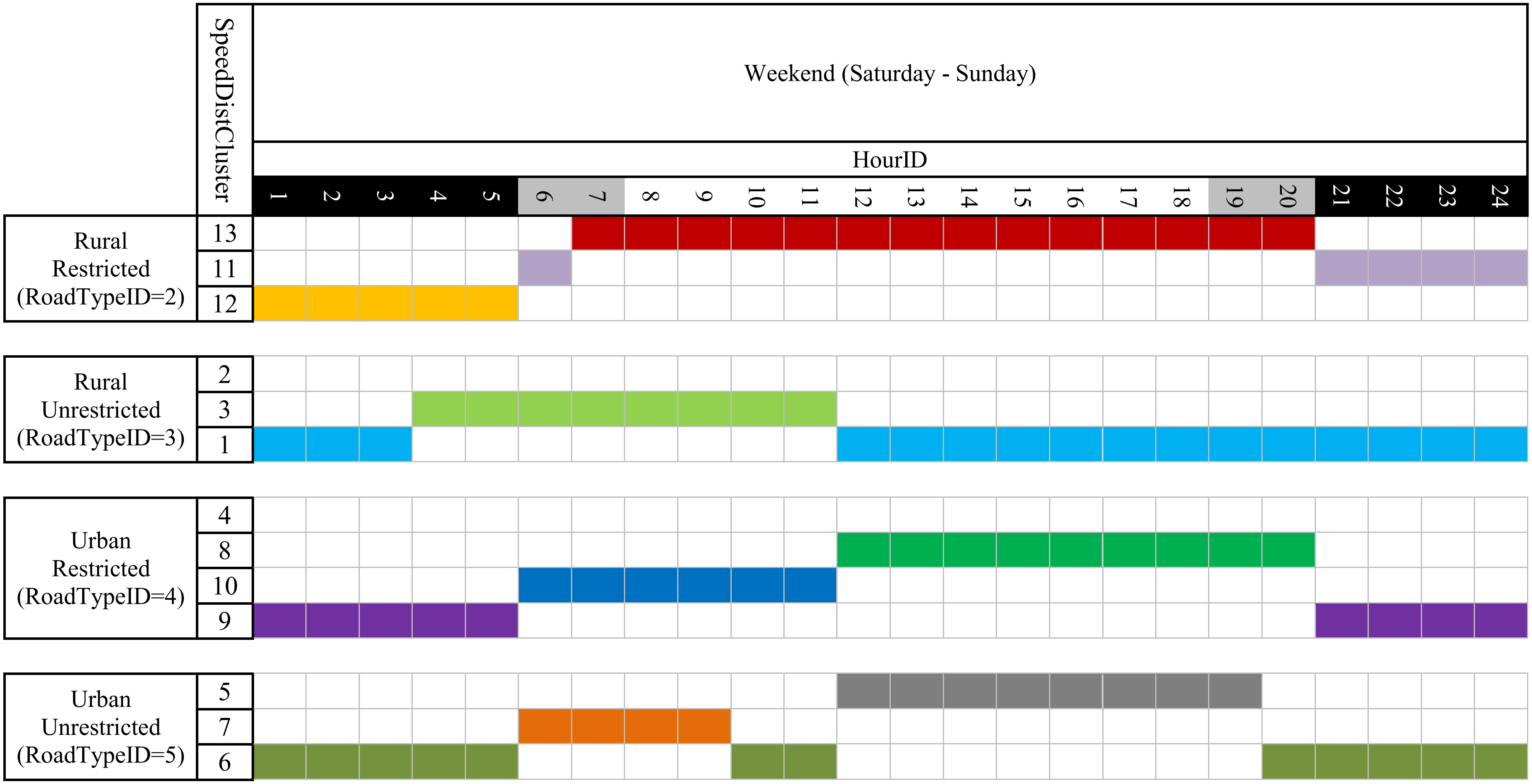


Figure 4-3. RoadTypeID and HourDayID of the 13 SpeedDistClusters for Weekday Days

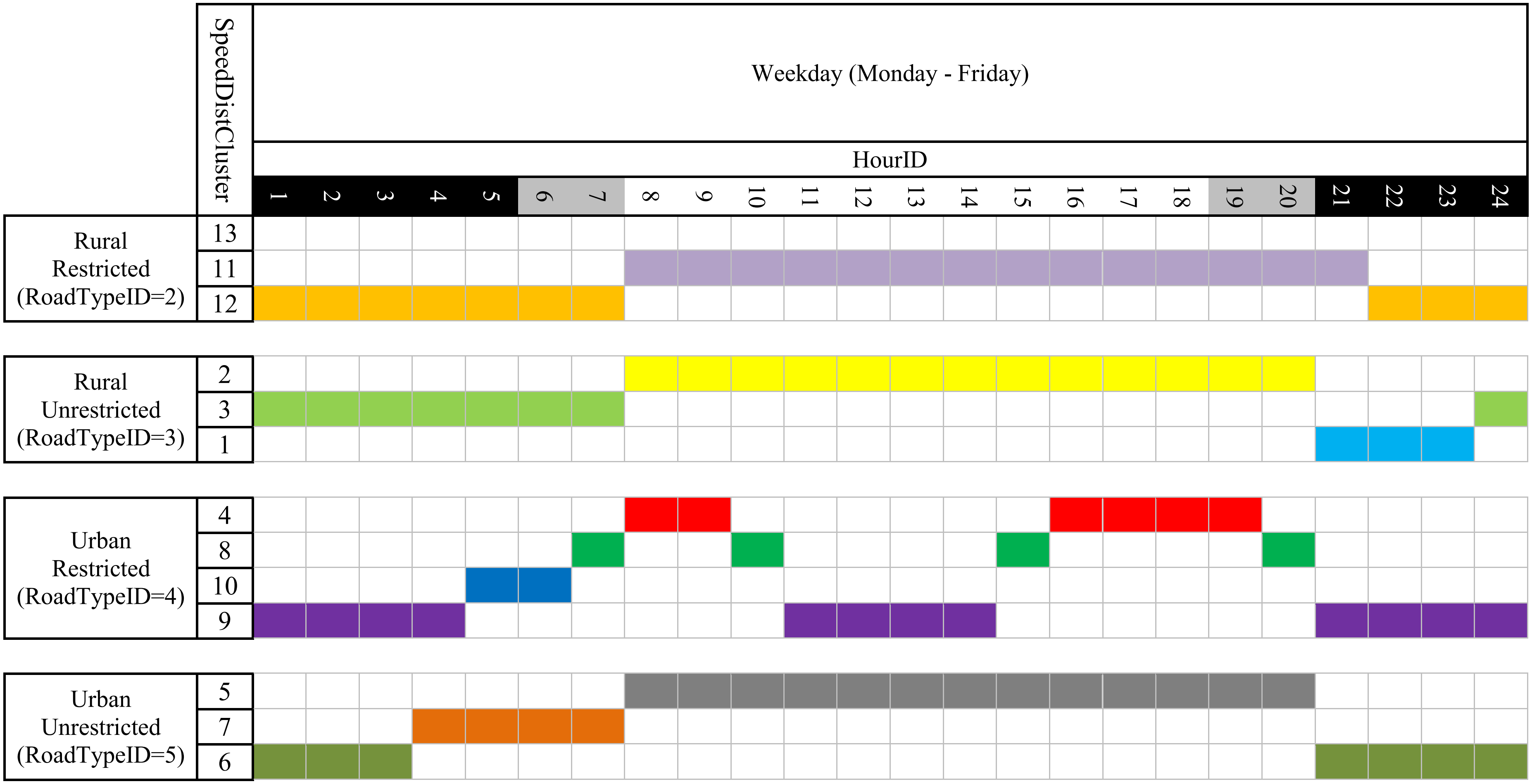


Figure 4-4. Average Cumulative Speed Distributions for the 13 SpeedDistClusters Using TomTom Data

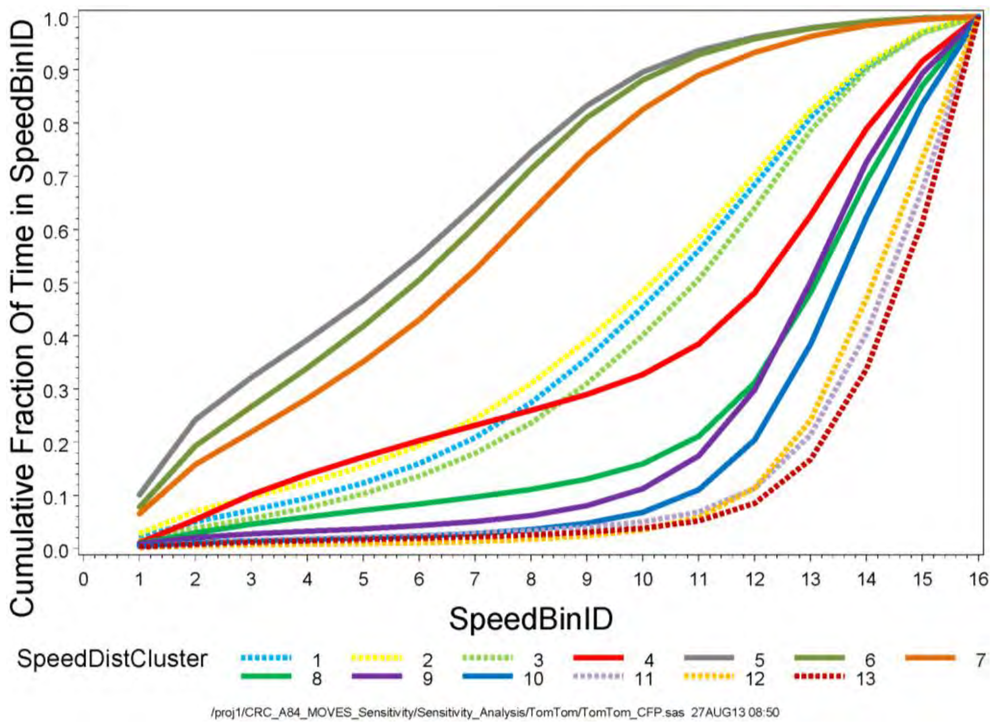
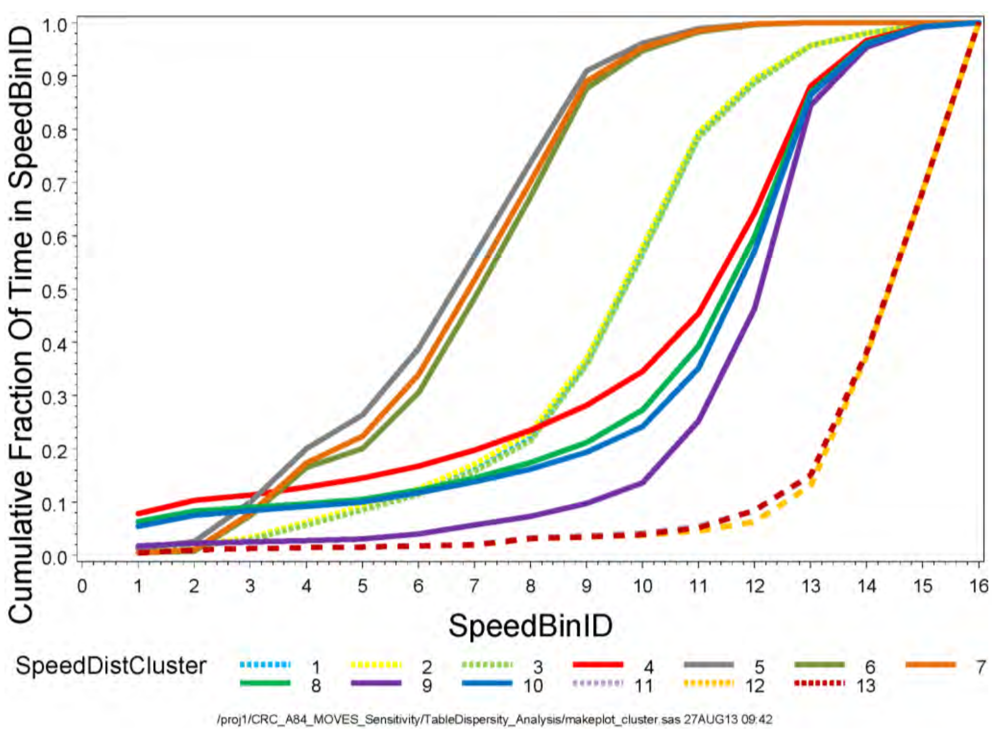


Figure 4-5. Average Cumulative Speed Distributions for the 13 SpeedDistClusters Using County-Submitted Data



4.5 Data Preparation

The dataset for each of the Primary5 tables was prepared separately, although many of the operations were performed in parallel on each of the tables. The steps prior to analysis were:

QA/QC

- Read into SAS the five raw data tables as output from MOVES[8].
- Read in the five raw defaults tables as output from MOVES and merge to the data tables.
- The data in three tables (SourceTypeAgeDistribution, RoadTypeDistribution, AvgSpeedDistribution) is formatted as fractions that should sum to 1 for a given SourceTypeID or RoadTypeID. When the sum did not equal 1, that SourceTypeID or RoadTypeID was deleted from the table.
- Some counties in Texas submitted numbers for population (in SourceTypeYear) or VMT (in HPMSVtypeYear) that were several orders of magnitude greater than the maximum numbers submitted by any other state. These were removed from the tables.
- Two states and one county provided incomplete submittals that would not contain the information needed for this study. All records for Oregon, Ohio, and Knox County of Tennessee were removed from all of the Primary5 tables.

Calculations

- An "average county" was defined for each table, using the median values of each data table[9].
- Vector angles (see Section 4.6) with respect to the MOVES defaults and with respect to the average county were calculated.
- Means were calculated for the two tables with continuous distributions: SourceTypeAgeDistribution (mean AgeID), AvgSpeedDistribution (mean SpeedBinID).
- Fractions were calculated for the two tables with absolute numbers: population fractions were calculated for SourceTypeYear, and VMT fractions were calculated for HPMSVtypeYear. For example, the "fraction" for SourceTypeID=21 in a given county was the population of SourceTypeID=21 vehicles divided by the total population of all SourceTypeIDs for that county.

---

[8] /proj1/CRC_A84_MOVES_Sensitivity/TableDispersity_Analysis/read_counties.sas
[9] /proj1/CRC_A84_MOVES_Sensitivity/TableDispersity_Analysis/makevectors_*.sas

- The total population and total VMT were calculated for each county, and the median total population across all counties and median total VMT across all counties were found.

- For the AvgSpeedDistribution table, some anomalous results were removed by excluding any distributions that represented a VMT of 0 for that RoadTypeID or SourceTypeID. Also for that table, any distributions with a mean SpeedBinID fraction equal to 1 were deleted, since these did not appear to represent real speed distributions.

- Also for the AvgSpeedDistribution table, for the purpose of analysis, the data were grouped into 13 SpeedDistClusters as described in Section 4.4. Separate analysis of each of 2,496 speed distributions for each county (13 SourceTypeIDs * 4 RoadTypeIDs * 48 HourDayIDs) would have been computationally overwhelming. Therefore, 13 clusters representing groups of similar speed distributions were created. Each cluster represents driving on a single RoadTypeID and includes several HourDayID values. See the discussion of cluster analysis in Section 4.4.

- For each of the five tables, the percentile ranking for each county's submittal, with respect to all county submittals, was calculated. This was done slightly differently for each of the Primary5 tables, due to differences in the structures of the tables.

  — **SourceTypeAgeDistribution:** The percentiles were calculated based on mean AgeID and were calculated separately for each SourceTypeID.

  — **SourceTypeYear:** The percentiles were based on the population fractions and were calculated separately for each SourceTypeID.

  — **HPMSVtypeYear:** The percentiles were based on the VMT fractions and were calculated separately for each HPMSVtypeID.

  — **RoadTypeDistribution:** The percentiles were based on the vector angle between each county and the "average county's" road type distribution and were calculated separately for each SourceTypeID.

  — **AvgSpeedDistribution:** The percentiles were based on the average SpeedBinID for a given SpeedDistCluster and were calculated separately for each SourceTypeID.

- When the distributions for SourceTypeAgeDistribution and AvgSpeedDistribution were ordered by percentile, long sequences of duplicate distributions were found. Often, these appear to have resulted from one state submitting the same distribution for a number of counties, or from one state submitting the MOVES default distribution for a number of counties. Duplicates were removed from the dataset.

- For the SourceTypeAgeDistribution and AvgSpeedDistribution tables, at each of the 10, 90, and 50 percentiles, three different county distributions were plotted and compared. The first of the three was automatically used as the distribution to represent that percentile, unless the plot showed it to be problematic, in which case the second or third of the plotted distributions was chosen instead. Problematic distributions included those distributions having with only one or two bars instead of a range or bars or those where all bars were at the same level.

4.6 Calculation of Vector Angles

A given distribution of fractional values can be thought of as a vector, with the number of elements determined by the distribution under consideration. In this study, that means 4 elements for the four RoadTypeIDs in RoadTypeDistribution, or 16 elements for the 16 SpeedBinIDs in AvgSpeedDistribution, or 31 elements for the 31 AgeIDs in SourceTypeAgeDistribution. If a MOVES default distribution vector is compared to a county-submitted distribution vector, and if the two distributions are not identical, then there will be a non-zero angle between the two vectors. The greater the difference between the two vectors, the greater the angle between them will be. Unlike a simple comparison of the means of two different distributions, which compares one number to one other number, the use of vector space allows each element in a given distribution to be compared, and the amount of difference between all of the elements summarized in a resultant angle.

Calculation of the angle between two vectors uses the vector dot product and the law of cosines, as shown in Equation 1:

$$D \cdot S = \sum_{i=1}^{n} D_i S_i = \|D\|\|S\| cos\theta$$

In other words, the angle θ can be found by calculating the dot product between the vector D (default) and S (submitted), and then dividing by the product of the magnitudes (i.e., the lengths) of the two vectors.

For example, consider a comparison of the age distribution for Fairfax, Virginia and the national MOVES default, shown in Table 4-8. The first column of the table shows the age; Columns 2 and 3 of the table show the fractional portion of the fleet in each of the 31 AgeID bins. Columns 4 and 5 show the same information, but the fractions are now cumulative. Comparing the cumulative fractions allows the vector comparison to take into account the proximity of two cells in the distribution. Including the effect of proximity is relevant for continuous distributions such as the age distribution or speed distribution; it is not used for non-continuous distributions like the road type distribution. The dot product of the D and S vectors is the sum of the products of each cell of the cumulative distributions; the individual products are shown in Column 6, and the sum at the bottom of the table. The magnitude of a vector is calculated as the square root of the sum of the squares of each cell. The individual squares for the default and submitted vectors are shown in Columns 7 and 8, and the sums at the bottom of the table. From these calculations, using the equation above, we find that the angle between the two vectors is 1.3 degrees (=invcos(19.988/(sqrt(19.656)*sqrt(20.336)))). To complete the process, the angle must be calculated a second time, this time with the accumulations done in descending order, instead of ascending as shown in the table. This is done to “average out” the end effects at the beginning and end of the distribution introduced by the accumulation process. This gives an angle of 2.6 degrees. The final result is the average of the two angles: 1.9 degrees between the default and submitted vectors.

SourceTypeAgeDistribution Table

For the SourceTypeAgeDistribution table, the mean AgeID was calculated for each county’s submittal of age distributions, separately for each SourceTypeID. The angle between the MOVES defaults and the county’s submittal was also calculated. In Figures A-1 through A-13 in Appendix A, the angles and the mean AgeIDs are plotted against each other, for each SourceTypeID. On each plot, the default mean AgeID is shown with a red dot, at an angle of 0 (with respect to itself). The submitted data are shown with black plus symbols, and it can be seen that as the mean AgeID moves above or below the default, the vector angle increases. Additionally, an “average” county was calculated, at the median of the submitted county data, and it is shown with a green dot.

Figure A-1 shows that the default age distribution for motorcycles is a much lower age distribution than that of any county’s submittal. Figure A-2 shows that the default age distribution is near the low end for passenger cars, but that there are some counties below the defaults, and many counties above them. For some SourceTypeIDs, the default distribution and the average distribution are very similar, with the mean AgeIDs close together, and the angle between them relatively small.

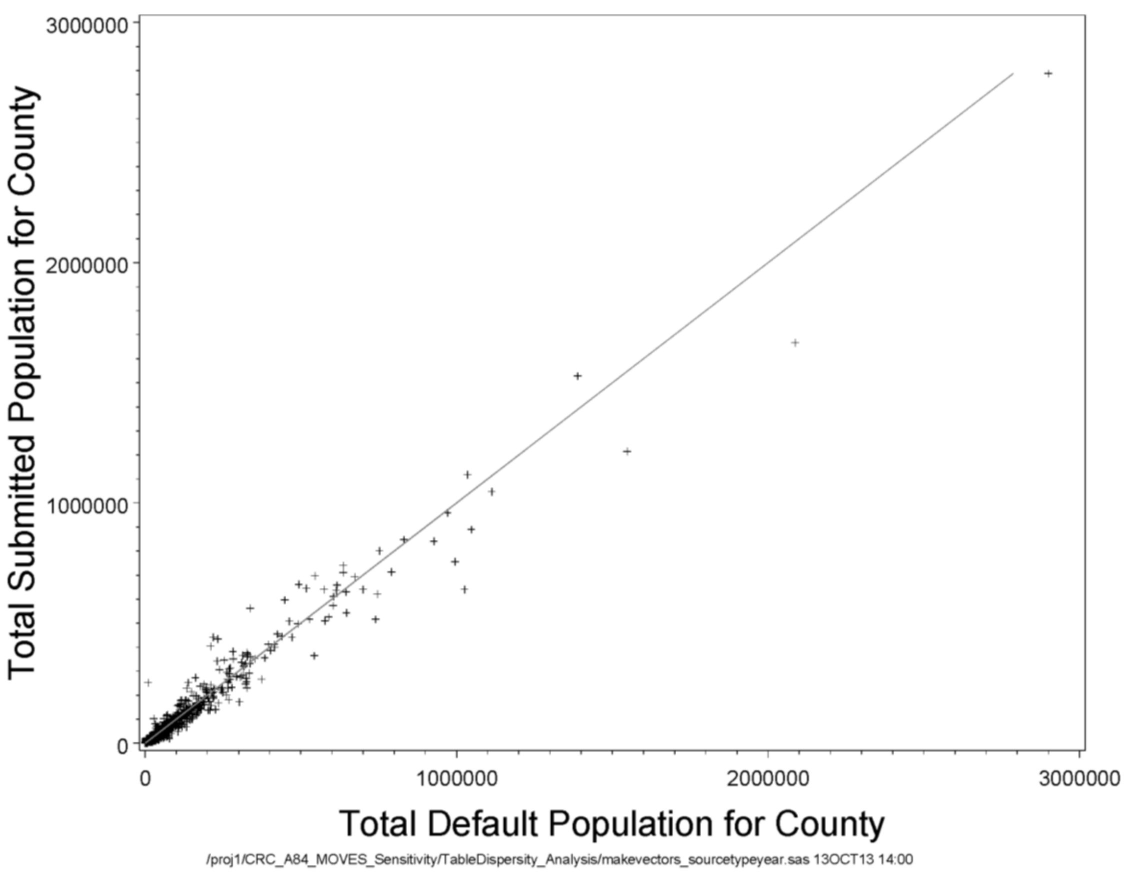
Table 4-8. Comparison of Age Distributions: MOVES Defaults and Submitted Data for Fairfax, Virginia

1	2	3	4	5	6	7	8
AgeID	**MOVES Default AgeFraction**	**Submitted AgeFraction**	**D Default accumulated up**	**S Submitted accumulated up**	**D*S**	**D*D**	**S*S**
0	0.071	0.057	0.071	0.057	0.004	0.005	0.003
1	0.062	0.077	0.133	0.134	0.018	0.018	0.018
2	0.052	0.059	0.186	0.193	0.036	0.034	0.037
3	0.058	0.064	0.244	0.257	0.063	0.059	0.066
4	0.064	0.075	0.308	0.333	0.102	0.095	0.111
5	0.065	0.070	0.372	0.402	0.150	0.139	0.162
6	0.063	0.066	0.435	0.468	0.204	0.189	0.219
7	0.060	0.065	0.495	0.533	0.264	0.245	0.284
8	0.059	0.067	0.554	0.600	0.332	0.307	0.360
9	0.060	0.062	0.614	0.661	0.406	0.377	0.437
10	0.060	0.056	0.674	0.717	0.483	0.455	0.514
11	0.060	0.056	0.734	0.773	0.568	0.539	0.598
12	0.052	0.046	0.787	0.819	0.644	0.619	0.671
13	0.042	0.038	0.828	0.857	0.710	0.686	0.734
14	0.035	0.031	0.864	0.888	0.767	0.746	0.789
15	0.030	0.024	0.893	0.912	0.814	0.798	0.831
16	0.024	0.022	0.917	0.934	0.856	0.841	0.872
17	0.020	0.016	0.937	0.950	0.890	0.878	0.903
18	0.015	0.012	0.951	0.962	0.916	0.905	0.926
19	0.011	0.009	0.962	0.972	0.935	0.926	0.944
20	0.008	0.007	0.970	0.979	0.950	0.941	0.959
21	0.007	0.006	0.977	0.985	0.963	0.955	0.970
22	0.006	0.004	0.983	0.989	0.972	0.967	0.978
23	0.005	0.003	0.988	0.991	0.979	0.976	0.983
24	0.003	0.002	0.991	0.994	0.985	0.983	0.988
25	0.003	0.002	0.994	0.996	0.990	0.988	0.991
26	0.002	0.002	0.996	0.997	0.994	0.993	0.994
27	0.002	0.001	0.998	0.998	0.996	0.996	0.997
28	0.001	0.001	0.999	0.999	0.998	0.998	0.998
29	0.001	0.001	0.999	1.000	0.999	0.999	0.999
30	0.001	0.000	1.000	1.000	1.000	1.000	1.000
Total					19.988	19.656	20.336

SourceTypeYear Table

The SourceTypeYear table contains the number of vehicles in the population for each SourceTypeID. For this table, the mean population for a county (across the SourceTypeIDs within that county) would not be a meaningful value; nor would the vector angle comparing the distribution of the vehicles to the default distribution be very useful. It was decided that for this table, the most useful comparison would be a comparison of each county's submitted population with its default population, separately for each SourceTypeID. The results of these comparisons are shown in Figures B-1 through B-13 in Appendix B. Figure 4-6 shows the same comparison, but now with the total population for each county. It can be seen from the figures that for the most part, the populations fall relatively near the 1:1 line; however, there are notable outliers on the plots, and trends with "rays" of points forming a line above or below the 1:1 line.

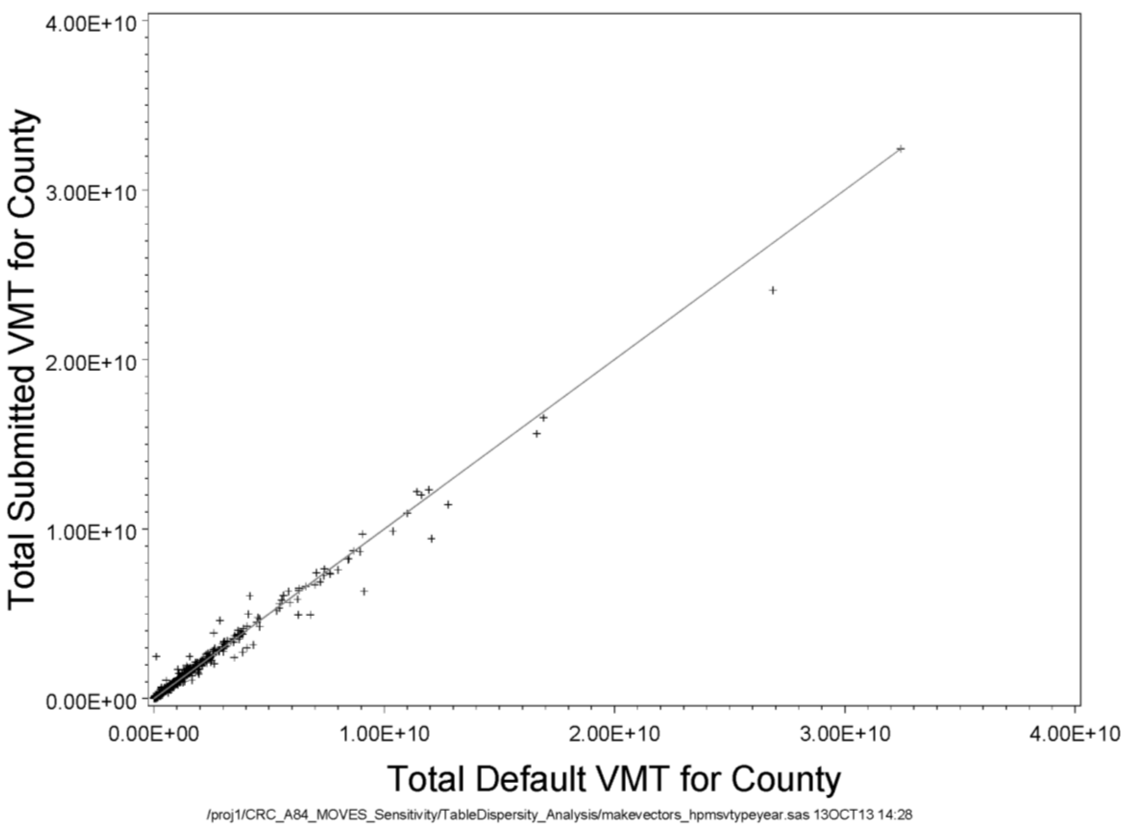
Figure 4-6. Submitted and Default Populations for All SourceTypeIDs



HPMSVtypeYear Table

The HPMSVtypeYear table contains the total annual Vehicle Miles Travelled (VMT) for each of the six HPMSVtypeIDs. This table is similar to the SourceTypeYear table, in that the mean VMT for a county (across the HPMSVtypeIDs within that county) would not be a meaningful value; nor would the vector angle comparing the distribution of the VMT to the default distribution be very useful. It was decided that for this table, the most useful comparison would be a comparison of each county's submitted VMT with its default VMT, separately for each HPMSVtypeID. The results of these comparisons are shown in Figures C-1 through C-13 in Appendix C. Figure 4-7 shows the same comparison, but now with the total VMT for each county. It can be seen from the figures that for the most part, the VMTs fall relatively near the 1:1 line, even more so than for the populations in the SourceTypeYear tables; however, there are notable outliers on the plots, and trends with "rays" of points forming a line above or below the 1:1 line.

Figure 4-7. Submitted and Default VMT for All HPMSVtypeIDs



RoadTypeDistribution Table

For the RoadTypeDistribution table, the distribution of driving among the four RoadTypeIDs was considered to be the most important characteristic, which can be compared using the vector angle between two road type distributions. In this case, since the MOVES defaults are by county, and the goal was to make comparisons across the counties, the angle was calculated between each county's submitted distribution and the distribution for an "average county," found by taking the median distribution across all counties. Since the RoadTypeID numbers (2 through 5) are categorical variables, rather than ordinal variables, the mean RoadTypeID does not contain the same useful meaning that the mean AgeID does in the SourceTypeAgeDistribution table. However, the mean RoadTypeID can give some idea of the degree of shift among RoadTypeIDs, with a lower mean RoadTypeID tending to indicate more rural driving and a higher mean RoadTypeID tending to indicate more urban driving. Therefore, both mean RoadTypeID and the angle between the submitted and average distributions were calculated for each county, separately for each SourceTypeID. In Figures D-1 through D-13 in Appendix D, the angles and the mean RoadTypeIDs are plotted against each other, for each SourceTypeID. On each plot, the "average" road type distribution (which is the median distribution over all the counties, for each SourceTypeID) is shown with a green dot, at an angle of 0 (with respect to itself). The submitted data are shown with black plus symbols, and it can be seen that as the mean RoadTypeID moves above or below the default, the vector angle increases.

AvgSpeedDistribution Table

As discussed in Section 4.4, the large number of distributions of AvgSpeedBinIDs built into the MOVES AvgSpeedDistribution table were collapsed into 13 SpeedDistClusters. Each cluster was selected to include combinations of RoadTypeID and HourDayID with similar speed distributions. The mean AvgSpeedBinID was calculated for each county's submittal of speed distributions, separately for each combination of SpeedDistCluster and SourceTypeID (169 combinations). The angle between the MOVES defaults and the county's submittal was also calculated. In Figures E-1 through E-13 in Appendix E, the angles and the mean AvgSpeedBinIDs are plotted against each other, for each SpeedDistCluster. Rather than show all 169 plots, only the 13 plots for SourceTypeID=21 (Passenger Cars) are shown. On each plot, the default mean AvgSpeedBinID is shown with a red dot, at an angle of 0 (with respect to itself). The submitted data are shown with black plus symbols, and it can be seen that as the mean AvgSpeedBinID moves above or below the default, the vector angle increases. Additionally, an "average" county was calculated, at the median of the submitted county data, and it is shown with a green dot.

It can be seen from the figures that within a cluster, there can be a great variation in mean AvgSpeedBinID and vector angle, both among the submitted data and when comparing the submitted data to the defaults. There are also patterns of similarity between the plots. For example, SpeedDistClusters 11, 12, and 13 are the three clusters that include the driving for RoadTypeID=2, and the distributions of mean AvgSpeedBinIDs and angles are very similar for each of these plots. In other words, the submitted driving distributions show very little change over the hours of the day.

5.0 Task 2: MOVES Sensitivity Analysis

5.1 Approach and Rationale

The goal of almost any sensitivity analysis is to determine the effect of changing one variable on the value of another variable. Sensitivity studies done on actual physical or chemical systems can be expensive and time consuming. However, if a model of the system exists or can be developed, then a sensitivity analysis on the model can be performed relatively inexpensively while providing results that are useful – but only to the degree that the model accurately represents the actual process. In this task, we design and examine the results of a sensitivity analysis on the MOVES model.

Response variables – This sensitivity analysis will investigate influences of the Primary5 tables on four response variables: the "total" fleet emissions for HC, CO, NOx, and PM2.5. "Total" means the sum of the exhaust emissions and evaporative emissions for all SourceTypeIDs combined. As the inputs to the MOVES model are changed, each of the four response variables will change. We will express the change in emissions from a perturbation of input values as a percent change relative to the total emissions for a base case emissions result. A negative change means a decrease in emissions relative to the base case; a positive change means an increase in emissions relative to the base case.

Base case – All MOVES runs in the set of experiments in this study were made for a weekday in July 2011. Thus, the results obtained are relevant only to a weekday in July 2011 and not to weekends, other months of the year, or other calendar years. The base case emissions are defined as the emissions produced for a 30,902 vehicle population and a 397,907,960 annual VMT when all values of the Primary5 tables are at their median values and the non-Primary5 tables are at constant values.

Tables to be Perturbed – The sensitivity analysis will be an analysis of a few "experiments." Each experiment will be performed by making a set of MOVES runs in which one of the MOVES tables has been perturbed while all other tables have been held constant. Thus, there will be five experiments – one experiment for each of the Primary5 tables. We are not quantifying the effects of the non-Primary5 tables, the effects of total vehicle population, or the effects of total VMT. The non-Primary5 table values will be held constant for all MOVES runs using the non-Primary5 table values submitted by Montgomery County, TX. The total vehicle population will be held constant at 30,902. The total VMT will be held constant at 397,907,960 miles per year.

Degree of Perturbation – In any sensitivity analysis the size of the change in the response variable depends on the size of the change in the input perturbed variable. Thus, determining which inputs have larger or smaller influences on the outputs (emissions) depends on the size of the changes made to the inputs. The advantage of the values in the tables submitted by the states for the NEI is that they provide a measure of the range of values that may actually occur in various circumstances around the country. Accordingly, MOVES emissions results when using the submitted values will reveal an estimate of the relative importance of inputs on fleet emissions in an estimate of a realistic national context.

Simply running MOVES for all of the counties would not provide the sensitivity information that is desired because the fleet of each county is likely to operate under a set of conditions characteristic of that county but not characteristic of all counties taken together. The approach to be used in this sensitivity analysis will be to consider the range of values for a given Primary5 table for all counties. For each Primary5 table, a set of values characteristic of counties that represent the 10, 50, and 90 percentiles of all counties will be selected as MOVES input value for each Primary5 table. Note that the counties that represent the 10, 50, and 90 percentiles were chosen by evaluating a summary statistic for the table, for example, average vehicle AgeID for the SourceTypeAgeDistribution table. However, the inputs that we used for each MOVES run in each experiment was always an individual county's actual submitted values (or derivatives of actual submitted values, as described in Section 5.2, and never an average of several county values. We chose this approach so that the results would be more realistic.

The values submitted for each of the counties were analyzed to determine the 10, 50, and 90 percentile values. The 50 percentile is also known as the median. The range of emissions effects between 10 and 90 percentile represent the middle 80 percent of the values that were submitted. The 10 and 90 percentile values were chosen, rather than the most extreme values, to avoid using values that may have been unrealistically extreme (as an artifact of single-digit vehicle populations submitted for some rural counties, for example), while still using values that reflect the largest part of the range of values submitted for all counties as a whole.

5.2 Experimental Designs and Implementation

A set of MOVES runs for each of the experiments was designed to provide values of inputs and the associated output emissions values that were needed to conduct the sensitivity analysis. This subsection discusses creating derivative tables of the county-submitted data that will have the properties needed for the sensitivity analysis and the design of the runs for each experiment. Table 5-1 shows key attributes of the five experiments.

Table 5-1. Sensitivity Analysis Experiment Structures for Testing Primary5 Tables

	SourceTypeAgeDist	SourceTypeYear	HPMSVtypeYear	RoadTypeDistribution	AvgSpeedDistribution	Non-Primary5 Tables
Description of Table Values:	*AgeFraction by SourceTypeID*	*Population (Absolute) by SourceTypeID*	*VMT (Absolute) by HPMSVTypeID*	*RoadTypeFraction by SourceTypeID*	*SpeedBinFraction by SourceTypeID, RoadTypeID, HourDayID*	
Baseline:	50th percentile age distribution based on mean age	50th percentile population fraction multiplied by median total population (=30,902)	50th percentile VMT fraction multiplied by median total VMT (=397,907,960)	Roadtype distribution at 50th percentile of angles used	Hour-by-hour data for county at 50th percentile of mean speeds for cluster	Montgomery County, TX
Experiment 1 Two runs. 1: age distribution at 90th percentile of mean ages, -1: at 10th percentile.	Percentiles based on mean age	Same as baseline.	Same as baseline.	Same as baseline.	Same as baseline.	Montgomery County, TX
Experiment 2 Two runs. 1: at 90th percentile population fraction multiplied by median total population, -1: at 10th percentile population fraction multiplied by median total population.	Same as baseline.	Percentiles based on population fractions	Same as baseline.	Same as baseline.	Same as baseline.	Montgomery County, TX
Experiment 3 Two runs. 1: at 90th percentile VMT fraction multiplied by median total VMT, -1: at 10th percentile VMT fraction multiplied by median total VMT.	Same as baseline.	Same as baseline.	Percentiles based on VMT fractions	Same as baseline.	Same as baseline.	Montgomery County, TX
Experiment 4 Ten runs. Fraction of VMT for given roadtype set at 1, 0, or 0.5 for each run according to mixture design.	Same as baseline.	Same as baseline.	Same as baseline.	Fractions were set at different levels	Same as baseline.	Montgomery County, TX
Experiment 5 32 runs with combinations of 1 and -1 for each of 13 clusters. 1 at 90th percentile of mean speeds for cluster, -1 at 10th percentile of mean speeds for cluster. Actual hour-by-hour data for county at the chosen percentile used for MOVES inputs.	Same as baseline.	Same as baseline.	Same as baseline.	Same as baseline.	Percentiles based on mean speeds	Montgomery County, TX

The experiments used submitted data selected from the SourceTypeAgeDistribution table for vehicle age, the SourceTypeYear table for vehicle population, the HPMSVtypeYear for VMT, the RoadTypeDistribution for VMT, and the AvgSpeedDistribution for vehicle speed. The SourceTypeAgeDistribution, RoadTypeDistribution, and AvgSpeedDistribution tables contain values that are relative fractions of the quantity of interest. However, SourceTypeYear table and HPMSVtypeYear table contain absolute populations and VMTs, respectively. The absolute numbers were converted to relative numbers so that the sensitivity analysis can investigate the effects of shifts in the relative population and VMT rather than the effects of the absolute values of population and VMT. Accordingly, as will be discussed in Section 5.3, the SourceTypeYear table and HPMSVtypeYear table values were used to create derivative Relative Table Population and Relative Table VMT by dividing the original table values by the population and VMT of the county, respectively. However, for the purposes of creating the MOVES inputs for the sensitivity analysis runs, absolute population and VMT input values were created by multiplying the relative values by the median population (30,902) and median VMT (397,907,960).

As shown in Table 5-1, Experiment 1 tests the effects of the SourceTypeAgeDistribution table. Because the each SourceTypeID contributes emissions independently, no interactions occur among SourceTypeIDs. Therefore, only two MOVES runs are required to produce emissions values for each of the 13 SourceTypeIDs for 10 and 90 percentile age distributions. Table 5-1 also shows that the other four Primary5 tables are held at median conditions.

Just as for Experiment 1, Table 5-1 shows that Experiments 2 and 3 each require only two MOVES runs.

Experiment 4, which was designed to investigate the effects of the distribution of VMT among the four RoadTypeIDs, presented a different situation. An analysis of the dataset's relative VMT distributions among RoadTypeIDs indicated that the relative VMT fraction varied from 0 to nearly 1 for each of the four RoadTypeIDs. In this case, we used a mixture design in four variables, which are the relative fractions of VMT distributed among the four RoadTypeIDs. The four variables can be thought of as occupying a tetrahedron with each of the vertices representing 100% of the VMT driven on one of the RoadTypeIDs. To evaluate whether interactions existed (they should not) among the relative VMT fractions, we used a 10-run mixture design[10]. A mixture design is used because all four components are subject to the constraint that their sum must equal 1. The conditions for the 10 runs correspond to the 6 midpoints of the edges and the 4 vertices of the tetrahedron.

[10] J. Cornell, Experiments with Mixtures, John Wiley & Sons, Inc., Third Edition, 2002, page 23, Figure 2.1.

Experiment 5 was designed using a 2-level, 32-run fractional factorial design, 2^{13-8}_{IV}. This design can determine the main effects of 13 parameters with Resolution IV, which means the main effects are confounded with, at worst, the three-factor or higher-order interactions. The generators for the design were obtained from Wu and Hamada[11].

Interactions do exist between tables, but in this initial study of MOVES outputs we are concentrating primarily on the main effects of the Primary5 tables. However, the interaction between the RoadTypeDistribution and AvgSpeedDistribution will be quantified in the analysis in Section 5.4.

5.3 Selection of 10, 50, and 90 Percentile Values for Inputs of the Primary5 Tables

In the previous section, the preparation of each of the Primary5 data tables was described, including the organization of each table into percentiles. These percentiles were used to choose inputs for use in the MOVES runs: each run used a distribution found at either 10, 90, or 50 percentile. In the experiments using these distributions, 90 and 10 percentiles represented the +1 and -1 conditions, while 50 percentile, or median, represented the baseline condition. The resulting percentiles and distributions are described below for each of the Primary5 tables.

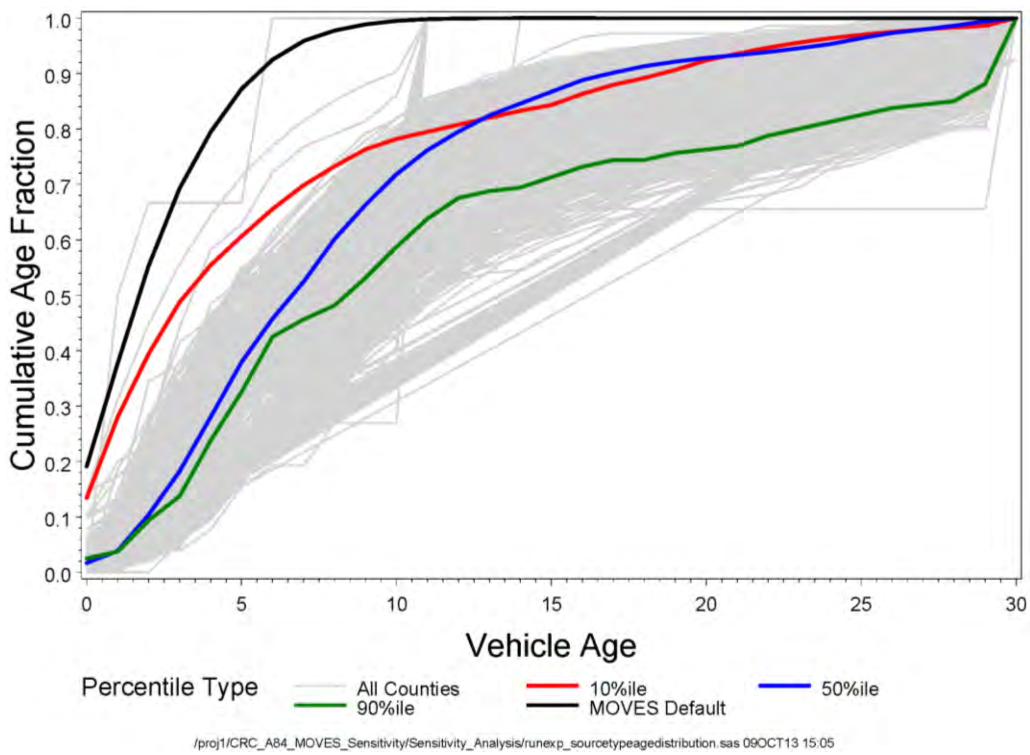
SourceTypeAgeDistribution Table

For this table, the percentiles were based on the mean AgeID of vehicles in each county, for each SourceTypeID. The mean AgeIDs were calculated from the AgeFractions – or the fraction of the fleet at each of 31 ages. The AgeFractions are the actual inputs for MOVES runs, not the mean AgeIDs.

In Figure 5-1, the cumulative fraction of the fleet of motorcycles in each of the 31 AgeIDs is shown graphically. The figure includes one line for the age distribution for the county that had a mean age at 10 percentile, one line for the distribution at 90 percentile, and one line for the distribution at the median. Additionally, there is one line for the MOVES default values that would be used if no county data were submitted. The distributions for all of the unselected counties are shown in gray in the background.

[11] C.F.J. Wu, M. Hamada, Experiments: Planning, Analysis, and Parameter Design Optimization, John Wiley & Sons, Inc., 2000, page 195, Table 4A.3.

Figure 5-1. Age Distributions for SourceTypeID=11 (Motorcycles)



From Figure 5-1, it is interesting to note that the default age distribution in MOVES for motorcycles is dominated by newer vehicles, more than any county's submitted data. It can also be seen that at 10 percentile, the age distribution includes more of the newer vehicles younger than AgeID 13 when compared to the median, and then is similar to the median for older vehicles. The line for 90 percentile is well below the median, indicating an older fleet.

As shown by Figure 5-2 for passenger cars, the 10 percentile, median, and 90 percentile distributions are well separated. The MOVES default is the most similar to 10 percentile. From Figure 5-3 for combination long-haul trucks, we can see that for the 5 newest vehicle ages, the median distribution is actually below the 90 percentile distribution – and it will later be seen that this has an effect on the final results in the MOVES runs.

Figures for the other SourceTypeIDs were similar, with the median either near the 10 percentile or closer to the median.

Figure 5-2. Age Distributions for SourceTypeID=21 (Passenger Cars)

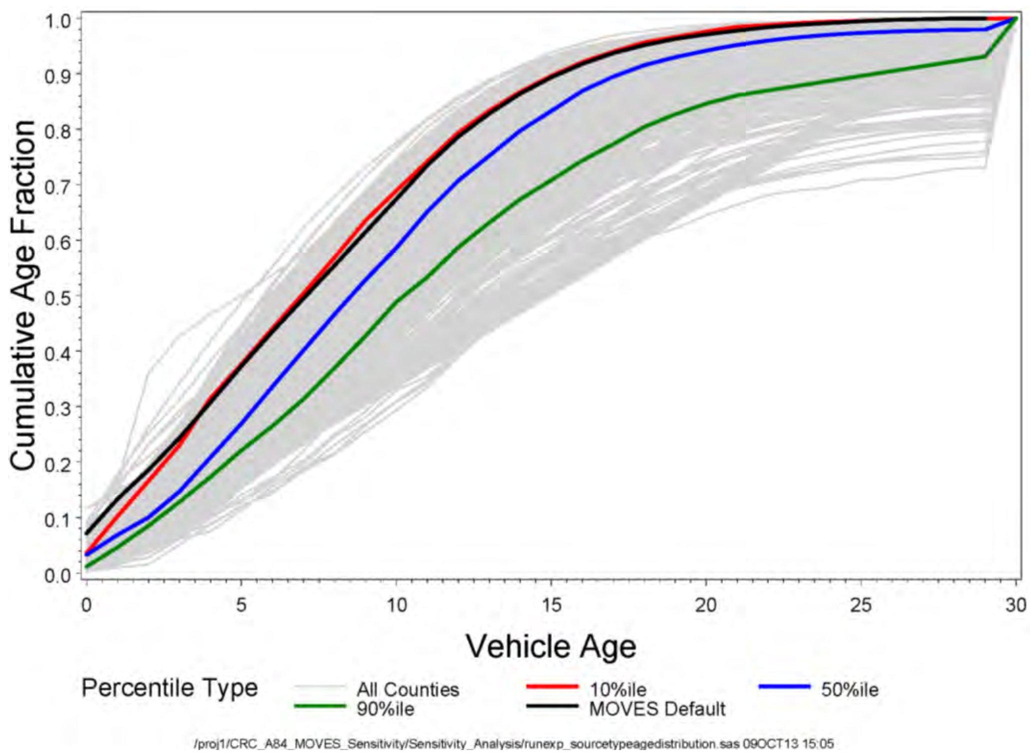
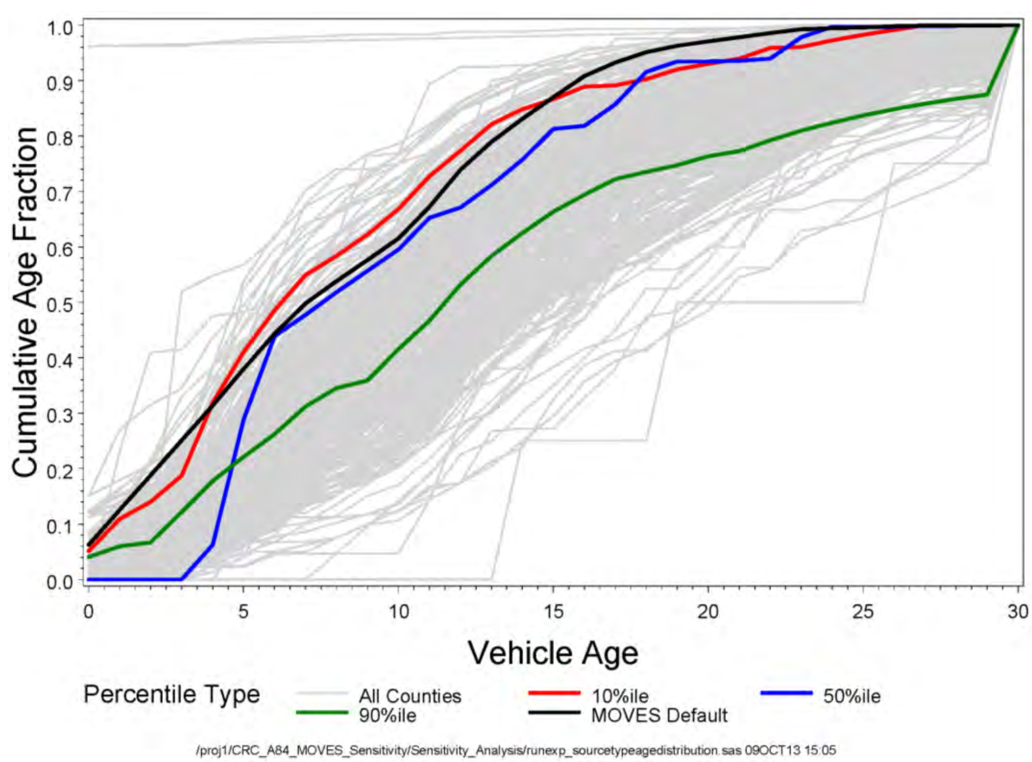


Figure 5-3. Age Distributions for SourceTypeID=62 (Comb. Long-Haul Truck)

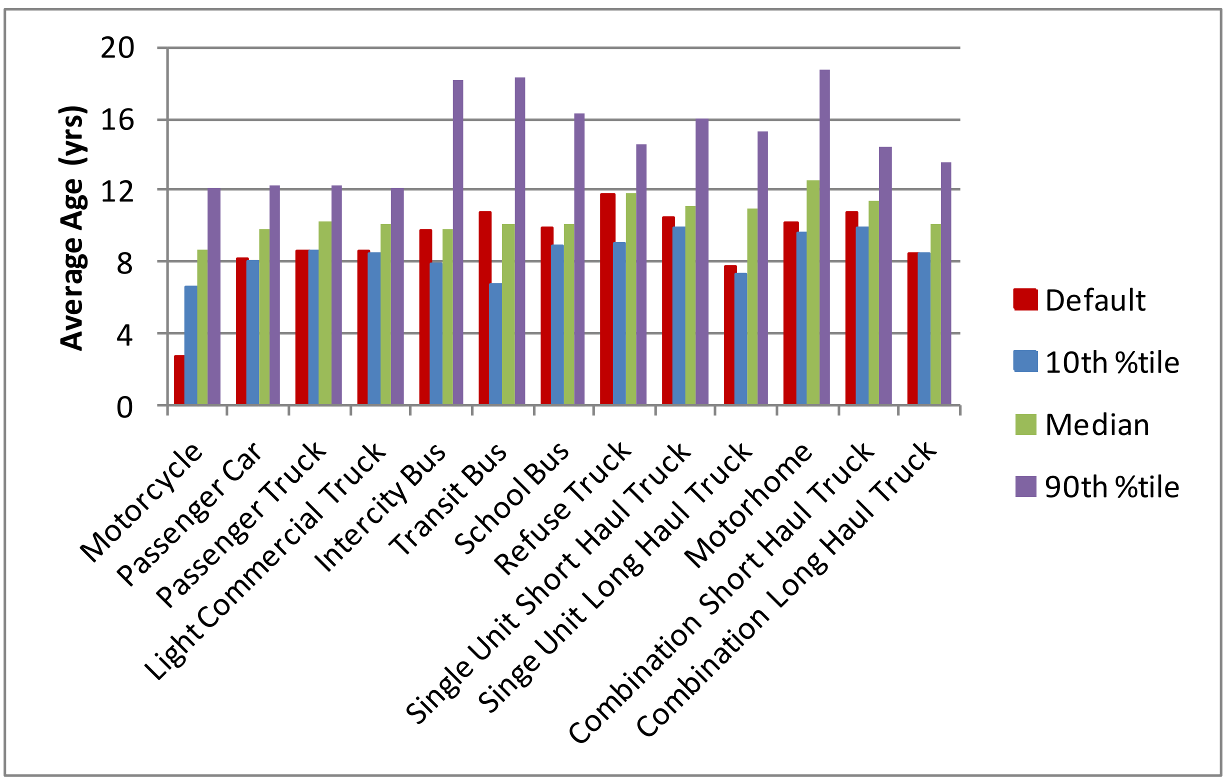


The mean AgeIDs for each SourceTypeID at each percentile are given in Table 5-2 and Figure 5-4. The mean AgeIDs were used to develop the percentiles for the distributions. Levels for age distribution were chosen based on average age. The median, 10th and 90th percentile average ages for the submitted data, by SourceTypeID, are shown in Figure 5-4, along with the averages of the default MOVES distributions used in the NEI if states did not submit data. The data show that for most source types, the average age of the MOVES default is at the lower end of the submitted data distribution. Many states submitted age distribution data with an older average age than MOVES defaults, particularly for buses and heavy trucks. Analysis by county showed that more populous urban areas tended to have younger fleets while rural areas had older fleets, so the population-weighted average age is generally below the median.

Table 5-2. Mean Ages for Each Condition and SourceTypeID

	11	21	31	32	41	42	43	51	52	53	54	61	62
10th pctl	6.7	8.1	8.6	8.5	8.0	6.8	8.9	9.0	9.9	7.3	9.6	9.9	8.5
50th pctl	8.7	9.8	10.3	10.2	9.8	10.2	10.2	11.8	11.2	10.9	12.6	11.4	10.2
90th pctl	12.1	12.3	12.3	12.2	18.2	18.3	16.3	14.5	15.9	15.3	18.8	14.4	13.6

Figure 5-4. Median, 10th and 90th Percentile Average Age, vs. MOVES Defaults

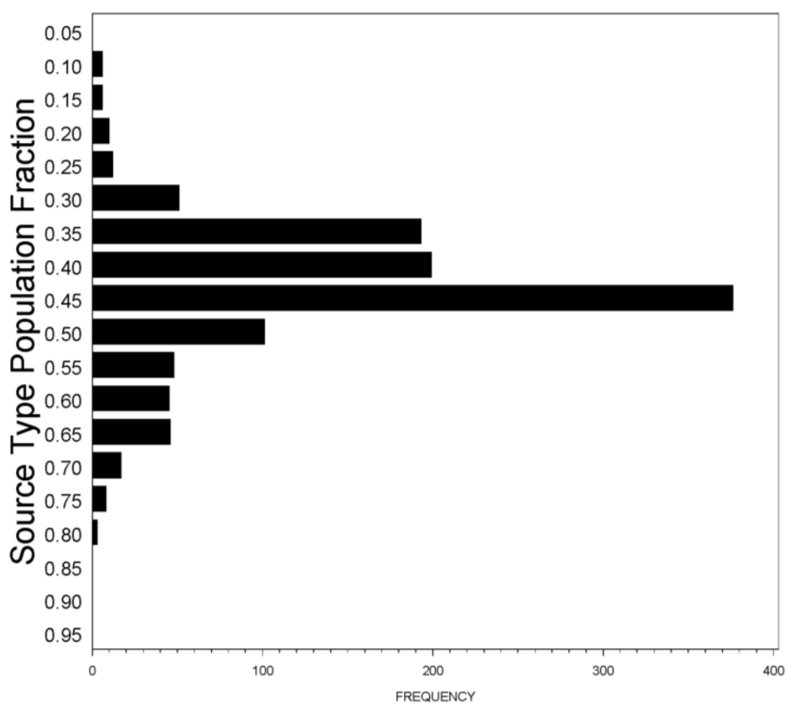


SourceTypeYear Table

For this table, the percentiles were based on the fractional distribution of vehicles of each SourceTypeID. The fractions for each county were calculated as the number of vehicles of a given SourceTypeID divided by the total number of vehicles in that county. Once the 10, 50, and 90 percentile fractions were selected, the population numbers that were used as MOVES inputs were found as the fraction for a given SourceTypeID, multiplied by the median total county population (30,902 vehicles).

The figures below show the range of fractional distributions for several of the SourceTypeIDs. Figure 5-5 shows that passenger cars comprise between 10 and 80% of vehicles over all counties, with a peak at about 45%. Passenger trucks are similar in Figure 5-6, but with a slightly lower peak fraction. In contrast, refuse trucks, shown in Figure 5-7, or single short haul trucks shown in Figure 5-8, are far less common in the county fleets. The 10, 50, and 90 percentiles were selected from the distributions as shown in the figures. The selections are listed in Table 5-3, and shown in Figure 5-9.

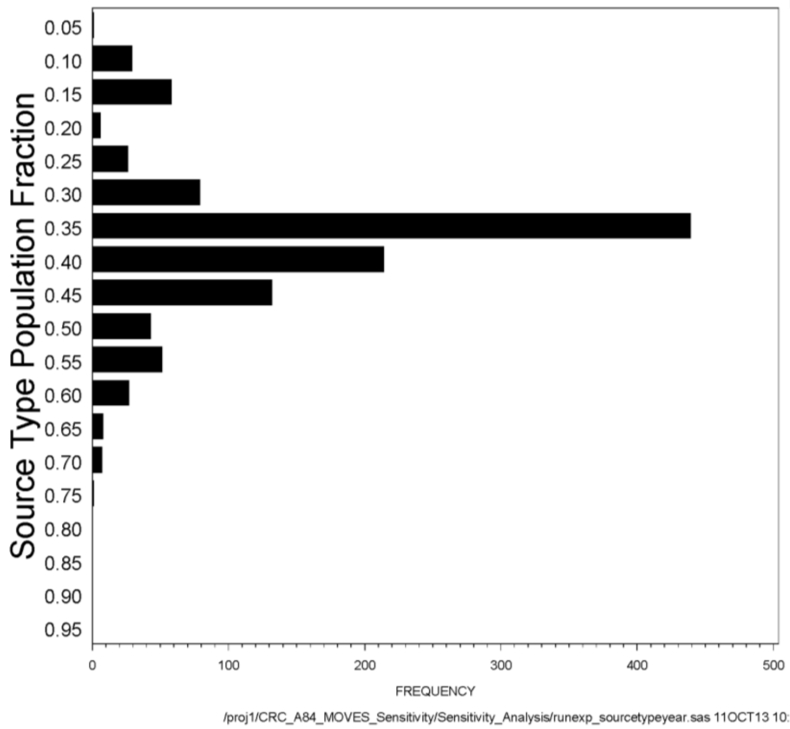
Figure 5-5. Population Fractions for SourceTypeID=21 (Passenger Cars)



FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
0	0	0.00	0.00
6	6	0.54	0.54
6	12	0.54	1.07
10	22	0.89	1.96
12	34	1.07	3.03
51	85	4.55	7.58
193	278	17.22	24.80
199	477	17.75	42.55
376	853	33.54	76.09
101	954	9.01	85.10
48	1002	4.28	89.38
45	1047	4.01	93.40
46	1093	4.10	97.50
17	1110	1.52	99.02
8	1118	0.71	99.73
3	1121	0.27	100.00
0	1121	0.00	100.00
0	1121	0.00	100.00
0	1121	0.00	100.00

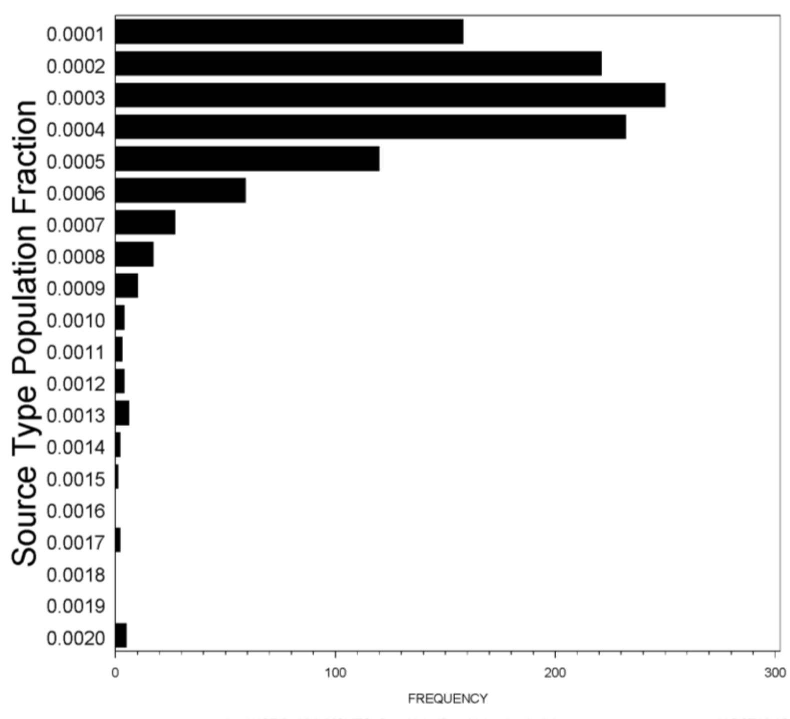
/proj1/CRC_A84_MOVES_Sensitivity/Sensitivity_Analysis/runexp_sourcetypeyear.sas 11OCT13 10:54

Figure 5-6. Population Fractions for SourceTypeID=31 (Passenger Trucks)



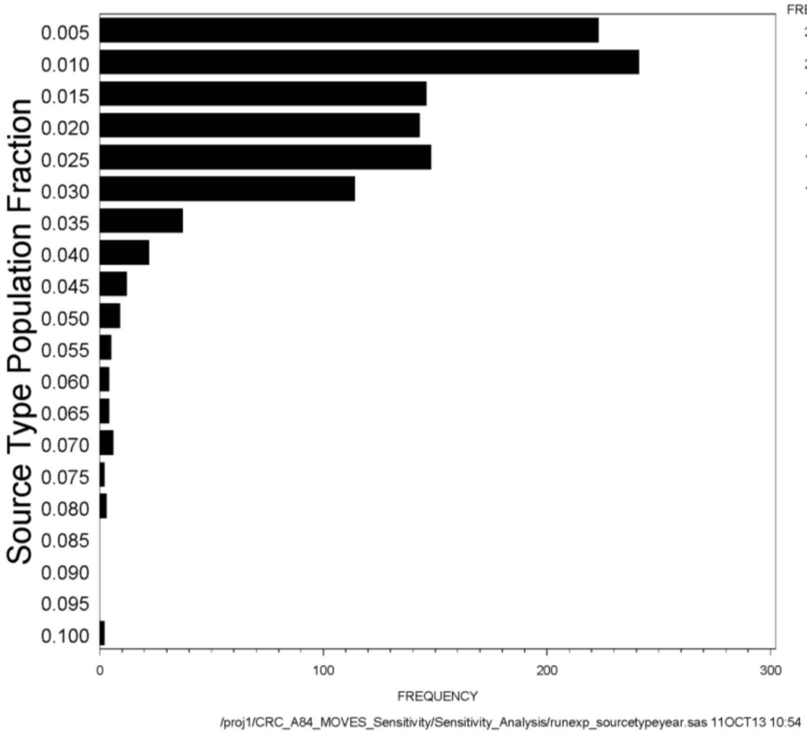
Source Type Population Fraction	FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
0.05	1	1	0.09	0.09
0.10	29	30	2.59	2.68
0.15	58	88	5.17	7.85
0.20	6	94	0.54	8.39
0.25	26	120	2.32	10.70
0.30	79	199	7.05	17.75
0.35	439	638	39.16	56.91
0.40	214	852	19.09	76.00
0.45	132	984	11.78	87.78
0.50	43	1027	3.84	91.61
0.55	51	1078	4.55	96.16
0.60	27	1105	2.41	98.57
0.65	8	1113	0.71	99.29
0.70	7	1120	0.62	99.91
0.75	1	1121	0.09	100.00
0.80	0	1121	0.00	100.00
0.85	0	1121	0.00	100.00
0.90	0	1121	0.00	100.00
0.95	0	1121	0.00	100.00

Figure 5-7. Population Fractions for SourceTypeID=51 (Refuse Trucks)



Source Type Population Fraction	FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
0.0001	158	158	14.09	14.09
0.0002	221	379	19.71	33.81
0.0003	250	629	22.30	56.11
0.0004	232	861	20.70	76.81
0.0005	120	981	10.70	87.51
0.0006	59	1040	5.26	92.77
0.0007	27	1067	2.41	95.18
0.0008	17	1084	1.52	96.70
0.0009	10	1094	0.89	97.59
0.0010	4	1098	0.36	97.95
0.0011	3	1101	0.27	98.22
0.0012	4	1105	0.36	98.57
0.0013	6	1111	0.54	99.11
0.0014	2	1113	0.18	99.29
0.0015	1	1114	0.09	99.38
0.0016	0	1114	0.00	99.38
0.0017	2	1116	0.18	99.55
0.0018	0	1116	0.00	99.55
0.0019	0	1116	0.00	99.55
0.0020	5	1121	0.45	100.00

Figure 5-8. Population Fractions for SourceTypeID=52 (Single Unit Short Haul Trucks)

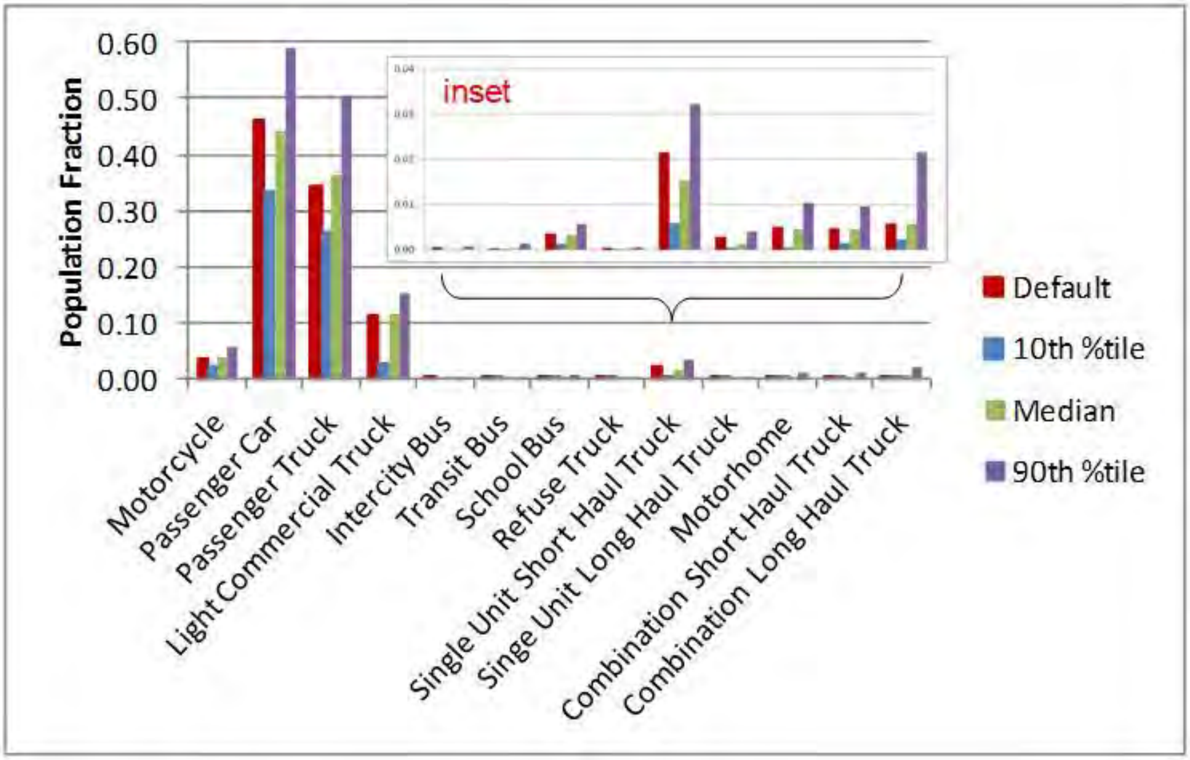


FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
223	223	19.89	19.89
241	464	21.50	41.39
146	610	13.02	54.42
143	753	12.76	67.17
148	901	13.20	80.37
114	1015	10.17	90.54
37	1052	3.30	93.84
22	1074	1.96	95.81
12	1086	1.07	96.88
9	1095	0.80	97.68
5	1100	0.45	98.13
4	1104	0.36	98.48
4	1108	0.36	98.84
6	1114	0.54	99.38
2	1116	0.18	99.55
3	1119	0.27	99.82
0	1119	0.00	99.82
0	1119	0.00	99.82
0	1119	0.00	99.82
2	1121	0.18	100.00

Table 5-3. Relative Population Fractions for Each Condition and SourceTypeID

	11	21	31	32	41	42	43	51	52	53	54	61	62
10^{th} pctl	0.0215	0.3355	0.2635	0.0264	0.0000	0.0000	0.0013	0.0001	0.0059	0.0005	0.0005	0.0015	0.0023
50^{th} pctl	0.0366	0.4428	0.3643	0.1141	0.0002	0.0003	0.0032	0.0003	0.0153	0.0013	0.0045	0.0045	0.0057
90^{th} pctl	0.0555	0.5903	0.5058	0.1508	0.0008	0.0013	0.0056	0.0006	0.0321	0.0040	0.0103	0.0096	0.0215
Median total population: 30,902													

Figure 5-9. Median, 10th and 90th Percentile Population Fractions vs. MOVES Defaults



HPMSVtypeYear Table

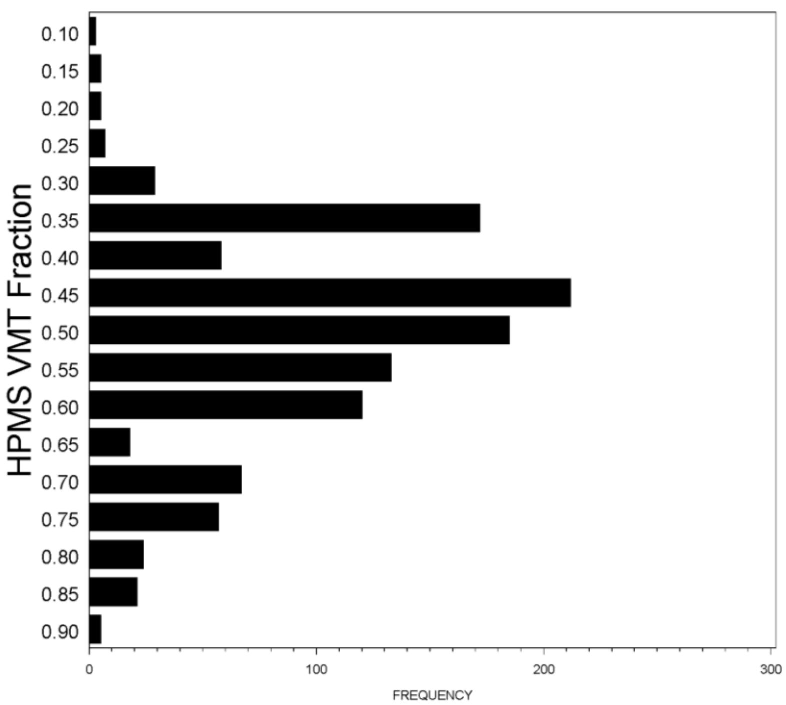
For this table, the percentiles were based on the fractional distribution of VMT for each HPMSVTypeID. The fractions for each county were calculated as the number of miles for a given SourceTypeID divided by the total number of miles in that county. Once 10, 50, and 90 percentile fractions were selected, the VMT numbers that were used as MOVES inputs were found as the fraction for a given SourceTypeID, multiplied by the median total county VMT (397,907,960 miles).

The figures below show the range of relative VMT fractions for several of the HPMSVTypeIDs. In Figure 5-10, we see that passenger cars comprise between 10 and 90% of VMT over all counties, with a peak at about 45%. This is similar to the fractional distributions that were seen for passenger car populations. Passenger trucks are similar in Figure 5-11, but with a slightly lower peak fraction. Finally, single unit short haul trucks are shown in Figure 5-12, with a much lower range of VMT fractions. The 10, 50, and 90 percentiles were selected from the distributions as shown in the figures. The selections are listed in Table 5-4 and shown in Figure 5-13. Levels for VMT were based on VMT fraction by Highway Performance Monitoring System (HPMS) vehicle class, the classification MOVES uses for VMT. Figure 5-13 shows the median, 10th and 90th percentile fraction for each HPMSVtypeID, along with the MOVES national defaults. As shown, the defaults fall within the range of data, and track the median values well. The data show a large difference for some vehicle classes between the 10th and 90th percentile values, reflecting large variation across the country in how VMT is split. For passenger cars, the 90th percentile fraction is around 2 times that of the 10th percentile fraction; for light trucks, 3 times; for combination trucks, the 90th percentile fraction is nearly 7 times higher than the 10th percentile fraction.

Table 5-4. Relative VMT Fractions for Each Condition and HPMSVtypeID

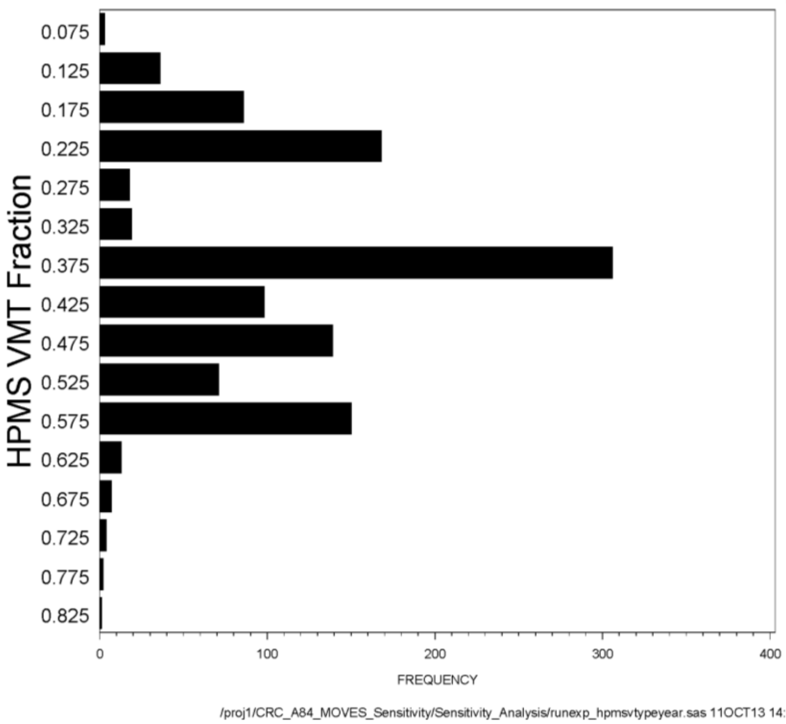
	10	20	30	40	50	60
10th pctl	0.003	0.355	0.192	0.002	0.011	0.016
50th pctl	0.006	0.505	0.375	0.004	0.030	0.051
90th pctl	0.013	0.719	0.584	0.007	0.048	0.110
Median total VMT: 397,907,960						

Figure 5-10. VMT Fractions for HPMSVtypeID=20 (Passenger Cars)



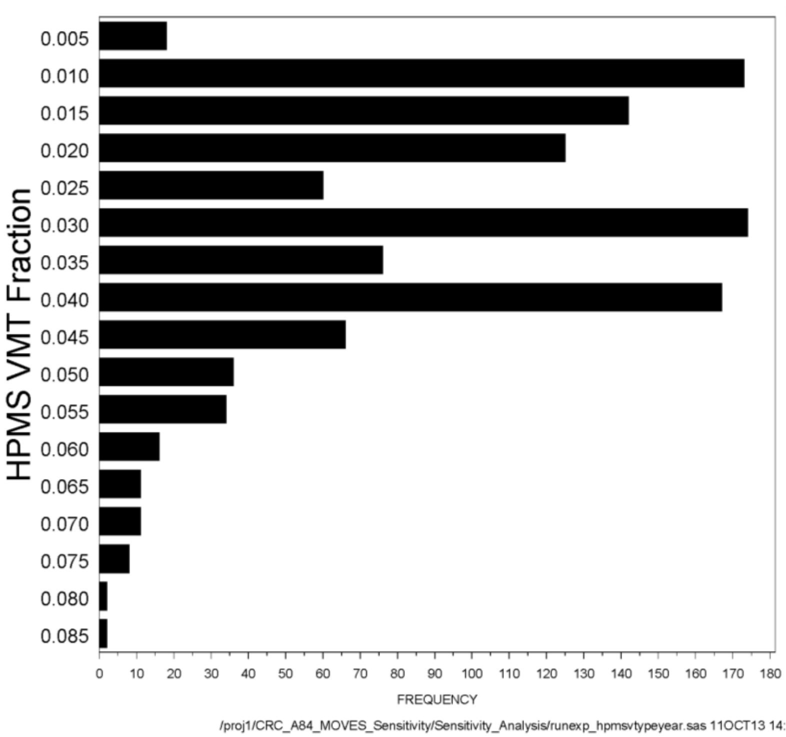
HPMS VMT Fraction	FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
0.10	3	3	0.27	0.27
0.15	5	8	0.45	0.71
0.20	5	13	0.45	1.16
0.25	7	20	0.62	1.78
0.30	29	49	2.59	4.37
0.35	172	221	15.34	19.71
0.40	58	279	5.17	24.89
0.45	212	491	18.91	43.80
0.50	185	676	16.50	60.30
0.55	133	809	11.86	72.17
0.60	120	929	10.70	82.87
0.65	18	947	1.61	84.48
0.70	67	1014	5.98	90.45
0.75	57	1071	5.08	95.54
0.80	24	1095	2.14	97.68
0.85	21	1116	1.87	99.55
0.90	5	1121	0.45	100.00

Figure 5-11. VMT Fractions for HPMSVtypeID=30 (Light Trucks)



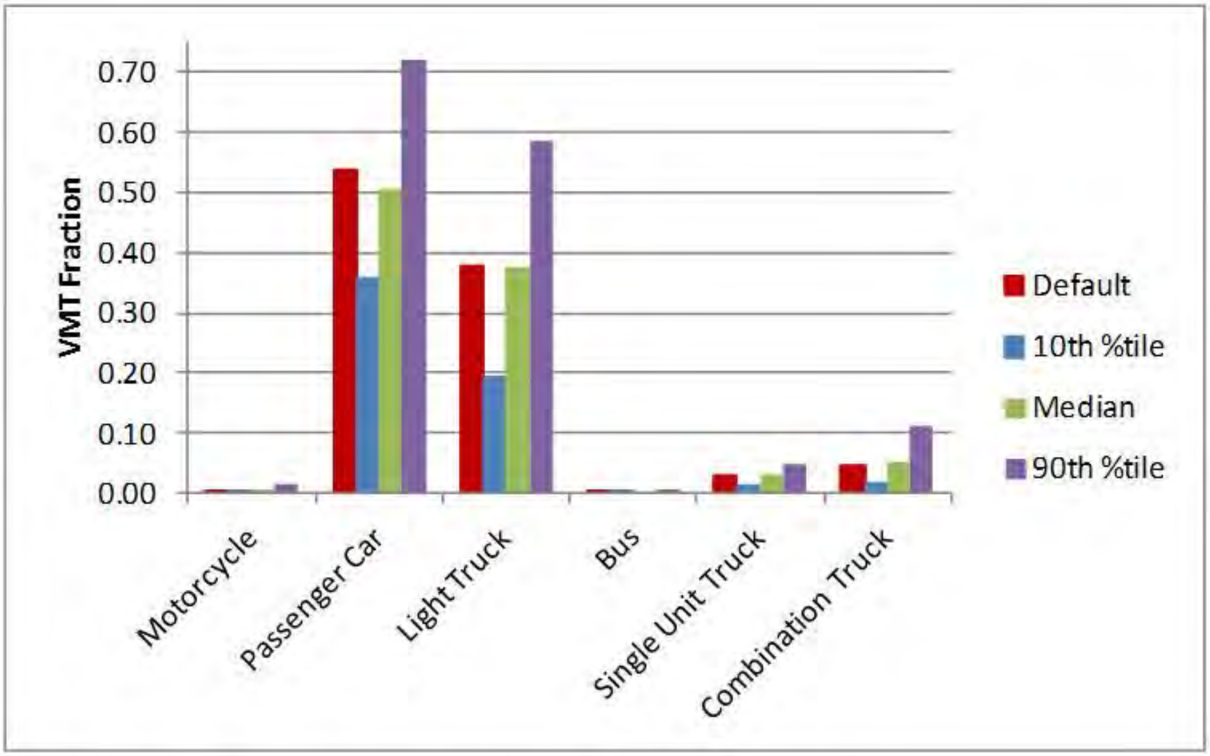
HPMS VMT Fraction	FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
0.075	3	3	0.27	0.27
0.125	36	39	3.21	3.48
0.175	86	125	7.67	11.15
0.225	168	293	14.99	26.14
0.275	18	311	1.61	27.74
0.325	19	330	1.69	29.44
0.375	306	636	27.30	56.74
0.425	98	734	8.74	65.48
0.475	139	873	12.40	77.88
0.525	71	944	6.33	84.21
0.575	150	1094	13.38	97.59
0.625	13	1107	1.16	98.75
0.675	7	1114	0.62	99.38
0.725	4	1118	0.36	99.73
0.775	2	1120	0.18	99.91
0.825	1	1121	0.09	100.00

Figure 5-12. VMT Fractions for HPMSVtypeID=50 (Single Unit Trucks)



HPMS VMT Fraction	FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
0.005	18	18	1.61	1.61
0.010	173	191	15.43	17.04
0.015	142	333	12.67	29.71
0.020	125	458	11.15	40.86
0.025	60	518	5.35	46.21
0.030	174	692	15.52	61.73
0.035	76	768	6.78	68.51
0.040	167	935	14.90	83.41
0.045	66	1001	5.89	89.30
0.050	36	1037	3.21	92.51
0.055	34	1071	3.03	95.54
0.060	16	1087	1.43	96.97
0.065	11	1098	0.98	97.95
0.070	11	1109	0.98	98.93
0.075	8	1117	0.71	99.64
0.080	2	1119	0.18	99.82
0.085	2	1121	0.18	100.00

Figure 5-13. Median, 10th and 90th Percentile VMT Fractions, vs. MOVES Defaults



RoadTypeDistribution Table

The RoadTypeDistribution table was different from the other tables in that while it contains fractional distributions (as do the SourceTypeAgeDistribution and AvgSpeedDistribution tables), the different RoadTypeIDs are merely categorical variables not ordinal variables. RoadTypeIDs are not numerically related to each other, so “average road type” for a county has no meaning. Therefore, because a mixture design (described in Section 5.2) was used to determine the relative emissions importance of driving on each of the 4 RoadTypeIDs, the 10 and 90 percentiles for each RoadTypeID were not needed. However, some sort of “baseline” distribution of RoadTypeIDs was needed for use in the MOVES runs for Experiments 1, 2, 3, and 5. Therefore, the county with the median vector angle was found, separately for each SourceTypeID, and used as the baseline county. As examples, the baseline distributions used for passenger cars, motorhomes, and combination short haul trucks are shown in Figures 5-14, 5-15, and 5-16.

Figure 5-14. Baseline RoadTypeID Distribution for SourceTypeID=21 (Passenger Cars)

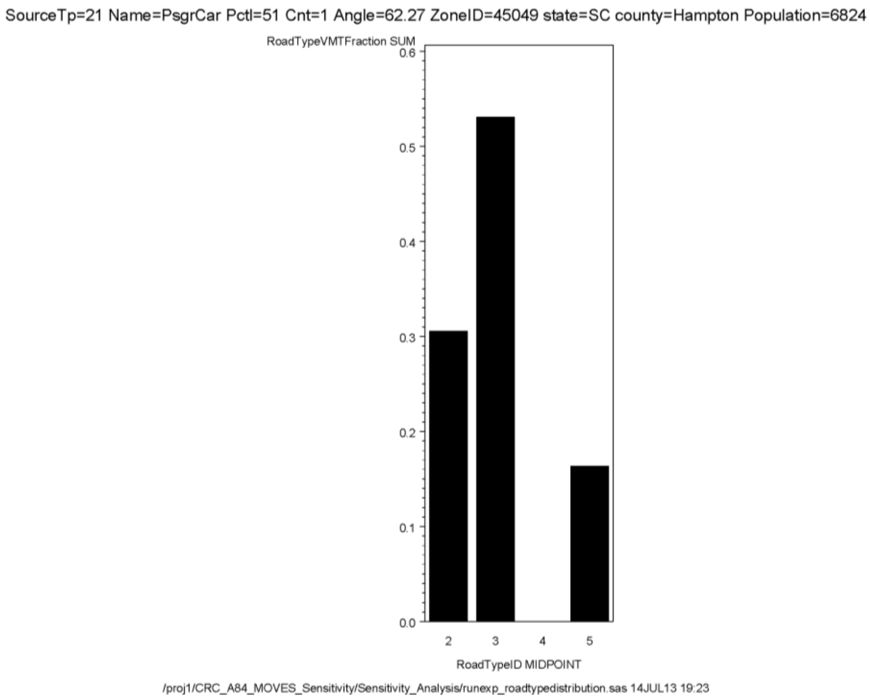


Figure 5-15. Baseline RoadTypeID Distribution for SourceTypeID=42 (Motorhomes)

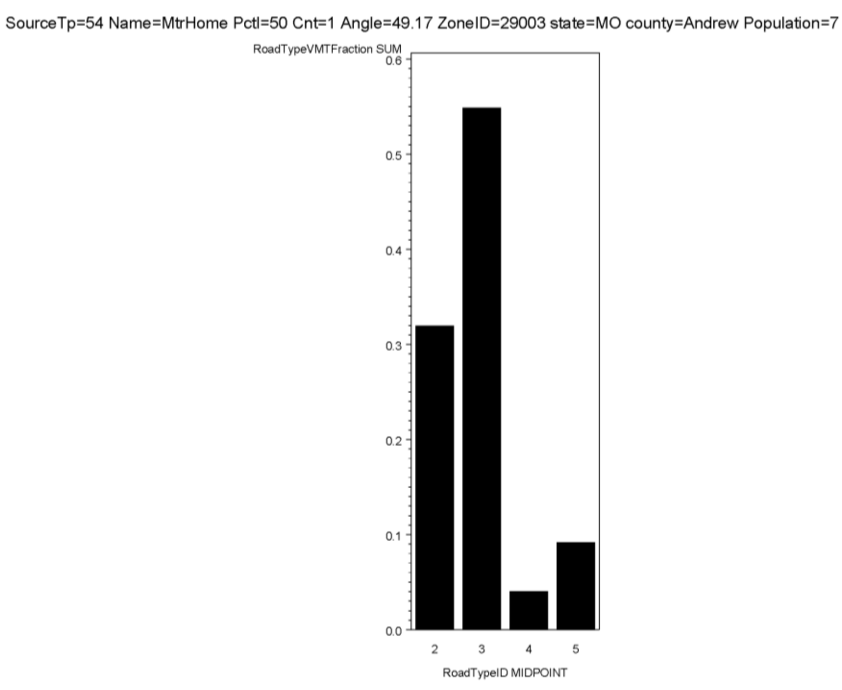
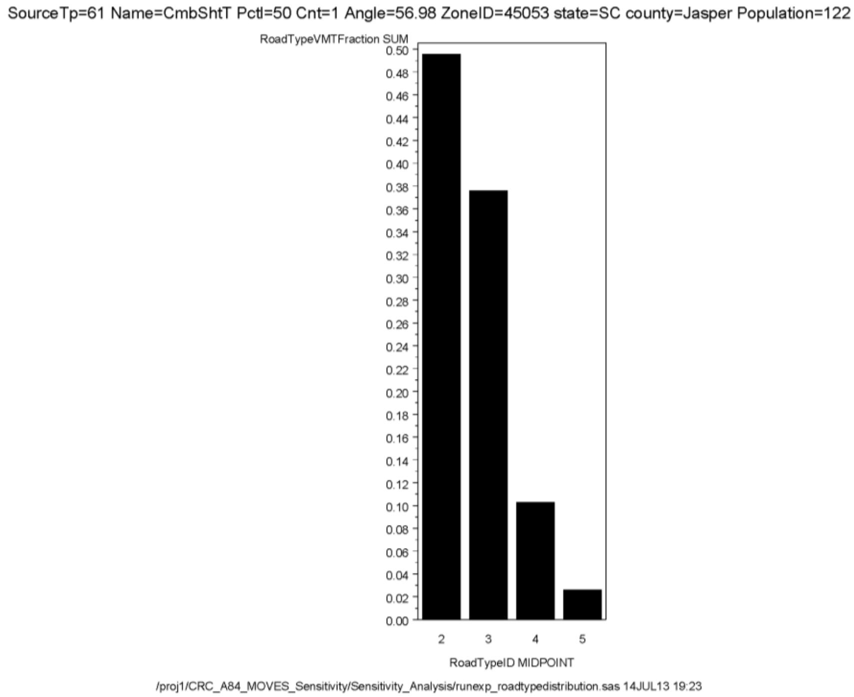


Figure 5-16. Baseline RoadTypeID Distribution for SourceTypeID=61 (Combination Short Haul Trucks)



AvgSpeedDistribution Table

The structure of the AvgSpeedDistribution table is similar to that of the SourceTypeAgeDistribution table having a series of ordered fractional values – in this case, one value for each of 16 SpeedBinIDs. However, the SourceTypeAgeDistribution table contains one stratification level (SourceTypeID), so that for each county there are 13 total distributions, one for each SourceTypeID. The AvgSpeedDistribution table uses four stratification levels: SourceTypeID (13 levels), RoadTypeID (4 levels), Hour of the Day (24 levels), and Weekday/Weekend (2 levels), resulting in 2496 distributions per county. To create a manageable dataset, the RoadTypeIDs, Hours, and Days were grouped into 13 clusters that had similar speed distributions within a cluster and different speed distributions between clusters. (See Section 4.4 for a description of the development of the 13 clusters.)

For this table, the percentiles were based on the mean AvgSpeedBinIDs of vehicles in each county, for each SourceTypeID and SpeedDistCluster. The mean AvgSpeedBinIDs for each SourceTypeID and SpeedDistCluster were calculated from the AvgSpeedFractions, which are the fractions of the driving time in each of 16 AvgSpeedBinIDs. The AvgSpeedFractions are the actual inputs for MOVES runs, not the mean AvgSpeedBinIDs

In Figure 5-17, the cumulative AvgSpeedFraction of the driving time in each of the 16 AvgSpeedBinIDs is shown graphically, for passenger cars in SpeedDistCluster 4 (rush hour on urban restricted roads). The figure includes one line for the speed distribution for the county that had a mean AvgSpeedBinID at 10 percentile, one line for the distribution at 90 percentile, and one line for the distribution at the median. Additionally, there is one line for the MOVES default values that would be used if no county data were submitted. The distributions for all of the unselected counties are shown in gray in the background. Figures 5-18 and 5-19 show similar distributions for passenger cars in SpeedDistClusters 9 (non-rush hour on urban restricted roads) and 2 (weekdays on rural un-restricted roads), respectively. The distributions for SpeedDistClusters 4 and 9 are similar, except that SpeedDistCluster 4 contains more low speed driving presumably caused by congestion during rush hour. Those two figures also show that 10 percentile and median are exactly the same as the defaults. Finally, SpeedDistCluster 2, on rural roads, shows a differently shaped distribution, and neither 10 nor 90 percentile distributions are the same as the default (although the median uses the default).

Figure 5-17. Speed Bin Distributions for SpeedDistCluster 4, Passenger Cars

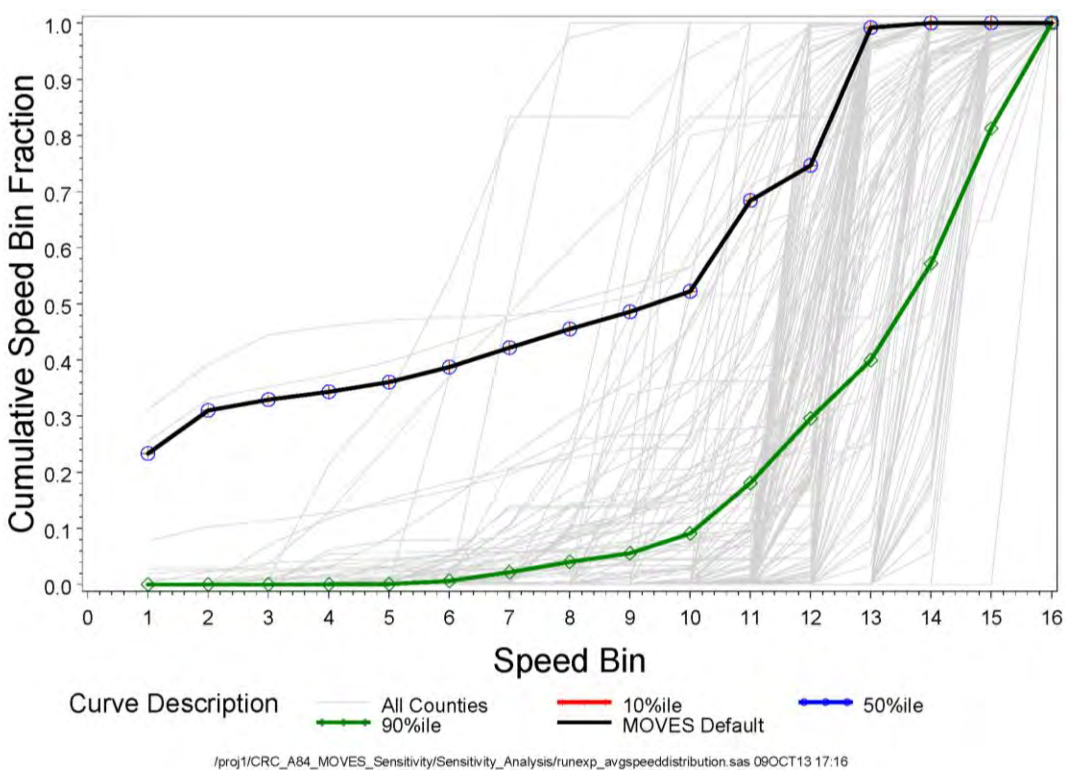


Figure 5-18. Speed Bin Distributions for SpeedDistCluster 9, Passenger Cars

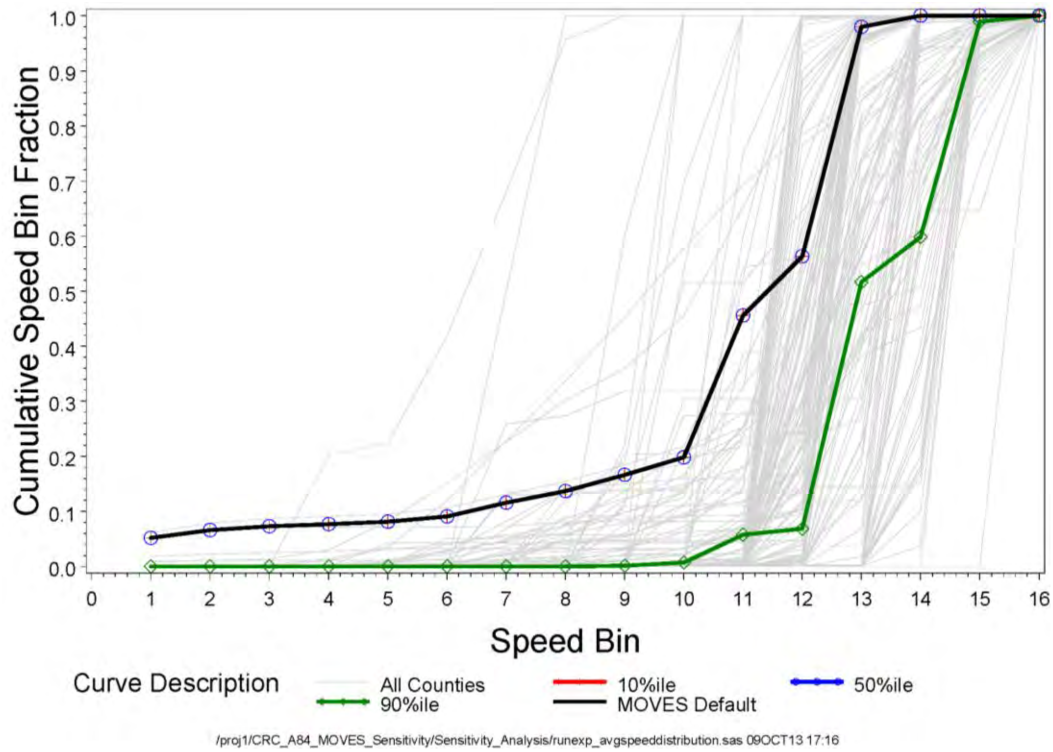
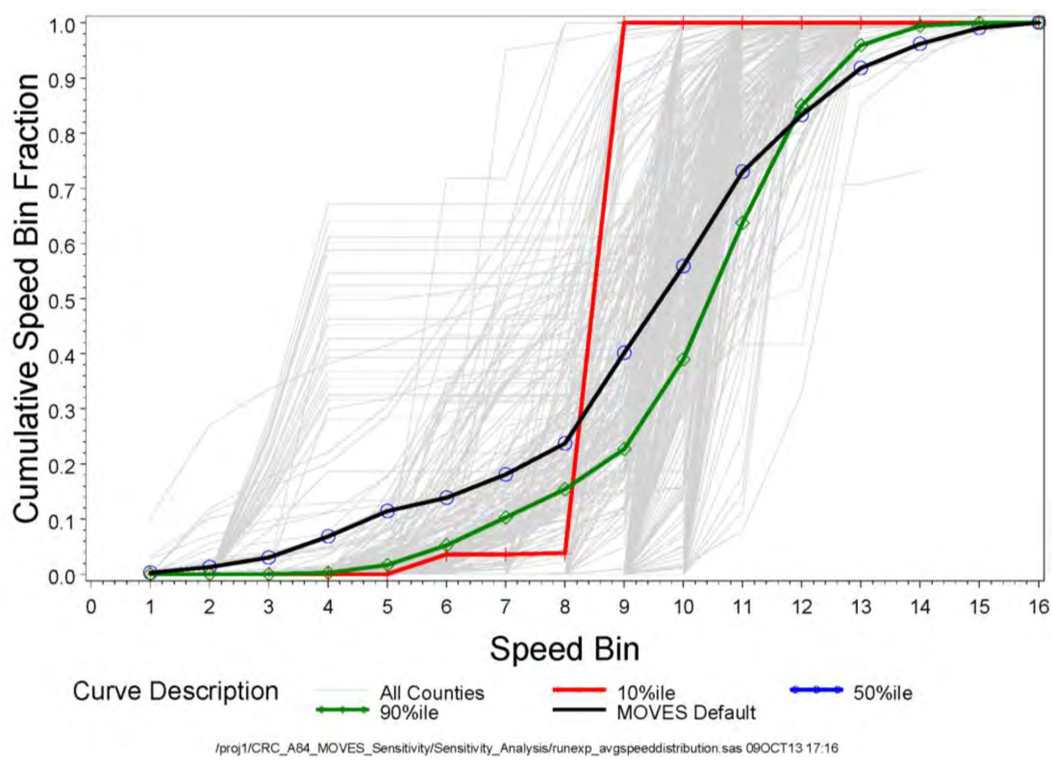


Figure 5-19. Speed Bin Distributions for SpeedDistCluster 2, Passenger Cars



Tables 5-5 through 5-17 give the mean AvgSpeedBinID for each SpeedDistCluster, respectively. The tables show 10, 50, and 90 percentiles for each SourceTypeID. It can be seen from the table that there are a number of instances of 10 percentile and median having the same mean AvgSpeedBinIDs, or the median and 90 percentile having the same mean AvgSpeedBinIDs. This results from the high number of counties submitting the default distributions, and both the median and 10 or 90 percentiles representing the same default distribution. Results by cluster, averaged across source bin, are shown in Figure 5-20.

Table 5-5. Mean AvgSpeedBinIDs for Cluster 1

	11	21	31	32	41	42	43	51	52	53	54	61	62
10^{th} pctl	8.9	9.0	9.0	9.0	9.0	9.0	8.9	8.9	9.0	9.0	8.9	9.0	8.9
50^{th} pctl	9.8	9.8	9.8	9.8	9.4	9.4	9.4	9.8	9.8	9.8	9.8	10.4	10.4
90^{th} pctl	10.8	10.8	10.8	10.8	10.7	10.7	10.8	10.6	10.8	10.8	10.8	10.8	10.6

Table 5-6. Mean AvgSpeedBinIDs for Cluster 2

	11	21	31	32	41	42	43	51	52	53	54	61	62
10^{th} pctl	8.8	8.9	8.9	8.9	8.9	8.9	8.8	8.8	8.9	8.9	8.7	8.9	8.8
50^{th} pctl	9.8	9.8	9.8	9.8	9.4	9.4	9.4	9.8	9.8	9.8	9.8	10.4	10.4
90^{th} pctl	10.6	10.6	10.6	10.6	10.7	10.7	10.7	10.6	10.6	10.6	10.6	10.6	10.6

Table 5-7. Mean AvgSpeedBinIDs for Cluster 3.

	11	21	31	32	41	42	43	51	52	53	54	61	62
10^{th} pctl	9.0	9.1	9.0	9.1	9.0	9.0	9.0	9.0	9.0	9.1	8.9	9.1	9.0
50^{th} pctl	9.8	9.8	9.8	9.8	9.4	9.4	9.4	9.8	9.8	9.8	9.8	10.4	10.4
90^{th} pctl	11.0	10.9	10.9	10.9	10.7	10.7	11.0	10.7	10.9	10.9	11.0	10.9	10.7

Table 5-8. Mean AvgSpeedBinIDs for Cluster 4.

	11	21	31	32	41	42	43	51	52	53	54	61	62
10th pctl	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7
50th pctl	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7
90th pctl	13.6	13.5	13.5	13.5	13.6	13.6	13.6	13.6	13.5	13.5	13.5	13.5	13.6

Table 5-9. Mean AvgSpeedBinIDs for Cluster 5.

	11	21	31	32	41	42	43	51	52	53	54	61	62
10^{th} pctl	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7
50^{th} pctl	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7
90^{th} pctl	8.9	8.8	8.8	8.9	8.9	8.9	8.9	8.9	8.8	8.8	8.9	8.9	8.9

Table 5-10. Mean AvgSpeedBinIDs for Cluster 6.

	11	21	31	32	41	42	43	51	52	53	54	61	62
10th pctl	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0
50th pctl	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0
90th pctl	9.3	9.3	9.2	9.3	9.1	9.1	9.3	9.1	9.2	9.3	9.3	9.3	9.0

Table 5-11. Mean AvgSpeedBinIDs for Cluster 7.

	11	21	31	32	41	42	43	51	52	53	54	61	62
10th pctl	5.8	5.8	5.8	5.8	5.8	5.8	5.8	5.8	5.8	5.8	5.8	5.8	5.8
50th pctl	5.8	5.8	5.8	5.8	5.8	5.8	5.8	5.8	5.8	5.8	5.8	5.8	5.8
90th pctl	9.2	9.2	9.2	9.2	9.0	9.0	9.2	9.0	9.2	9.2	9.2	9.2	9.0

Table 5-12. Mean AvgSpeedBinIDs for Cluster 8.

	11	21	31	32	41	42	43	51	52	53	54	61	62
10th pctl	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4
50th pctl	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4
90th pctl	13.6	13.6	13.6	13.6	13.6	13.6	13.6	13.6	13.6	13.6	13.6	13.6	13.6

Table 5-13. Mean AvgSpeedBinIDs for Cluster 9.

	11	21	31	32	41	42	43	51	52	53	54	61	62
10th pctl	10.9	10.9	10.9	10.9	10.9	10.9	10.9	10.9	10.9	10.9	10.9	10.9	10.9
50th pctl	11.1	11.1	11.1	11.1	11.1	11.1	11.1	11.1	11.1	11.1	11.1	11.1	11.1
90th pctl	13.8	13.8	13.8	13.8	13.8	13.8	13.8	13.8	13.8	13.8	13.8	13.8	13.8

Table 5-14. Mean AvgSpeedBinIDs for Cluster 10.

	11	21	31	32	41	42	43	51	52	53	54	61	62
10th pctl	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3	8.3
50th pctl	9.0	9.0	9.0	9.0	9.0	9.0	9.0	9.0	9.0	9.0	9.0	9.0	9.0
90th pctl	13.6	13.6	13.6	13.6	13.6	13.6	13.6	13.6	13.6	13.6	13.6	13.6	13.6

Table 5-15. Mean AvgSpeedBinIDs for Cluster 11.

	11	21	31	32	41	42	43	51	52	53	54	61	62
10th pctl	13.9	13.7	13.7	13.7	13.9	13.9	13.8	13.6	13.7	13.7	13.7	13.7	13.6
50th pctl	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4
90th pctl	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4

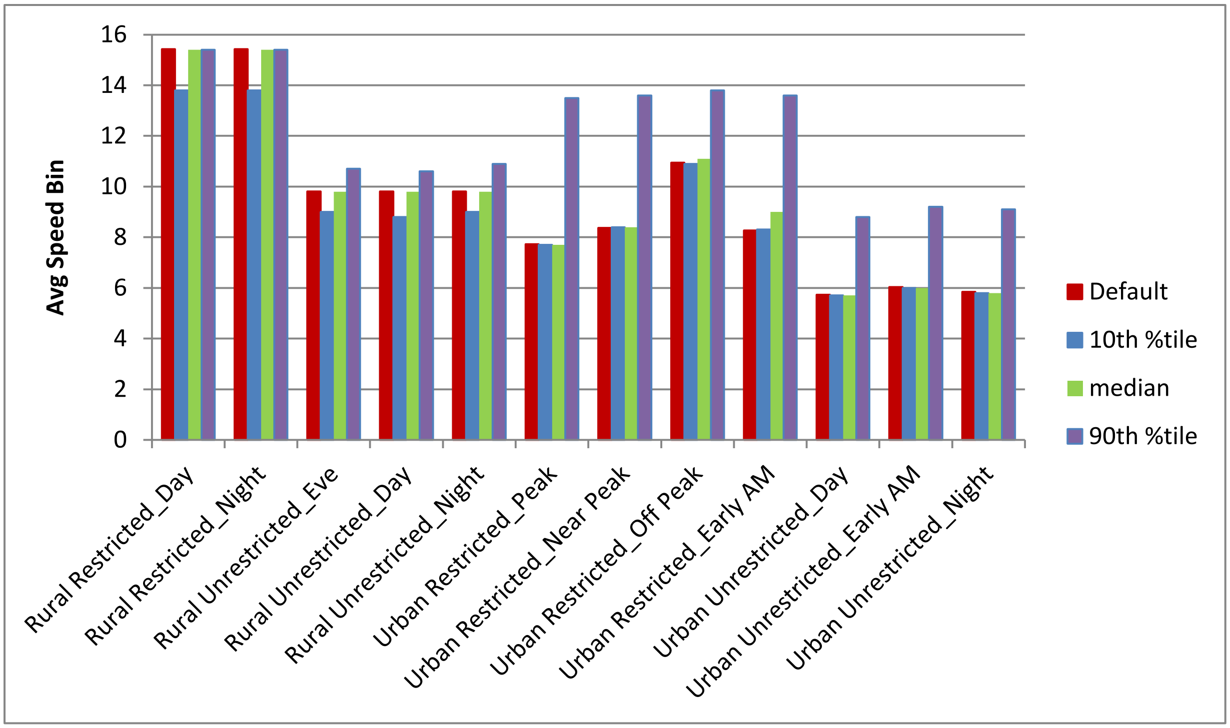
Table 5-16. Mean AvgSpeedBinIDs for Cluster 12.

	11	21	31	32	41	42	43	51	52	53	54	61	62
10th pctl	13.9	13.8	13.8	13.8	13.9	13.9	13.8	13.6	13.8	13.8	13.8	13.8	13.6
50th pctl	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4
90th pctl	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4

Table 5-17. Mean AvgSpeedBinIDs for Cluster 13.

	11	21	31	32	41	42	43	51	52	53	54	61	62
10th pctl	13.9	13.7	13.7	13.7	13.9	13.9	13.8	13.5	13.7	13.7	13.5	13.7	13.5
50th pctl	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4
90th pctl	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4	15.4

Figure 5-20. Median, 10th and 90th Percentile average speed bins, vs. MOVES defaults



5.4 Calculation of Percent Emission Changes from MOVES Output

The baseline run and the two Experiment 1 MOVES runs provided the emissions information for the analysis of the AvgAgeDistribution table for a weekday in July 2011. The total HC, CO, NX, and PM emissions results for each of the SourceTypeIDs were obtained from the MOVES output. The emissions were summed for each emission type and test condition and are tabulated in Columns 4, 5, 6, and 7 of Table 5-18. The second column indicates the 10 percentile average age distribution as -1 (young age), the 50 percentile average age distribution as 0 (median age), and the 90 percentile average age distribution as 1 (old age).

Table 5-18. Analysis of Results from Experiment 1 for the AvgAgeDistribution Table

Source TypeID	AvgAge Setting	AvgAge (yr)	ΣHC (t/day)	ΣCO (t/day)	ΣNX (t/day)	ΣPM2.5 (t/day)	ΔHC (t/day)	ΔCO (t/day)	ΔNX (t/day)	ΔPM2.5 (t/day)	ΔHC (%)	ΔCO (%)	ΔNX (%)	ΔPM2.5 (%)
	-1	6.7	1.0980	11.738	2.513	0.07498	-0.0013	-0.0258	-0.0003	0	-0.12	-0.22	-0.01	0
11	0	8.7	1.0993	11.764	2.514	0.07498	0.0000	0.0000	0.0000	0	0	0	0	0
	1	12.1	1.1017	11.766	2.514	0.07498	0.0024	0.0027	0.0001	0	0.22	0.02	0.00	0
	-1	8.1	1.0039	10.815	2.396	0.07333	-0.0954	-0.9488	-0.1171	-0.00165	-8.68	-8.07	-4.66	-2.20
21	0	9.8	1.0993	11.764	2.514	0.07498	0.0000	0.0000	0.0000	0	0	0	0	0
	1	12.3	1.2622	13.323	2.642	0.07890	0.1630	1.5597	0.1281	0.00391	14.83	13.26	5.10	5.22
	-1	8.6	0.9926	10.769	2.380	0.07387	-0.1066	-0.9942	-0.1332	-0.00112	-9.70	-8.45	-5.30	-1.49
31	0	10.3	1.0993	11.764	2.514	0.07498	0.0000	0.0000	0.0000	0	0	0	0	0
	1	12.3	1.2378	13.703	2.653	0.07675	0.1386	1.9398	0.1391	0.00177	12.61	16.49	5.53	2.36
	-1	8.5	1.0655	11.395	2.475	0.07426	-0.0338	-0.3681	-0.0387	-0.00072	-3.07	-3.13	-1.54	-0.96
32	0	10.2	1.0993	11.764	2.514	0.07498	0.0000	0.0000	0.0000	0	0	0	0	0
	1	12.2	1.1313	12.166	2.562	0.07640	0.0320	0.4027	0.0482	0.00142	2.91	3.42	1.92	1.89
	-1	8	1.0992	11.763	2.510	0.07476	0.0000	-0.0006	-0.0035	-0.00023	0.00	-0.01	-0.14	-0.30
41	0	9.8	1.0993	11.764	2.514	0.07498	0.0000	0.0000	0.0000	0	0	0	0	0
	1	18.2	1.0997	11.767	2.527	0.07637	0.0004	0.0030	0.0134	0.00138	0.04	0.03	0.53	1.84
	-1	6.8	1.0990	11.762	2.510	0.07473	-0.0003	-0.0017	-0.0036	-0.00025	-0.02	-0.01	-0.14	-0.34
42	0	10.2	1.0993	11.764	2.514	0.07498	0.0000	0.0000	0.0000	0	0	0	0	0
	1	18.3	1.0995	11.766	2.519	0.07517	0.0002	0.0022	0.0058	0.00019	0.02	0.02	0.23	0.25
	-1	8.9	1.0990	11.762	2.511	0.07477	-0.0003	-0.0020	-0.0030	-0.00022	-0.02	-0.02	-0.12	-0.29
43	0	10.2	1.0993	11.764	2.514	0.07498	0.0000	0.0000	0.0000	0	0	0	0	0
	1	16.3	1.1019	11.796	2.520	0.07544	0.0026	0.0321	0.0068	0.00046	0.24	0.27	0.27	0.61
	-1	9	1.0990	11.762	2.511	0.07467	-0.0003	-0.0016	-0.0024	-0.00032	-0.03	-0.01	-0.09	-0.42
51	0	11.8	1.0993	11.764	2.514	0.07498	0.0000	0.0000	0.0000	0	0	0	0	0
	1	14.5	1.0993	11.765	2.516	0.07504	0.0001	0.0011	0.0023	0.00006	0.00	0.01	0.09	0.08
	-1	9.9	1.0932	11.712	2.488	0.07281	-0.0061	-0.0514	-0.0257	-0.00217	-0.55	-0.44	-1.02	-2.89
52	0	11.2	1.0993	11.764	2.514	0.07498	0.0000	0.0000	0.0000	0	0	0	0	0
	1	15.9	1.1085	11.874	2.534	0.07656	0.0093	0.1109	0.0200	0.00157	0.84	0.94	0.80	2.10
	-1	7.3	1.0984	11.757	2.511	0.07473	-0.0008	-0.0068	-0.0026	-0.00025	-0.07	-0.06	-0.10	-0.33
53	0	10.9	1.0993	11.764	2.514	0.07498	0.0000	0.0000	0.0000	0	0	0	0	0
	1	15.3	1.1002	11.772	2.516	0.07519	0.0009	0.0089	0.0029	0.00021	0.08	0.08	0.12	0.28
	-1	9.6	1.0977	11.740	2.510	0.07487	-0.0015	-0.0233	-0.0035	-0.00011	-0.14	-0.20	-0.14	-0.15
54	0	12.6	1.0993	11.764	2.514	0.07498	0.0000	0.0000	0.0000	0	0	0	0	0
	1	18.8	1.1034	11.823	2.523	0.07524	0.0041	0.0593	0.0094	0.00026	0.38	0.50	0.37	0.34
	-1	9.9	1.0977	11.754	2.476	0.07284	-0.0016	-0.0100	-0.0372	-0.00215	-0.14	-0.09	-1.48	-2.86
61	0	11.4	1.0993	11.764	2.514	0.07498	0.0000	0.0000	0.0000	0	0	0	0	0
	1	14.4	1.1017	11.782	2.592	0.07884	0.0025	0.0184	0.0786	0.00385	0.22	0.16	3.13	5.14
	-1	8.5	1.0913	11.725	2.388	0.06225	-0.0079	-0.0390	-0.1254	-0.01273	-0.72	-0.33	-4.99	-16.98
62	0	10.2	1.0993	11.764	2.514	0.07498	0.0000	0.0000	0.0000	0	0	0	0	0
	1	13.6	1.0991	11.765	2.544	0.07177	-0.0001	0.0011	0.0306	-0.00321	-0.01	0.01	1.22	-4.28

The table can be used to determine the effect of a change in the age distribution of a particular SourceTypeID on the total emissions. For example, the second row of the table shows the baseline HC emissions of 1.0993 tons/day. This value is the sum of all HC emissions for all SourceTypeIDs. The first row of the table shows that when the SourceTypeID=11 (motorcycles) age distribution changes from the median (second row) to the 10 percentile (first row), the total HC drops to 1.0980 tons/day, which is a change of -0.0013 tons/day and a relative change of -0.12%. Note that the age distributions of no other SourceTypeIDs were allowed to change while obtaining the result in the first row.

The baseline run and the two Experiment 2 MOVES runs provided the emissions information for the analysis of the SourceTypeYear table for a weekday in July 2011. The total HC, CO, NX, and PM emissions results for each of the SourceTypeIDs were obtained from the MOVES output. The emissions were summed for each emission type and test condition and are tabulated in Columns 4, 5, 6, and 7 of Table 5-19. The second column indicates the 10 percentile relative vehicle population as -1 (low fraction), the 50 percentile relative vehicle population as 0 (median fraction), and the 90 percentile relative vehicle population as 1 (high fraction).

Table 5-19. Analysis of Results from Experiment 2 for the SourceTypeYear Table

Source TypeID	SourceType Population Setting	Relative Population Fraction	ΣHC (t/day)	ΣCO (t/day)	ΣNX (t/day)	ΣPM2.5 (t/day)	ΔHC (t/day)	ΔCO (t/day)	ΔNX (t/day)	ΔPM2.5 (t/day)	ΔHC (%)	ΔCO (%)	ΔNX (%)	ΔPM2.5 (%)
	-1	0.0215	1.0896	11.763	2.514	0.07498	-0.0097	0.000	0.000	0.00000	-0.88	0.00	0.00	0.00
11	0	0.0366	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.0555	1.1113	11.764	2.514	0.07498	0.0121	0.000	0.000	0.00000	1.10	0.00	0.00	0.00
	-1	0.3355	1.0417	11.564	2.486	0.07470	-0.0575	-0.200	-0.027	-0.00028	-5.23	-1.70	-1.08	-0.38
21	0	0.4428	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.5903	1.1783	12.038	2.551	0.07537	0.0790	0.275	0.037	0.00039	7.19	2.33	1.49	0.52
	-1	0.2635	1.0447	12.012	2.567	0.07556	-0.0545	0.248	0.054	0.00058	-4.96	2.11	2.13	0.77
31	0	0.3643	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.5058	1.2178	12.377	2.587	0.07552	0.1185	0.613	0.073	0.00053	10.78	5.21	2.90	0.71
	-1	0.0264	0.9968	10.654	2.350	0.07208	-0.1025	-1.109	-0.163	-0.00291	-9.32	-9.43	-6.49	-3.87
32	0	0.1141	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.1508	1.1270	11.925	2.522	0.07500	0.0277	0.161	0.009	0.00001	2.52	1.37	0.34	0.02
	-1	0	1.0984	11.758	2.496	0.07388	-0.0009	-0.005	-0.018	-0.00111	-0.08	-0.05	-0.71	-1.48
41	0	0.0002	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.0008	1.0995	11.766	2.519	0.07534	0.0003	0.002	0.006	0.00036	0.03	0.02	0.23	0.48
	-1	0	1.0987	11.760	2.506	0.07452	-0.0005	-0.003	-0.008	-0.00047	-0.05	-0.03	-0.30	-0.62
42	0	0.0003	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.0013	1.0997	11.767	2.520	0.07537	0.0004	0.004	0.006	0.00039	0.04	0.03	0.25	0.52
	-1	0.0013	1.1006	11.767	2.527	0.07585	0.0013	0.003	0.013	0.00087	0.12	0.03	0.52	1.15
43	0	0.0032	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.0056	1.0987	11.766	2.507	0.07456	-0.0005	0.003	-0.006	-0.00042	-0.05	0.02	-0.25	-0.56
	-1	0.0001	1.0992	11.763	2.512	0.07490	-0.0001	-0.001	-0.001	-0.00008	-0.01	-0.01	-0.05	-0.11
51	0	0.0003	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.0006	1.0991	11.763	2.512	0.07484	-0.0002	-0.001	-0.002	-0.00014	-0.01	-0.01	-0.08	-0.19
	-1	0.0059	1.0957	11.682	2.519	0.07542	-0.0035	-0.082	0.006	0.00043	-0.32	-0.69	0.24	0.58
52	0	0.0153	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.0321	1.1073	11.927	2.514	0.07481	0.0081	0.164	0.000	-0.00018	0.73	1.39	0.01	-0.23
	-1	0.0005	1.0989	11.757	2.513	0.07497	-0.0004	-0.006	0.000	-0.00001	-0.03	-0.05	-0.01	-0.02
53	0	0.0013	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.004	1.1013	11.792	2.519	0.07528	0.0020	0.029	0.005	0.00030	0.18	0.24	0.22	0.40
	-1	0.0005	1.0955	11.707	2.505	0.07472	-0.0038	-0.057	-0.009	-0.00027	-0.34	-0.48	-0.36	-0.36
54	0	0.0045	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.0103	1.1017	11.785	2.515	0.07501	0.0025	0.022	0.001	0.00002	0.22	0.18	0.04	0.03
	-1	0.0015	1.0976	11.752	2.482	0.07334	-0.0016	-0.011	-0.032	-0.00165	-0.15	-0.10	-1.25	-2.20
61	0	0.0045	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.0096	1.0951	11.746	2.430	0.07064	-0.0042	-0.018	-0.084	-0.00435	-0.38	-0.15	-3.33	-5.80
	-1	0.0023	1.1038	11.776	2.561	0.07783	0.0045	0.012	0.047	0.00284	0.41	0.10	1.89	3.79
62	0	0.0057	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.0215	1.1113	11.817	2.640	0.08261	0.0121	0.054	0.126	0.00762	1.10	0.46	5.02	10.16

The table can be used to determine the effect of a change in the relative population fraction of a particular SourceTypeID on the total emissions. For example, the second row of the table shows the baseline HC emissions of 1.0993 tons/day, as it always is for the baseline in this study. This value is the sum of all HC emissions for all SourceTypeIDs. The first row of the table shows that when the SourceTypeID=11 (motorcycles) the relative population fraction changes from the median of 0.0366 (second row) to the 10 percentile value of 0.0215 (first row), the total HC drops to 1.0896 tons/day, which is a change of -0.0097 tons/day and a relative change of -0.88%. Note that the relative population fractions of no other SourceTypeIDs were allowed to change while obtaining the result in the first row.

The baseline run and the two Experiment 3 MOVES runs provided the emissions information for the analysis of the HPMSVtypeYear table for a weekday in July 2011. The total HC, CO, NX, and PM emissions results for each of the SourceTypeIDs were obtained from the MOVES output. The method of determining the influences on emissions for this experiment are somewhat different than for Experiments 1 and 2 since the SourceTypeID VMT distributions are moved in the same direction for all SourceTypeIDs within their common HPMSVtypeID. In Table 5-20 the SourceTypeIDs that are in a common HPMSVtypeID have the same background color in the third column. For example, SourceTypeID = 31 and 32 are both in HPMSVtypeID=30 and are shown with a green background in the table. Again, the emissions were summed for each emission type and test condition and are tabulated in Columns 4, 5, 6, and 7 of Table 5-20. The second column indicates the 10 percentile relative VMT fraction as -1 (low fraction), the 50 percentile relative VMT fraction as 0 (median fraction), and the 90 percentile relative VMT fraction as 1 (high fraction).

The table can be used to determine the effect of a change in the relative VMT fraction of a particular SourceTypeID on the total emissions. For example, the eighth row of the table shows the baseline HC emissions of 1.0993 tons/day, as it always is for the baseline in this study. This value is the sum of all HC emissions for all SourceTypeIDs. The seventh row of the table shows that when the SourceTypeID=31 (passenger trucks) the relative VMT fraction changes from the median of 0.375 (eighth row) to the 10 percentile value of 0.192 (seventh row), the total HC drops to 1.0276 tons/day, which is a change of -0.0717 tons/day and a relative change of -3.48%. Note that the relative VMT fractions of SourceTypeID=32 was allowed to change at the same time, however by acquiring the separate MOVES outputs for SourceTypeID=31 and SourceTypeID=32, the separate effects on emissions are obtained.

Table 5-20. Analysis of Results from Experiment 3 for the HPMSVtypeYear Table

Source TypeID	HPMSType VMT Setting	HPMSVtype VMT Fraction	ΣHC (t/day)	ΣCO (t/day)	ΣNX (t/day)	ΣPM2.5 (t/day)	ΔHC (t/day)	ΔCO (t/day)	ΔNX (t/day)	ΔPM2.5 (t/day)	ΔHC (%)	ΔCO (%)	ΔNX (%)	ΔPM2.5 (%)
	-1	0.003	1.0943	11.715	2.512	0.07491	-0.0050	-0.049	-0.002	-0.00007	-0.45	-0.41	-0.08	-0.09
11	0	0.006	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.013	1.1134	11.902	2.519	0.07518	0.0142	0.138	0.006	0.00020	1.29	1.17	0.23	0.26
	-1	0.355	1.0610	10.704	2.385	0.07357	-0.0383	-1.059	-0.128	-0.00142	-3.48	-9.01	-5.11	-1.89
21	0	0.505	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.719	1.1538	13.273	2.696	0.07700	0.0545	1.509	0.183	0.00202	4.96	12.83	7.27	2.69
	-1	0.192	1.0276	10.124	2.261	0.07268	-0.0717	-1.640	-0.252	-0.00230	-6.52	-13.94	-10.04	-3.07
31	0	0.375	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.584	1.1809	13.631	2.801	0.07760	0.0816	1.867	0.287	0.00262	7.43	15.87	11.44	3.49
	-1	0.192	1.0743	11.264	2.415	0.07300	-0.0250	-0.499	-0.098	-0.00199	-2.27	-4.24	-3.91	-2.65
32	0	0.375	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.584	1.1277	12.332	2.626	0.07724	0.0285	0.569	0.112	0.00226	2.59	4.83	4.46	3.01
	-1	0.002	1.0989	11.761	2.506	0.07449	-0.0004	-0.002	-0.008	-0.00049	-0.03	-0.02	-0.32	-0.66
41	0	0.004	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.007	1.0997	11.766	2.523	0.07554	0.0004	0.003	0.009	0.00056	0.04	0.02	0.36	0.75
	-1	0.002	1.0989	11.761	2.508	0.07467	-0.0004	-0.002	-0.005	-0.00032	-0.03	-0.02	-0.20	-0.42
42	0	0.004	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.007	1.0997	11.766	2.519	0.07535	0.0004	0.002	0.006	0.00036	0.04	0.02	0.23	0.48
	-1	0.002	1.0982	11.758	2.505	0.07440	-0.0010	-0.006	-0.009	-0.00059	-0.09	-0.05	-0.35	-0.78
43	0	0.004	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.007	1.1004	11.770	2.524	0.07565	0.0012	0.007	0.010	0.00067	0.11	0.06	0.40	0.89
	-1	0.011	1.0987	11.760	2.507	0.07450	-0.0006	-0.003	-0.007	-0.00048	-0.05	-0.03	-0.27	-0.64
51	0	0.03	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.048	1.0998	11.767	2.520	0.07546	0.0005	0.003	0.007	0.00047	0.05	0.03	0.26	0.63
	-1	0.011	1.0857	11.626	2.431	0.07045	-0.0136	-0.137	-0.082	-0.00453	-1.23	-1.17	-3.28	-6.04
52	0	0.03	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.048	1.1126	11.898	2.595	0.07944	0.0133	0.135	0.081	0.00445	1.21	1.15	3.22	5.94
	-1	0.011	1.0980	11.752	2.507	0.07458	-0.0012	-0.011	-0.007	-0.00040	-0.11	-0.10	-0.27	-0.53
53	0	0.03	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.048	1.1005	11.775	2.520	0.07538	0.0012	0.011	0.007	0.00039	0.11	0.10	0.26	0.52
	-1	0.011	1.0973	11.724	2.506	0.07474	-0.0019	-0.039	-0.008	-0.00024	-0.18	-0.34	-0.32	-0.32
54	0	0.03	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.048	1.1012	11.802	2.521	0.07522	0.0019	0.039	0.008	0.00024	0.17	0.33	0.31	0.31
	-1	0.016	1.0899	11.714	2.328	0.06532	-0.0094	-0.050	-0.186	-0.00966	-0.85	-0.42	-7.38	-12.89
61	0	0.051	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.11	1.1151	11.848	2.828	0.09134	0.0159	0.085	0.314	0.01636	1.44	0.72	12.50	21.82
	-1	0.016	1.0645	11.647	2.150	0.05310	-0.0348	-0.116	-0.364	-0.02189	-3.16	-0.99	-14.48	-29.19
62	0	0.051	1.0993	11.764	2.514	0.07498	0.0000	0.000	0.000	0	0	0	0	0
	1	0.11	1.1581	11.960	3.130	0.11204	0.0589	0.197	0.616	0.03705	5.36	1.67	24.51	49.42

The 10 runs of Experiment 4 provided the emissions information for the analysis of the RoadTypeDistribution table for a weekday in July 2011. The total HC, CO, NX, and PM emissions results for each of the SourceTypeIDs were obtained from the MOVES output. A linear regression was performed[12] in SAS for each of the SourceTypeIDs using a model statement derived from the canonical polynomial[13] associated with the 10-run, 4-component mixture design used to generate the data. That is, the model statement was constructed to be able to obtain regression coefficients for main effects and two-factor interactions with a zero

[12] /proj1/CRC_A84_MOVES_Sensitivity/Sensitivity_Analysis/DOE/model_exp04.sas
[13] Cornell, J., Experiments with Mixtures, John Wiley & Sons, Inc., Third Edition, 2002, page 25.

intercept. The regression results indicated that for all SourceTypeIDs the coefficients of all two-factor interactions were zero, as was expected. Thus, for each SourceTypeID the model simplifies to a linear combination of the relative VMT fraction for each of the four RoadTypeIDs and the emissions factor for 100% of the VMT for each of the RoadTypeIDs. The third, fourth, fifth, and sixth columns of the first fifty-two rows of Table 5-21 give the emission factors determined by the regression for 100% of the VMT for each of the RoadTypeIDs.

To arrive at the overall effect of changes to the distribution of VMT on the four different RoadTypeIDs, the third, fourth, fifth, and sixth columns of the first fifty-two rows of Table 5-21 are summed by RoadTypeID to produce the four rows at the bottom of the table labeled "Fleet" with a RoadTypeID indicator. The last row of the table gives the baseline values. A comparison of the last five rows of the table reveals the effect of VMT on different RoadTypeIDs. For example, if all VMT were on RoadTypeID=5 (urban restricted), the total NOx would increase from the baseline of 2.514 tons/day to 3.171 tons/day, an increase of 0.657 tons/day, and a relative increase of 26.15%.

The emissions values can also be used to calculate the fleet emissions for any distribution of VMT on the RoadTypeIDs by using the linear combination relationship described earlier. We will use that relationship later when evaluating the emissions effects of the AvgSpeedDistribution table.

The 32 runs of Experiment 5 provided the emissions information for the analysis of the AvgSpeedDistribution table for a weekday in July 2011. The total HC, CO, NX, and PM emissions results for each of the SourceTypeIDs were obtained from the MOVES output. A linear regression was performed[14] in SAS for each of the SourceTypeIDs using a model statement derived from the alias structure[15] associated with the 32-run, 13-effect (the 13 SpeedDistClusters) fractional factorial design used to generate the data. The model statement was constructed to be able to obtain regression coefficients for the 13 main effects, 15 two-factor interactions, and 3 three-factor interactions that were consistent with the alias structure analysis. The regression results indicated that for all SpeedDistClusters the coefficients of the two-factor interactions and the three-factor interactions were zero, as was expected.

---

[14] /proj1/CRC_A84_MOVES_Sensitivity/Sensitivity_Analysis/DOE/model_exp05.sas
[15] The alias structure was determined by the SAS program /proj1/CRC_A84_MOVES_Sensitivity/Sensitivity_Analysis/DOE/fractional_factorial.sas using the experimental design generators found in C.F.J. Wu, M. Hamada, Experiments: Planning, Analysis, and Parameter Design Optimization, John Wiley & Sons, Inc., 2000, page 195, Table 4A.3 for the 2^{13-8}_{IV} design.

Table 5-21. Analysis of Results from Experiment 4 for the RoadTypeDistribution Table

Source TypeID	RoadTypeID Fraction	ΣHC (t/day)	ΣCO (t/day)	ΣNX (t/day)	ΣPM2.5 (t/day)	ΔHC (t/day)	ΔCO (t/day)	ΔNX (t/day)	ΔPM2.5 (t/day)	ΔHC (%)	ΔCO (%)	ΔNX (%)	ΔPM2.5 (%)
11	100% RoadTypeID=2	0.0318	0.114	0.005	0.00022								
	100% RoadTypeID=3	0.0339	0.114	0.005	0.00016								
	100% RoadTypeID=4	0.0358	0.127	0.005	0.00027								
	100% RoadTypeID=5	0.0397	0.116	0.004	0.00017								
21	100% RoadTypeID=2	0.3507	5.402	0.532	0.00729								
	100% RoadTypeID=3	0.3562	3.748	0.515	0.00516								
	100% RoadTypeID=4	0.3814	4.833	0.624	0.00791								
	100% RoadTypeID=5	0.4283	4.609	0.664	0.00604								
31	100% RoadTypeID=2	0.4278	6.570	0.683	0.00777								
	100% RoadTypeID=3	0.4331	4.712	0.663	0.00572								
	100% RoadTypeID=4	0.4563	5.788	0.781	0.00853								
	100% RoadTypeID=5	0.5091	5.372	0.804	0.00691								
32	100% RoadTypeID=2	0.1322	2.001	0.228	0.00436								
	100% RoadTypeID=3	0.1380	1.599	0.237	0.00426								
	100% RoadTypeID=4	0.1455	1.828	0.281	0.00527								
	100% RoadTypeID=5	0.1665	1.841	0.316	0.00594								
41	100% RoadTypeID=2	0.0007	0.004	0.018	0.00094								
	100% RoadTypeID=3	0.0008	0.005	0.017	0.00105								
	100% RoadTypeID=4	0.0010	0.006	0.019	0.00117								
	100% RoadTypeID=5	0.0012	0.007	0.021	0.00155								
42	100% RoadTypeID=2	0.0007	0.005	0.013	0.00069								
	100% RoadTypeID=3	0.0008	0.005	0.010	0.00068								
	100% RoadTypeID=4	0.0011	0.007	0.014	0.00089								
	100% RoadTypeID=5	0.0011	0.006	0.011	0.00080								
43	100% RoadTypeID=2	0.0021	0.026	0.026	0.00136								
	100% RoadTypeID=3	0.0024	0.022	0.019	0.00126								
	100% RoadTypeID=4	0.0032	0.029	0.027	0.00182								
	100% RoadTypeID=5	0.0033	0.024	0.023	0.00162								
51	100% RoadTypeID=2	0.0006	0.005	0.011	0.00060								
	100% RoadTypeID=3	0.0008	0.005	0.010	0.00070								
	100% RoadTypeID=4	0.0010	0.006	0.012	0.00081								
	100% RoadTypeID=5	0.0013	0.008	0.014	0.00111								
52	100% RoadTypeID=2	0.0225	0.361	0.127	0.00597								
	100% RoadTypeID=3	0.0286	0.362	0.128	0.00700								
	100% RoadTypeID=4	0.0339	0.429	0.162	0.00886								
	100% RoadTypeID=5	0.0456	0.458	0.189	0.01072								
53	100% RoadTypeID=2	0.0018	0.027	0.010	0.00051								
	100% RoadTypeID=3	0.0023	0.026	0.010	0.00059								
	100% RoadTypeID=4	0.0028	0.033	0.013	0.00077								
	100% RoadTypeID=5	0.0037	0.034	0.015	0.00091								
54	100% RoadTypeID=2	0.0043	0.092	0.014	0.00035								
	100% RoadTypeID=3	0.0048	0.067	0.012	0.00037								
	100% RoadTypeID=4	0.0058	0.091	0.015	0.00048								
	100% RoadTypeID=5	0.0075	0.086	0.015	0.00054								
61	100% RoadTypeID=2	0.0114	0.067	0.266	0.01258								
	100% RoadTypeID=3	0.0149	0.082	0.266	0.01477								
	100% RoadTypeID=4	0.0181	0.095	0.295	0.01666								
	100% RoadTypeID=5	0.0249	0.125	0.346	0.02452								
62	100% RoadTypeID=2	0.0371	0.139	0.490	0.02699								
	100% RoadTypeID=3	0.0526	0.180	0.522	0.03210								
	100% RoadTypeID=4	0.0666	0.215	0.592	0.03616								
	100% RoadTypeID=5	0.0996	0.299	0.747	0.05470								
Fleet	100% RoadTypeID=2	1.0235	14.813	2.424	0.06963	-0.0758	3.050	-0.090	-0.00535	-6.89	25.93	-3.58	-7.14
	100% RoadTypeID=3	1.0691	10.929	2.413	0.07383	-0.0301	-0.835	-0.101	-0.00116	-2.74	-7.10	-4.00	-1.54
	100% RoadTypeID=4	1.1524	13.486	2.840	0.08958	0.0531	1.722	0.326	0.01460	4.83	14.64	12.99	19.47
	100% RoadTypeID=5	1.3317	12.986	3.171	0.11553	0.2325	1.222	0.657	0.04054	21.15	10.39	26.15	54.07
	Baseline	1.0993	11.764	2.514	0.07498	0	0	0	0	0	0	0	0

Table 5-22 contains the results of the regressions as the emissions in Columns 14, 15, 16, and 17. The analysis of Experiment 5 is more complex than the first four experiments. Table 5-22 gives the results and effects for four separate groups. Each group represents the case when 100% of the VMT is driven in just one RoadTypeID. For example, the top group in the table shows the case where all VMT is driven on RoadTypeID=2 (rural restricted). This group has its own baseline, which is shown by the first row in the group. The next three subgroups show the results for the three SpeedDistClusters (11, 12, 13) that make up the RoadTypeID=2 group. Within each of the sub-groups the speed distributions are moved to the 10 percentile (indicated by -1) and to the 90 percentile (indicated by 1) independently. Because the RoadTypeID=2 group does not contain SpeedDistClusters 1, 2, 3, 4, 8, 9, 10, 5, 6, and 7, the indicators in those columns are missing as is designated by the dot. The columns to the right of the SpeedDistCluster indicators show the effects on emissions of changing the speed distributions of the three SpeedDistClusters when 100% of the VMT is driven on RoadTypeID=2. The emissions effects when each of the other three RoadTypeIDs are assumed to have 100% of the VMT are shown by the other three groups in Table 5-22.

However, the results in Table 5-22 do not reflect most real situations since the fleet of a county will rarely drive on just one type of roadway for 100% of its VMT. The effect of the VMT driven by the fleet on all four RoadTypeIDs needs to be brought to the analysis. This is done in Table 5-23. Table 5-23 has the same entries as Table 5-22 except for the last four columns, which are the calculated emissions effects for a fleet that is made up of 14% of VMT on RoadTypeID=2, 37% of VMT on RoadTypeID=3, 18% of VMT on RoadTypeID=4, and 31% of VMT on RoadTypeID=5. The values in these columns were obtained by multiplying the RoadTypeID VMT percentage for the group by the corresponding percent emissions change effect from Table 5-22. For example, Table 5-22 indicated that the HC emissions change for the -1 level of SpeedDistCluster=2 in RoadTypeID=3 is 2.68%. But that change would be observed only if 100% of the VMT would be driven on RoadTypeID=3. When only 37% of the VMT is driven on RoadTypeID=3, the effect on the HC emissions can be only 37% as large, i.e. 0.99% (=2.68% * 0.37), which is the value shown for the corresponding cell in Table 5-23.

Thus, the effects on the fleet are represented by Table 5-23. The multiplication of the emissions effect from the AvgSpeedDistribution table result from Experiment 5 with the emissions effect from the RoadTypeDistribution table result from Experiment 4 is an interaction between those two tables that has a strong influence on the calculated emissions.

Table 5-22. Analysis of Weekday Results from Experiment 5 for the AvgSpeedDistribution Table

Average SpeedBinID over all SourceTypeIDs for perturbed SpeedDistCluster	Road TypeID 2			Road TypeID 3			Road TypeID 4				Road TypeID 5														
	SpeedDistCluster													Assuming 100% of VMT is only in RoadTypeID that was Perturbed											
	11	12	13	1	2	3	4	8	9	10	5	6	7	**ΣHC (t/day)**	**ΣCO (t/day)**	**ΣNX (t/day)**	**ΣPM2.5 (t/day)**	**ΔHC (t/day)**	**ΔCO (t/day)**	**ΔNX (t/day)**	**ΔPM2.5 (t/day)**	**ΔHC (%)**	**ΔCO (%)**	**ΔNX (%)**	**ΔPM2.5 (%)**
100% VMT only RoadTypeID=2:																									
Baseline	0	0	0	.	.	.	.	.	.	.	.	.	.	1.0235	14.813	2.424	0.06963	0	0	0	0	0	0	0	0
13.8	-1	0	0	.	.	.	.	.	.	.	.	.	.	1.0153	12.999	2.318	0.07026	-0.0082	-1.814	-0.106	0.00063	-0.80	-12.25	-4.36	0.90
15.4	1	0	0	.	.	.	.	.	.	.	.	.	.	1.0235	14.813	2.424	0.06963	0	0	0	0	0.00	0.00	0.00	0.00
13.8	0	-1	0	.	.	.	.	.	.	.	.	.	.	1.0218	14.531	2.403	0.06955	-0.0017	-0.282	-0.020	-0.00008	-0.16	-1.91	-0.84	-0.11
15.4	0	1	0	.	.	.	.	.	.	.	.	.	.	1.0235	14.813	2.424	0.06963	0	0	0	0	0.00	0.00	0.00	0.00
13.7	0	0	-1	.	.	.	.	.	.	.	.	.	.	1.0235	14.813	2.424	0.06963	0	0	0	0	0.00	0.00	0.00	0.00
15.4	0	0	1	.	.	.	.	.	.	.	.	.	.	1.0235	14.813	2.424	0.06963	0	0	0	0	0.00	0.00	0.00	0.00
100% VMT only RoadTypeID=3:																									
Baseline	.	.	.	0	0	0	.	.	.	.	.	.	.	1.0691	10.929	2.413	0.07383	0	0	0	0	0	0	0	0
9.0	.	.	.	-1	0	0	.	.	.	.	.	.	.	1.0717	10.931	2.418	0.07460	0.0025	0.003	0.005	0.00077	0.24	0.02	0.20	1.05
10.7	.	.	.	1	0	0	.	.	.	.	.	.	.	1.0653	10.889	2.404	0.07364	-0.0039	-0.040	-0.009	-0.00019	-0.36	-0.36	-0.38	-0.26
8.8	.	.	.	0	-1	0	.	.	.	.	.	.	.	1.0978	10.928	2.400	0.08298	0.0287	-0.001	-0.012	0.00915	2.68	-0.01	-0.51	12.39
10.6	.	.	.	0	1	0	.	.	.	.	.	.	.	1.0521	10.711	2.334	0.07265	-0.0170	-0.218	-0.079	-0.00117	-1.59	-1.99	-3.28	-1.59
9.0	.	.	.	0	0	-1	.	.	.	.	.	.	.	1.0726	10.932	2.415	0.07523	0.0035	0.003	0.002	0.00140	0.33	0.03	0.08	1.90
10.9	.	.	.	0	0	1	.	.	.	.	.	.	.	1.0659	10.898	2.405	0.07373	-0.0033	-0.031	-0.007	-0.00010	-0.30	-0.28	-0.31	-0.14
100% VMT only RoadTypeID=4:																									
Baseline	.	.	.	.	.	.	0	0	0	0	.	.	.	1.1524	13.486	2.840	0.08958	0	0	0	0	0	0	0	0
7.7	.	.	.	.	.	.	-1	0	0	0	.	.	.	1.1524	13.486	2.840	0.08958	0	0	0	0	0.00	0.00	0.00	0.00
13.5	.	.	.	.	.	.	1	0	0	0	.	.	.	1.0932	13.763	2.613	0.08190	-0.0592	0.277	-0.227	-0.00768	-5.13	2.06	-7.98	-8.58
8.4	.	.	.	.	.	.	0	-1	0	0	.	.	.	1.1524	13.486	2.840	0.08958	0	0	0	0	0.00	0.00	0.00	0.00
13.6	.	.	.	.	.	.	0	1	0	0	.	.	.	1.1264	13.426	2.751	0.08595	-0.0259	-0.060	-0.089	-0.00364	-2.25	-0.44	-3.13	-4.06
10.9	.	.	.	.	.	.	0	0	-1	0	.	.	.	1.1545	13.449	2.841	0.09015	0.0022	-0.037	0.001	0.00056	0.19	-0.28	0.02	0.63
13.8	.	.	.	.	.	.	0	0	1	0	.	.	.	1.1277	13.545	2.752	0.08588	-0.0246	0.059	-0.088	-0.00370	-2.14	0.44	-3.11	-4.13
8.3	.	.	.	.	.	.	0	0	0	-1	.	.	.	1.1550	13.495	2.846	0.08992	0.0027	0.009	0.006	0.00034	0.23	0.07	0.20	0.38
13.6	.	.	.	.	.	.	0	0	0	1	.	.	.	1.1513	13.492	2.840	0.08940	-0.0010	0.006	0.000	-0.00018	-0.09	0.04	0.01	-0.20
100% VMT only RoadTypeID=5:																									
Baseline	.	.	.	.	.	.	.	.	.	.	0	0	0	1.3317	12.986	3.171	0.11553	0	0	0	0	0	0	0	0
5.7	.	.	.	.	.	.	.	.	.	.	-1	0	0	1.3317	12.986	3.171	0.11553	0	0	0	0	0.00	0.00	0.00	0.00
8.8	.	.	.	.	.	.	.	.	.	.	1	0	0	1.1675	11.742	2.600	0.09433	-0.1642	-1.244	-0.571	-0.02120	-12.33	-9.58	-18.00	-18.35
6.0	.	.	.	.	.	.	.	.	.	.	0	-1	0	1.3317	12.986	3.171	0.11553	0	0	0	0	0.00	0.00	0.00	0.00
9.2	.	.	.	.	.	.	.	.	.	.	0	1	0	1.3064	12.799	3.100	0.11223	-0.0253	-0.187	-0.071	-0.00330	-1.90	-1.44	-2.25	-2.86
5.8	.	.	.	.	.	.	.	.	.	.	0	0	-1	1.3317	12.986	3.171	0.11553	0	0	0	0	0.00	0.00	0.00	0.00
9.1	.	.	.	.	.	.	.	.	.	.	0	0	1	1.3146	12.860	3.134	0.11325	-0.0171	-0.126	-0.037	-0.00227	-1.28	-0.97	-1.17	-1.97

Table 5-23. Application of Weekday Results from Experiments 4 and 5 for the Fleet

Average SpeedBinID over all SourceTypeIDs for perturbed SpeedDistCluster	Road TypeID 2			Road TypeID 3			Road TypeID 4				Road TypeID 5														
	SpeedDistCluster													Assuming 100% of VMT is only in RoadTypeID that was Perturbed								Mix: 2=14%, 3=37%, 4=18%, 5=31%			
	11	12	13	1	2	3	4	8	9	10	5	6	7	**ΣHC (t/day)**	**ΣCO (t/day)**	**ΣNX (t/day)**	**ΣPM2.5 (t/day)**	**ΔHC (t/day)**	**ΔCO (t/day)**	**ΔNX (t/day)**	**ΔPM2.5 (t/day)**	**ΔHC (%)**	**ΔCO (%)**	**ΔNX (%)**	**ΔPM2.5 (%)**
100% VMT only RoadTypeID=2:																									
Baseline	0	0	0	.	.	.	.	.	.	.	.	.	.	1.0235	14.813	2.424	0.06963	0	0	0	0	0	0	0	0
13.8	-1	0	0	.	.	.	.	.	.	.	.	.	.	1.0153	12.999	2.318	0.07026	-0.0082	-1.814	-0.106	0.00063	-0.11	-1.71	-0.61	0.13
15.4	1	0	0	.	.	.	.	.	.	.	.	.	.	1.0235	14.813	2.424	0.06963	0	0	0	0	0.00	0.00	0.00	0.00
13.8	0	-1	0	.	.	.	.	.	.	.	.	.	.	1.0218	14.531	2.403	0.06955	-0.0017	-0.282	-0.020	-0.00008	-0.02	-0.27	-0.12	-0.02
15.4	0	1	0	.	.	.	.	.	.	.	.	.	.	1.0235	14.813	2.424	0.06963	0	0	0	0	0.00	0.00	0.00	0.00
13.7	0	0	-1	.	.	.	.	.	.	.	.	.	.	1.0235	14.813	2.424	0.06963	0	0	0	0	0.00	0.00	0.00	0.00
15.4	0	0	1	.	.	.	.	.	.	.	.	.	.	1.0235	14.813	2.424	0.06963	0	0	0	0	0.00	0.00	0.00	0.00
100% VMT only RoadTypeID=3:																									
Baseline	.	.	.	0	0	0	.	.	.	.	.	.	.	1.0691	10.929	2.413	0.07383	0	0	0	0	0	0	0	0
9.0	.	.	.	-1	0	0	.	.	.	.	.	.	.	1.0717	10.931	2.418	0.07460	0.0025	0.003	0.005	0.00077	0.09	0.01	0.07	0.39
10.7	.	.	.	1	0	0	.	.	.	.	.	.	.	1.0653	10.889	2.404	0.07364	-0.0039	-0.040	-0.009	-0.00019	-0.13	-0.13	-0.14	-0.10
8.8	.	.	.	0	-1	0	.	.	.	.	.	.	.	1.0978	10.928	2.400	0.08298	0.0287	-0.001	-0.012	0.00915	0.99	0.00	-0.19	4.59
10.6	.	.	.	0	1	0	.	.	.	.	.	.	.	1.0521	10.711	2.334	0.07265	-0.0170	-0.218	-0.079	-0.00117	-0.59	-0.74	-1.21	-0.59
9.0	.	.	.	0	0	-1	.	.	.	.	.	.	.	1.0726	10.932	2.415	0.07523	0.0035	0.003	0.002	0.00140	0.12	0.01	0.03	0.70
10.9	.	.	.	0	0	1	.	.	.	.	.	.	.	1.0659	10.898	2.405	0.07373	-0.0033	-0.031	-0.007	-0.00010	-0.11	-0.10	-0.11	-0.05
100% VMT only RoadTypeID=4:																									
Baseline	.	.	.	.	.	.	0	0	0	0	.	.	.	1.1524	13.486	2.840	0.08958	0	0	0	0	0	0	0	0
7.7	.	.	.	.	.	.	-1	0	0	0	.	.	.	1.1524	13.486	2.840	0.08958	0	0	0	0	0.00	0.00	0.00	0.00
13.5	.	.	.	.	.	.	1	0	0	0	.	.	.	1.0932	13.763	2.613	0.08190	-0.0592	0.277	-0.227	-0.00768	-0.92	0.37	-1.44	-1.54
8.4	.	.	.	.	.	.	0	-1	0	0	.	.	.	1.1524	13.486	2.840	0.08958	0	0	0	0	0.00	0.00	0.00	0.00
13.6	.	.	.	.	.	.	0	1	0	0	.	.	.	1.1264	13.426	2.751	0.08595	-0.0259	-0.060	-0.089	-0.00364	-0.40	-0.08	-0.56	-0.73
10.9	.	.	.	.	.	.	0	0	-1	0	.	.	.	1.1545	13.449	2.841	0.09015	0.0022	-0.037	0.001	0.00056	0.03	-0.05	0.00	0.11
13.8	.	.	.	.	.	.	0	0	1	0	.	.	.	1.1277	13.545	2.752	0.08588	-0.0246	0.059	-0.088	-0.00370	-0.38	0.08	-0.56	-0.74
8.3	.	.	.	.	.	.	0	0	0	-1	.	.	.	1.1550	13.495	2.846	0.08992	0.0027	0.009	0.006	0.00034	0.04	0.01	0.04	0.07
13.6	.	.	.	.	.	.	0	0	0	1	.	.	.	1.1513	13.492	2.840	0.08940	-0.0010	0.006	0.000	-0.00018	-0.02	0.01	0.00	-0.04
100% VMT only RoadTypeID=5:																									
Baseline	.	.	.	.	.	.	.	.	.	.	0	0	0	1.3317	12.986	3.171	0.11553	0	0	0	0	0	0	0	0
5.7	.	.	.	.	.	.	.	.	.	.	-1	0	0	1.3317	12.986	3.171	0.11553	0	0	0	0	0.00	0.00	0.00	0.00
8.8	.	.	.	.	.	.	.	.	.	.	1	0	0	1.1675	11.742	2.600	0.09433	-0.1642	-1.244	-0.571	-0.02120	-3.82	-2.97	-5.58	-5.69
6.0	.	.	.	.	.	.	.	.	.	.	0	-1	0	1.3317	12.986	3.171	0.11553	0	0	0	0	0.00	0.00	0.00	0.00
9.2	.	.	.	.	.	.	.	.	.	.	0	1	0	1.3064	12.799	3.100	0.11223	-0.0253	-0.187	-0.071	-0.00330	-0.59	-0.45	-0.70	-0.89
5.8	.	.	.	.	.	.	.	.	.	.	0	0	-1	1.3317	12.986	3.171	0.11553	0	0	0	0	0.00	0.00	0.00	0.00
9.1	.	.	.	.	.	.	.	.	.	.	0	0	1	1.3146	12.860	3.134	0.11325	-0.0171	-0.126	-0.037	-0.00227	-0.40	-0.30	-0.36	-0.61

5.5 Summary Results for Primary5 Inputs

The section summarizes the Task 2 approach and presents overall emission sensitivity results for each of the Primary5 inputs, as well as a comparison across inputs. For each of the five inputs listed, county data were selected to represent the 10th and 90th percentile distributions, as well as the median. Although Task 1 focused on quantifying distributions by "angles," for Task 2 the 10th and 90th percentile levels were chosen based on mean statistics that were more intuitive, as shown in Table 5-24.

Table 5-24. Statistics Used for Choosing 10th / 90th Percentile Levels

Input	Statistic Used
Age Distribution	Average age
Vehicle Population	Fraction of total population
VMT	Fraction of total VMT
Road Type Distribution	100% of each Road Type, National default mix
Average Speed Distribution	Mean of speed bins

Following analysis of the input data, the next task evaluated the sensitivity of MOVES emission results to the range of inputs observed in the submissions. This evaluation was performed at the source type level, so that the importance of variations in each source type on overall MOVES emission predictions could be evaluated. The impact of changing each source type's input from the 10^{th} to 90^{th} percentile level on total daily HC, CO, NOx and PM emission predictions was analyzed for a typical July day in Montgomery County, Texas (outside Houston). When one input was varied, all other inputs were held constant at median levels (for the remaining primary inputs) or default levels (all other inputs), so the magnitude of change for each input / source type combination could be evaluated and compared. Evaluating the change in total daily emissions, across all source types and emission processes, provided the proper weighting of inventory contribution from each source type and emission processes in the sensitivity analysis. The overall goal was to establish which source type / input combinations are most influential on total emissions, to help guide MOVES users towards data collection efforts that will yield the most improvement in local inventories.

Age Distribution

MOVES runs were conducted to estimate the impact from changing the age distribution of each source type from the 10^{th} the 90^{th} percentile average age level on total daily HC, CO, NOx and PM emissions. All other inputs were held at median or default levels. The results are shown in Figure 5-21, expressed as the relative change in total ton/day emissions across all emissions processes, source types and hours of the day. The sum of the differences for all source types equals the total difference for a pollutant, if all source types were changed from 10th to 90th percentile levels at the same time.

Figure 5-21. Change in Total Ton/Day Emissions Going from 10th to 90th Percentile Age Distribution, by Source Type

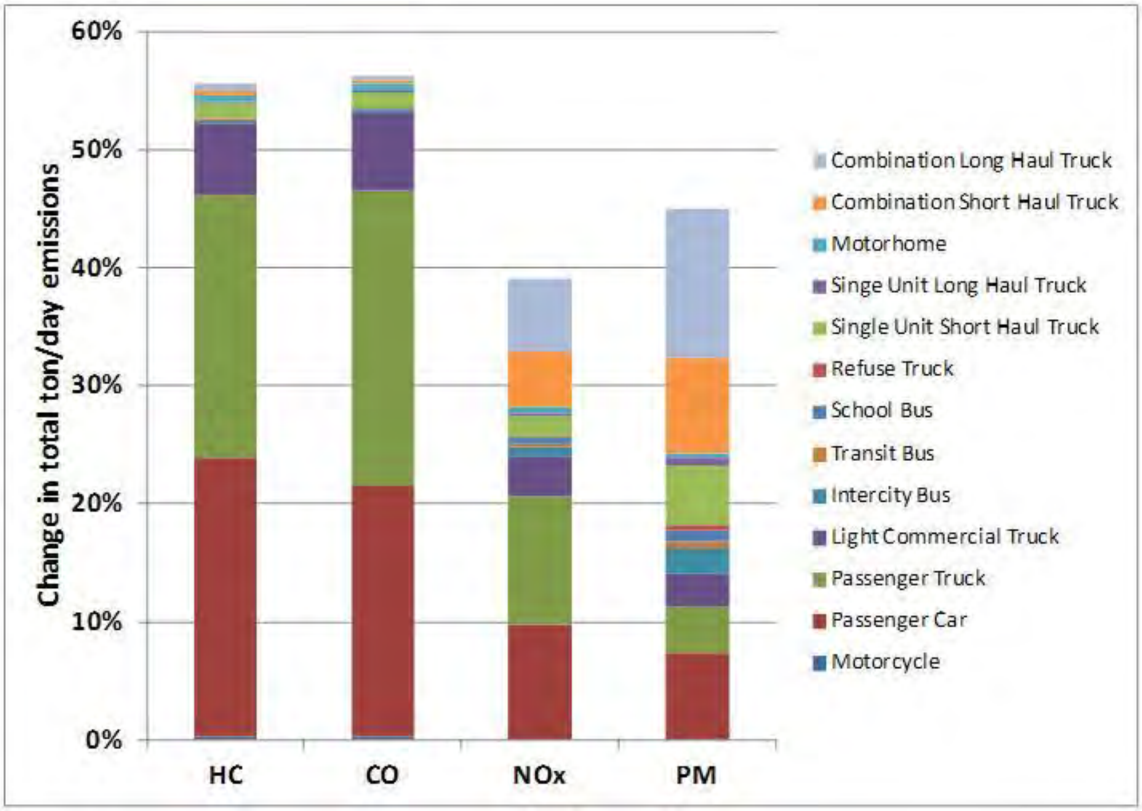


Figure 5-21 shows that HC and CO emissions increase about 55 percent, NO emissions 40 percent, and PM emissions 45 percent in total. Broken down by source type, passenger car and light truck (passenger and light commercial) comprise the majority of this increase for HC and CO, and about half for NOx. Heavier trucks, particularly combination trucks, contribute much of the change for NOx and the majority for PM. These results are in line with the overall inventory contributions of these sources.

Vehicle Population

Inputs for the population sensitivity case levels were developed by establishing a constant total vehicle population (across all source types), based on median source type population fractions. 10th and 90th percentile populations for a given source type were calculated by applying the 10th and 90th percentile population fractions for that source type to the median total population. The resulting lower/higher populations for a given source type were then subtracted/added to the median total. This allowed the impact of changing only the population of a single source type to be evaluated.

Results by source type are shown in Figure 5-22, expressed as the relative change in total ton/day emissions based on the change from 10th to 90th percentile population fractions. For some source types, the results show that adding population decreases overall emissions. This is due to internal MOVES logic for allocating VMT within HPMS classes. An increase in population can result in less VMT for source types within the same HPMS class; in some cases, the drop in VMT for adjacent source types has a larger effect on emissions than the increase in population, resulting in a net emissions decrease.

Figure 5-22. Change in Total Ton/Day Emissions Going from 10th to 90th Percentile Population Fraction, by Source Type

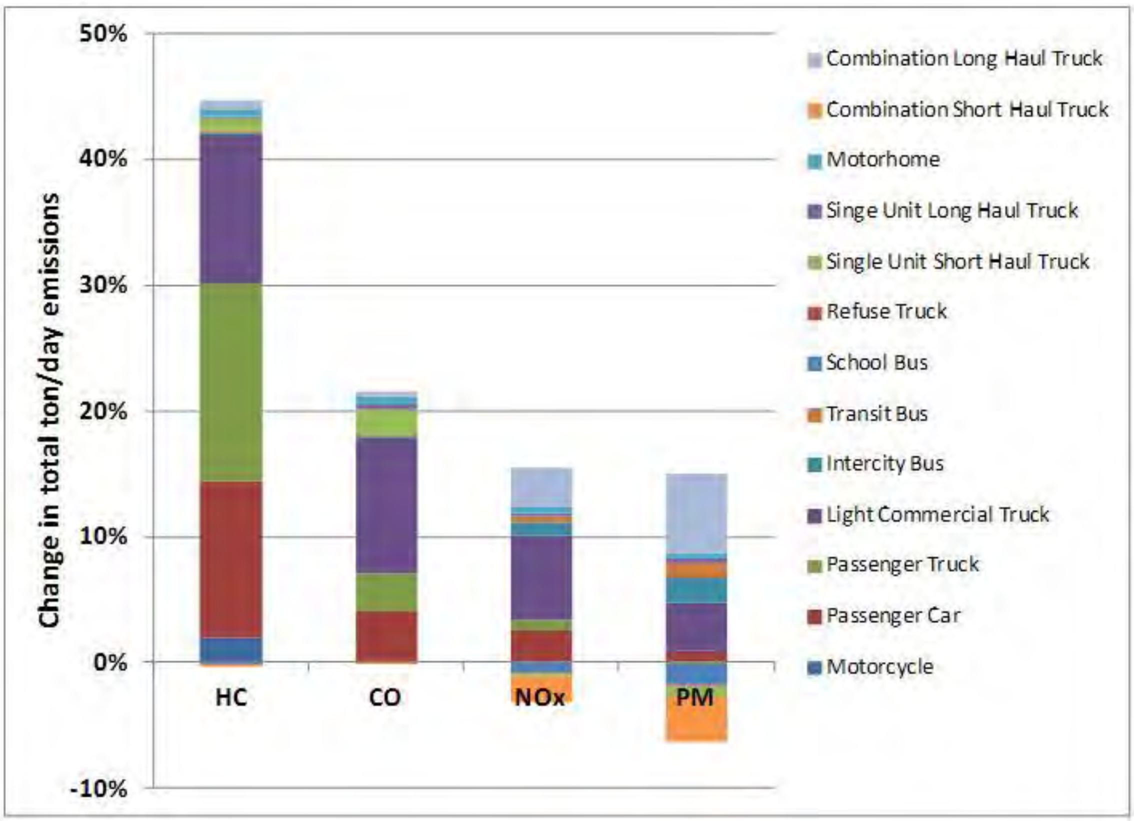
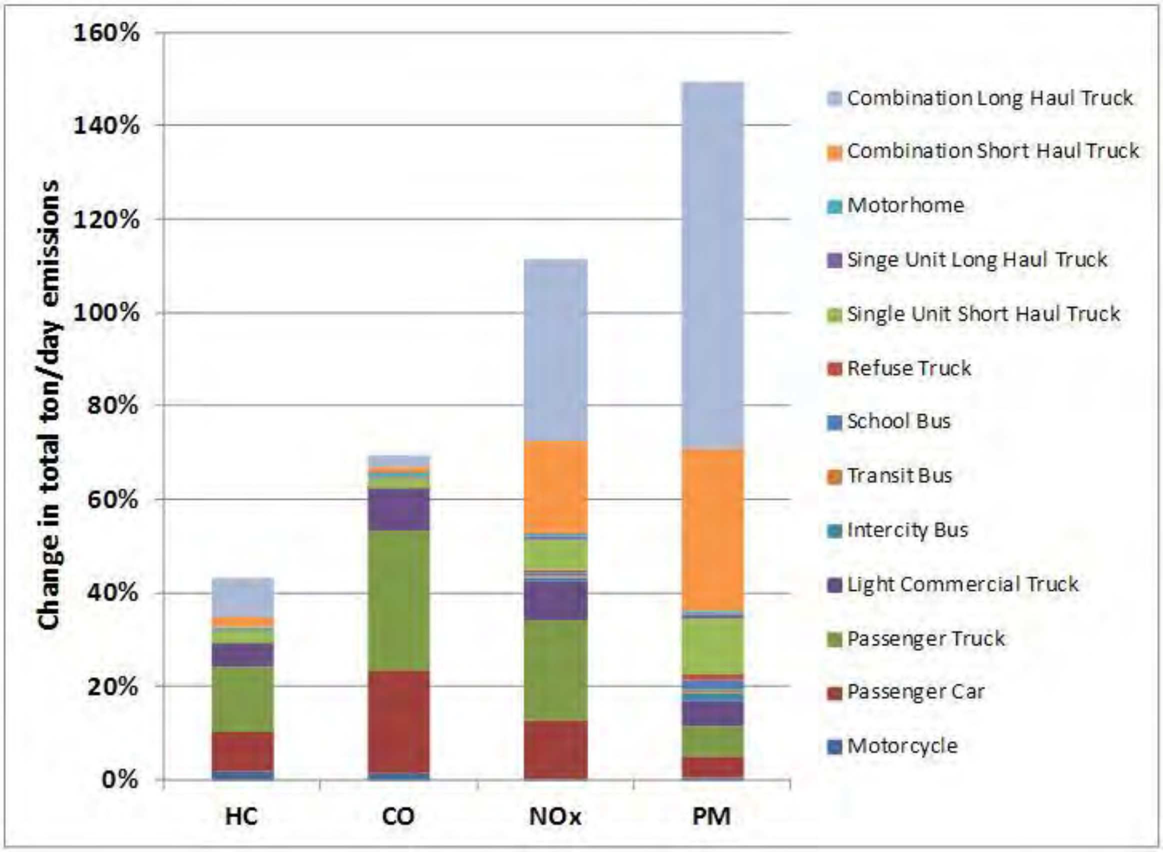


Figure 5-22 shows that MOVES is most sensitive to changes in population for HC, with a range in total emissions close to 45 percent between the 10th and 90th percentile cases. This reflects the importance of the start and evaporative emission processes on total HC, which are directly influenced by population. The importance of start emissions to total CO contributes to a variation in CO of over 20 percent. The variation for NOx and PM are lower, around 10-15 percent, as start emissions are less influential for diesel trucks in the model. Overall, light trucks, followed by passenger cars, contribute most of the sensitivity observed in this experiment. Emission results uniformly increase if population is varied for all source types within a given HPMS class at the same time. For MOVES users, varying population for all source types within an HPMS class simultaneously would be one approach to avoid unintentional emission decreases that may occur when varying populations of individual source types. Alternatively, users could consider varying VMT along with population.

Vehicle Miles Travelled

Inputs for the VMT sensitivity case levels were developed by establishing a constant total VMT (across all source types), based on median VMT fractions. 10th and 90th percentile VMTs were calculated by applying the 10th and 90th percentile VMT fractions for the parent HPMS class to the median total VMT. The resulting lower/higher VMT for a given HPMS class was then subtracted/added to the median total. Although VMT for the entire HPMS class was varied, allocations to source type within each HPMS class were kept intact (because populations were not varied). The impact of changing VMT on emissions was analyzed at the source type level. The resulting emission changes going from the 10th to 90th percentile VMT fraction are shown in Figure 5-23.

Figure 5-23. Change in Total Ton/Day Emissions Going from 10th to 90th Percentile VMT Fraction, by Source Type

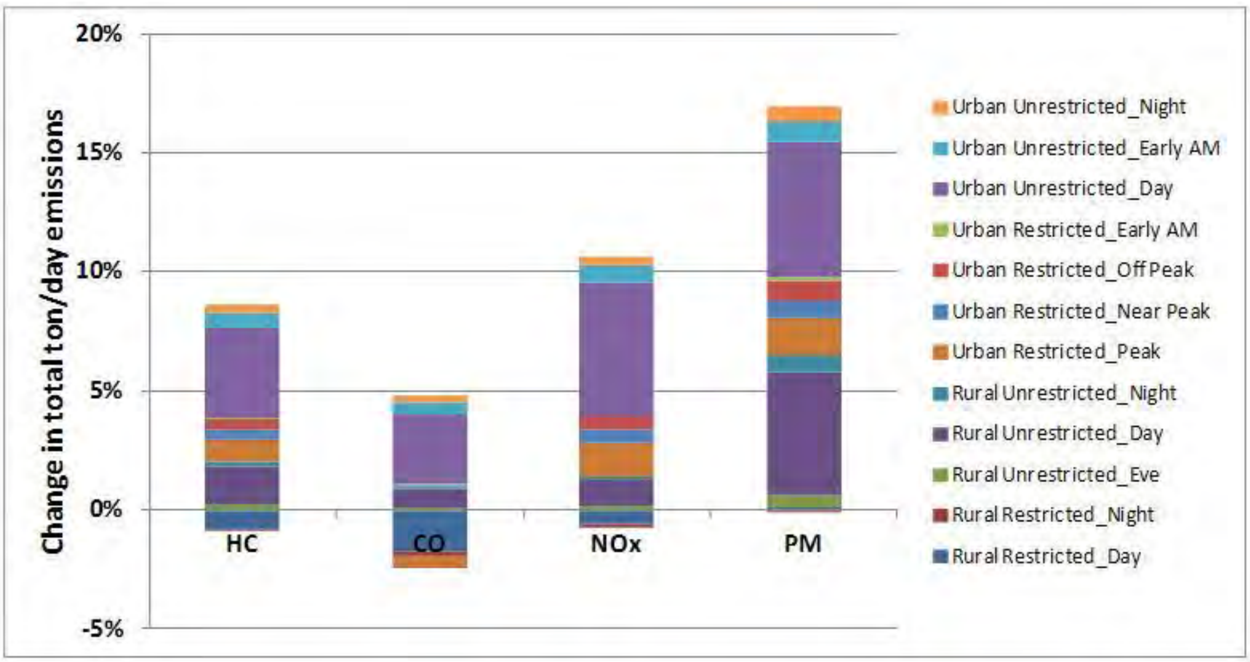


As expected given the influence of VMT, the total increases are large, from 40 to nearly 150 percent depending on pollutant. These increases are a consequence of total VMT more than doubling when all VMT fractions are increased simultaneously; while this is an extreme change in total VMT, it is within the range of variation observed in comparing state-submitted VMT to default VMT totals by county. Looking at individual source type results, of particular note is the influence of the large variation in combination truck VMT fraction on NOx and PM emissions. The increase in combination truck VMT alone caused total NOx emissions to vary over 50 percent, and total PM emissions to vary over 100 percent. This highlights the importance of improving heavy truck VMT estimates for local areas.

Average Speed and Road Type Distribution

Ultimately, Average Speed and Road Type Distributions were analyzed in conjunction for Task 2, and the impact of Speed depends on the mix of Road Type, and vice versa. For the MOVES sensitivity runs, the 10th and 90th percentile average speed distributions for each hour cluster were input into the model, and changes in total daily emissions evaluated. To evaluate the impact of a change in a single cluster on total daily emissions requires an assumption about road type distribution, which varies greatly across counties. To give a sense of overall sensitivity, the MOVES national default road type mix scenario was evaluated (14/37/18/31 percent of VMT on Rural Restricted, Rural Unrestricted, Urban Restricted, and Urban Unrestricted road types, respectively), with results shown in Figure 5-24. Since lower average speeds reflect more congestion and hence more emissions, these result are presented as the increase going from 90th to 10th percentile average speeds.

Figure 5-24. Change in Total Ton/Day Emissions Going from 90th to 10th Percentile Average Speed, by Hour Cluster (MOVES National Default Road Type VMT Distribution)



Sensitivity runs were also for the cases where 100% of each road type were in a given county; while this is not a realistic scenario for any county, it provides a bound for how each road type is influenced by variations in average speed on that road type. The emission sensitivity results for these cases are shown in Figures 5-25 through 5-28; since lower average speeds reflect more congestion and hence more emissions, there are presented as the increase going from 90th to 10th percentile average speeds. Results for a given county would be based on the road type distribution within that county.

Figure 5-25. Emission Results for Average Speed, 100% Urban Unrestricted

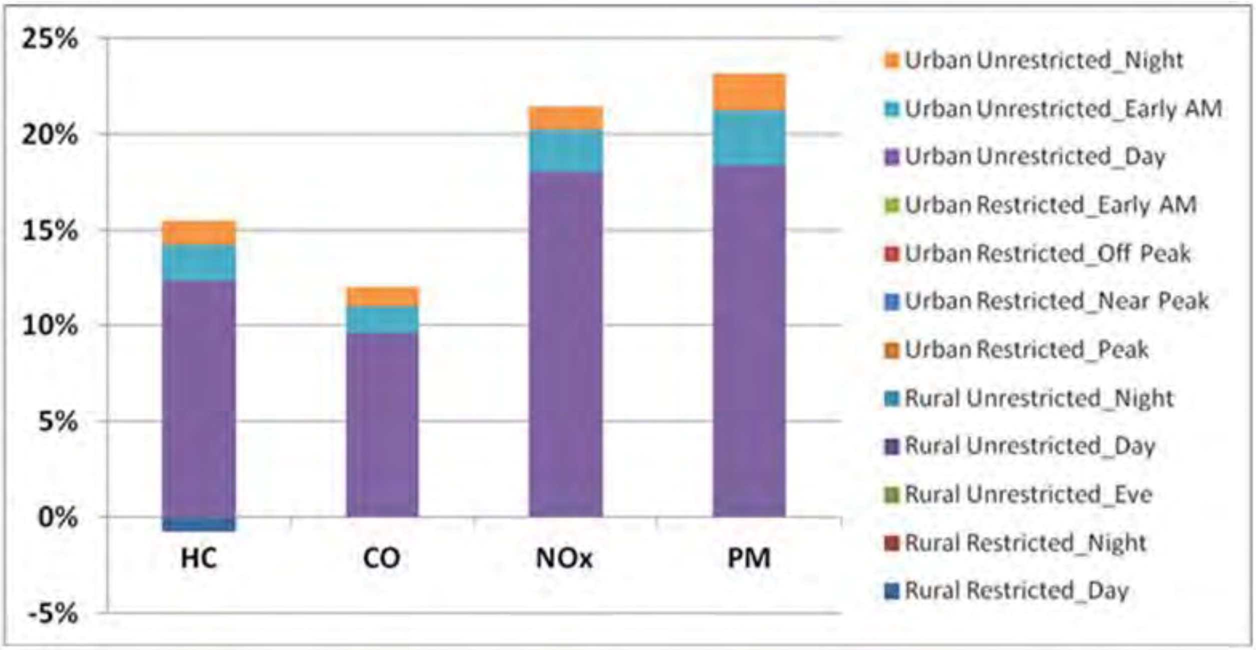


Figure 5-26. Emission Results for Average Speed, 100% Urban Restricted

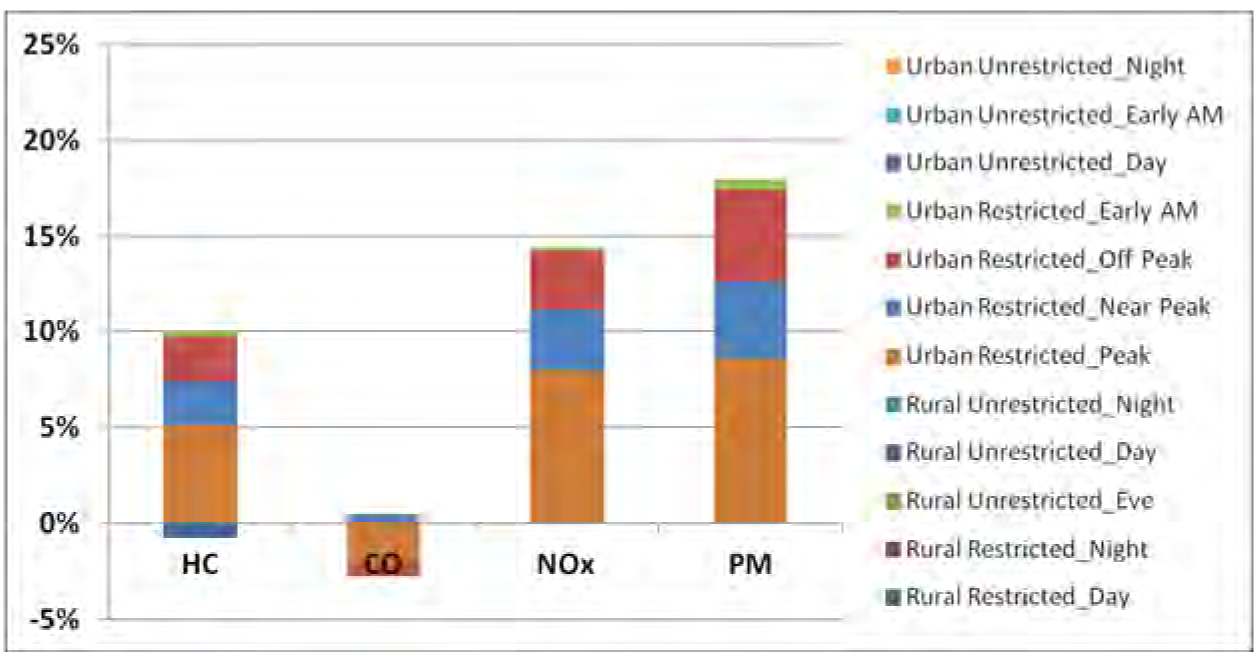


Figure 5-27. Emission Results for Average Speed, 100% Rural Unrestricted

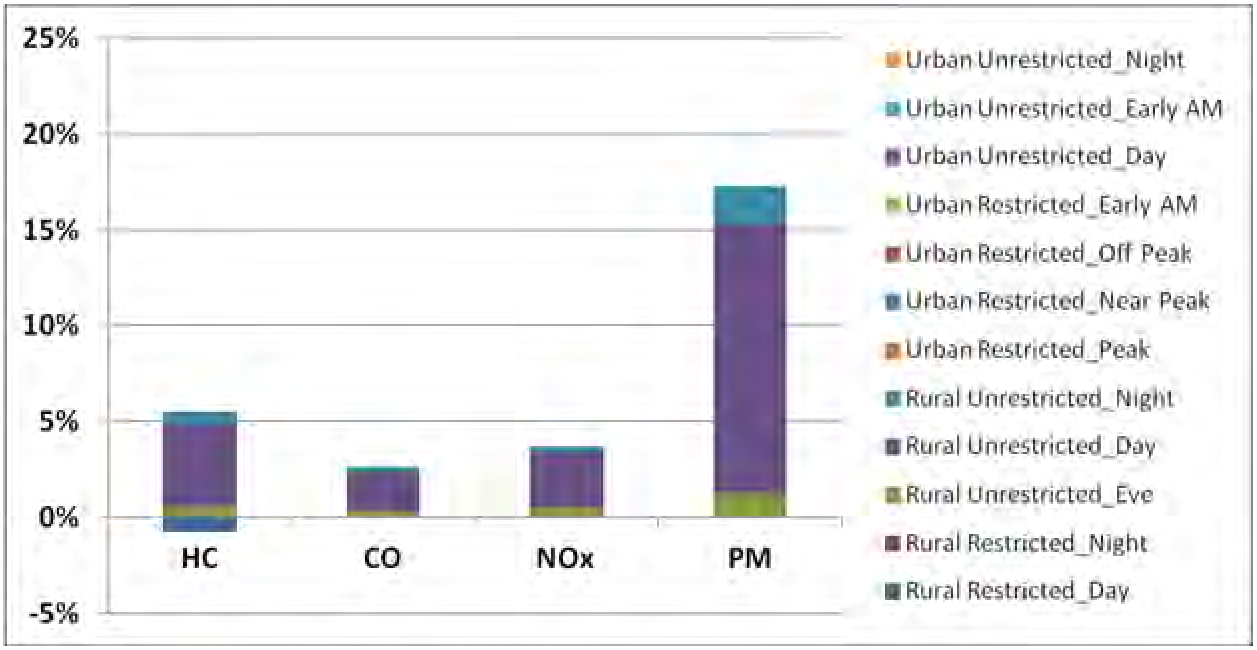
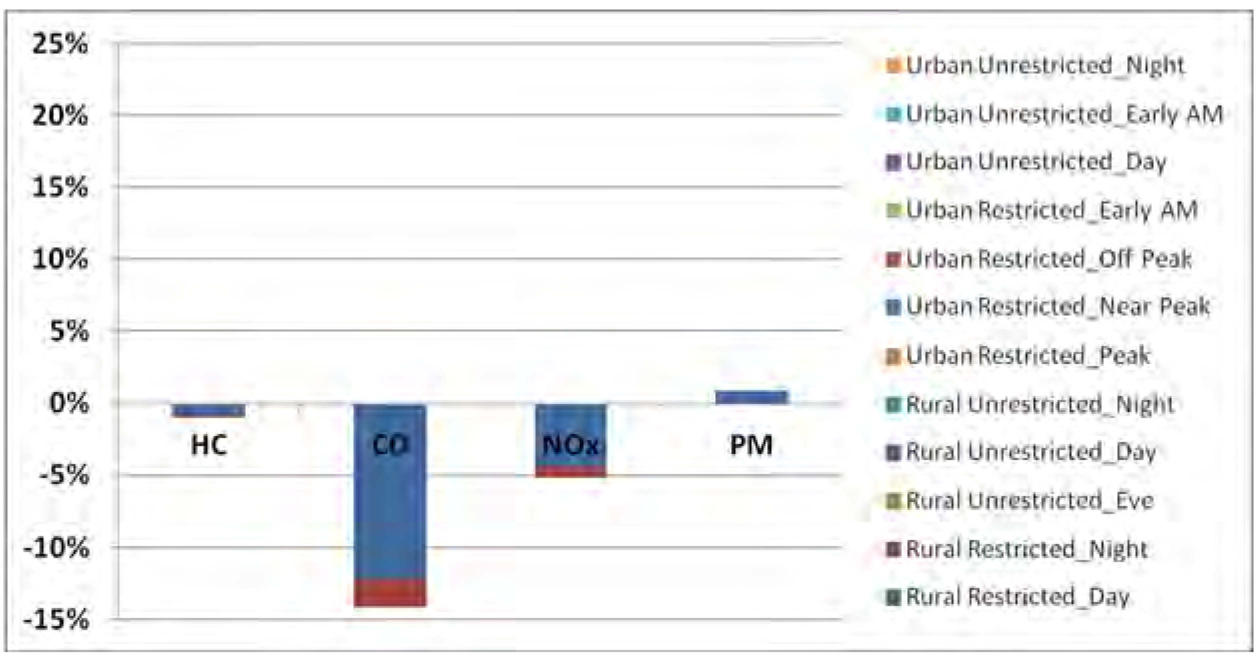


Figure 5-28. Emission Results for Average Speed, 100% Rural Restricted

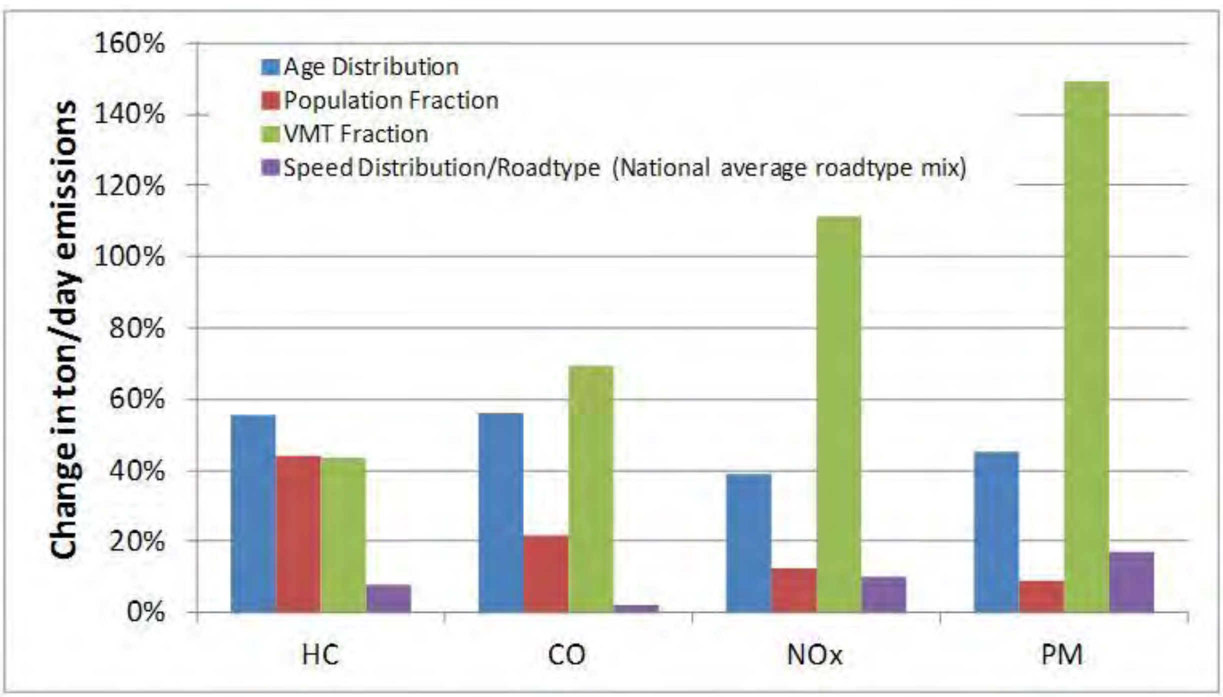


Overall the results show that that variations in average speed, generally from uncongested to more congested conditions as submitted by states, can increase total daily emissions up to 25 percent in total, with the largest variability coming on urban unrestricted roadway; this is not unexpected, as these roadways contain more stop-and-go driving that lead to higher emissions, and that are more sensitive to changes in average speed.

Comparison Across Inputs

The Task 2 sensitivity analysis was constructed to allow comparison across the primary inputs; this allows an assessment of how the contribution to total daily emissions of each input compares at the source type (or cluster) level, based on data submitted by states for the NEI. As an overview, Figure 5-29 shows the total increase in daily emissions if all source types went from the 10th to 90th percentile level for each of the inputs (90th to 10th for average speed). These are the compiled totals from the results for each input presented in Figures 5-21 through 5-24 above.

Figure 5-29. Total Change in Ton/Day Emissions Going from 10th to 90th Percentile Input for all Source Types/Hour Clusters, by Input



These results show that the importance of each input varies depending on pollutant. Overall, VMT changes contributed the highest change in emissions, with particularly large increases for NOx and PM. Age distribution is also very influential, particularly for HC and CO. Population is most influential for HC among the pollutants, which reflects the importance of start and evaporative emissions in total HC. The influence of average speed is the lowest, which may seem a surprising result – however, it should be noted that this analysis looked at total daily emissions, where on the whole variability in speed is relatively small and generally restricted to a few hours of the day. An analysis focused on hourly emissions at the project level would likely show a larger influence from speed.

Table 5-25 presents a ranking of the most influential inputs (i.e. largest change in total daily emissions when the input is varied from 10^{th} to 90^{th} percentile) at the source type / cluster level, for emissions changes five percent or greater. Note that for this analysis 100 percent road type VMT was used for the average speed clusters, representing an upper bound for this input. As shown, inputs for passenger cars and trucks tend to be towards the top of the list, as would be expected; combination trucks are very prominent for NOx and PM.

Table 5-25. Most Influential Inputs by Source Type / Cluster

	HC	
Source Type/Cluster	**Input Varied**	**Increase in Total Ton/Day Emissions**
Passenger Car	Age Distribution	23.5%
Passenger Truck	Age Distribution	22.3%
Passenger Truck	Population Fraction	15.7%
Passenger Truck	VMT Fraction	13.9%
Passenger Car	Population Fraction	12.4%
Urban Unrestricted_Day	Average Speed Distribution	12.3%
Light Commercial Truck	Population Fraction	11.8%
Combination Long Haul Truck	VMT Fraction	8.5%
Passenger Car	VMT Fraction	8.4%
Light Commercial Truck	Age Distribution	6.0%
Urban Restricted_Peak	Average Speed Distribution	5.1%
	CO	
Source Type/Cluster	**Input Varied**	**Increase in Total Ton/Day Emissions**
Passenger Truck	VMT Fraction	29.8%
Passenger Truck	Age Distribution	24.9%
Passenger Car	VMT Fraction	21.8%
Passenger Car	Age Distribution	21.3%
Light Commercial Truck	Population Fraction	10.8%
Urban Unrestricted_Day	Average Speed Distribution	9.6%
Light Commercial Truck	VMT Fraction	9.1%
Light Commercial Truck	Age Distribution	6.6%
	NOx	
Source Type/Cluster	**Input Varied**	**Increase in Total Ton/Day Emissions**
Combination Long Haul Truck	VMT Fraction	39.0%
Passenger Truck	VMT Fraction	21.5%
Combination Short Haul Truck	VMT Fraction	19.9%
Urban Unrestricted_Day	Average Speed Distribution	18.0%
Passenger Car	VMT Fraction	12.4%
Passenger Truck	Age Distribution	10.8%
Passenger Car	Age Distribution	9.8%
Light Commercial Truck	VMT Fraction	8.4%
Urban Restricted_Peak	Average Speed Distribution	8.0%
Light Commercial Truck	Population Fraction	6.8%
Single Unit Short Haul Truck	VMT Fraction	6.5%
Combination Long Haul Truck	Age Distribution	6.2%
	PM	
Source Type/Cluster	**Input Varied**	**Increase in Total Ton/Day Emissions**
Combination Long Haul Truck	VMT Fraction	78.6%
Combination Short Haul Truck	VMT Fraction	34.7%
Urban Unrestricted_Day	Average Speed Distribution	18.3%
Rural Unrestricted_Day	Average Speed Distribution	14.0%
Combination Long Haul Truck	Age Distribution	12.7%
Single Unit Short Haul Truck	VMT Fraction	12.0%
Urban Restricted_Peak	Average Speed Distribution	8.6%
Combination Short Haul Truck	Age Distribution	8.0%
Passenger Car	Age Distribution	7.4%
Passenger Truck	VMT Fraction	6.6%
Combination Long Haul Truck	Population Fraction	6.4%
Light Commercial Truck	VMT Fraction	5.7%
Single Unit Short Haul Truck	Age Distribution	5.0%

5.6 Task 1 & 2 Conclusions and Import for Practitioners

MOVES input data submitted to the National Emissions Inventory by state/local/tribal air agencies for over 1,400 U.S. counties were evaluated against EPA's best practice guidance, analyzed to determine the distribution of data provided, and assessed for impact of emissions. As the process for submitting and incorporating MOVES CDB data into the NEI is still evolving, ERG views the submission rate as a success; of course, more work is needed to raise this submission rate, but for an initial process the collection of this amount of data is larger than expected.

Task 1a found that the frequency of local data and adherence to best practice guidance varies depending on input; VMT inputs follow guidance to the highest degree, while fewer agencies provide average speed for their area. Detailed analysis of age distribution, population fraction, VMT fraction, road type distribution and average speed distribution under Task 1b showed large variation in these inputs. In particular, submitted age distributions were generally older than assumed by MOVES defaults. The amount of data submitted allowed a better assessment of the spread of data across states. This spread in general was larger than expected, which has important implications for improving regional and national on-road emission inventories. The focus by EPA, state air and transportation agencies, and research consortiums such as CRC and TRB on collecting better local data to replace national defaults is justified.

Under Task 2, a MOVES sensitivity analysis was performed based on levels determined from the spread of the submitted data. Sensitivity analysis on MOVES emissions confirms a strong influence on this spread of inputs on emissions predictions. Overall, MOVES *total daily* emissions varied up to 56 percent for HC, 70 percent for CO, 111 percent for NOx, and 149 percent for PM based on changes in a single input within the range of state-submitted data. The importance of each input and source type varied depending on pollutant; for HC and CO, dominated by light-duty gasoline sources, passenger car and truck age distribution and population were highly influential. For NOx and PM, a large variation in combination truck VMT fraction led to very large emission differences.

Task 1 and 2 presents the first comparison of the relative influence of emissions by input and source type based on state-submitted data, which will help MOVES users target areas where focused data collection will lead to the most improvement in emission inventory estimates. Overall, this work underscores the critical need for good local data in developing regional emission inventories.

6.0 Task 3: Recommendations for Improvements

The purpose of Task 3 was to provide recommendations for improving data submissions to the NEI, for the ultimate objective of improving NEI emission estimates and its use as a repository for the best local data for use in MOVES. To this end, ERG's recommendations center on two primary objectives: increasing the number of states submitting best practice MOVES inputs, and taking advantage of emerging data sources to improving the scope and quality of local MOVES inputs. Specific recommendations related to these two objectives are discussed below.

Conduct outreach to states that did not provide data. States that did not submit custom data (other than California) either haven't developed MOVES inputs because they aren't doing SIP /Conformity analysis, or have compiled inputs for SIP/Conformity but don't have the resources, time or priority to submit these for the NEI. Recommendations on how to address this depend on the case, but in general will require more focused outreach to the states that did not provide custom data. State air advocacy partnerships (e.g. NACAA) and regional planning organizations (e.g. LADCO, MARAMA) will be useful allies in this effort. These organizations provide some of the best forums for sharing of best practices between states, and generally have working groups already set up to address modeling issues. Working directly with individual states to understand barriers to data submission – i.e. lack of time, resources, data, or interest – will be invaluable, at least for a handful of states if not practical for all of the non-submitting states. For states that do submit data, a template for documentation is recommended; this recommendation stems from Task 1a, where the unevenness in documentation detail made it more challenging to assess what states actually did.

Use emission sensitivity results to establish priorities for outreach. Related to the outreach recommendation above, states low on resources may be overwhelmed with the prospect of generating all of the MOVES CDB inputs accepted in the NEI. As shown in Task 2, however, focusing on a small number of inputs could have a big impact and lead to considerable improvement in the NEI emissions of particular states. We recommend using the results from this study to prioritize specific inputs to address initially, such as heavy-duty VMT, or light-vehicle population and age distribution. This outreach could provide more detailed guidance on where to find these data within a state. It may be less daunting for a state still coming up to speed on MOVES to focus on a few high priority inputs with a clear pathway for finding the data, rather than the full range of MOVES CDB inputs.

Provide assistance in compiling and converting SIP/Conformity MOVES inputs. In some cases states have compiled MOVES inputs for use in SIP/Conformity analysis which are not completely consistent with the inputs submitted for the NEI. This may be a function of different timing between the regulatory analyses and NEI cycles, and/or different methodologies used for SIPs. For these cases, additional work would be required to use these inputs for NEI, and resources or priority are not put towards this effort. For example, Texas performs link-level modeling using MOVES emission rates and an external software application for SIP/Conformity analysis, as well as the completed inventory submitted for the NEI. While this process developed many custom MOVES inputs by county, it does not cross-walk directly to the MOVES CDB approach used by EPA in the NEI. Lack of resources for making this conversion means that MOVES input data submissions aren't made (as in the case of Texas), or aren't consistent with the best practice SIP inventory. To address this, states doing SIP/Conformity work and not submitting the inputs to the NEI process should be identified, and efforts made to provide assistance to develop the needed NEI inputs.

Take advantage of emerging data sources to broaden and improve MOVES inputs. With respect to MOVES inputs, more sophisticated data sources are becoming available to improve local inputs. These data sources should be considered when updating MOVES guidance, to provide a broader array of options for state and local users to improve MOVES activity and fleet inputs. For activity inputs, the proliferation of commercial GPS devices (or even smart phones), such as the data described earlier use in developing speed "clusters", is a new arena for local areas to use for developing speed distributions where previously little to no data existed. Since average speed was one of the least submitted data fields, a focus on compiling data from these services could yield immediate improvements in regional and national emission inventories; EPA is already taking advantage of this emerging data source to update national default speed distributions for the next version of MOVES. For heavy trucks, an analogous technology is telematics, used by fleet operators to track truck location and activity.

For fleet-related inputs such as age distribution or vehicle population, MOVES inputs are meant to reflect the fleet of vehicles within the modeling domain over the time period being modeled. Although vehicle registration data is a typical source for these inputs, as directed by EPA guidance, the actual fleet on the road may differ due to commuter flow, pass-through traffic, or unregistered vehicles. While registration data is a good foundation, supplemental data sources can be used to better characterize the population of vehicles not in the a specific area's registration. These include roadside studies using automatic license plate recognition (ALPR), where roadside observations are made using a camera and computer with plate recognition software that can detect out-of-state plates, unregistered vehicles, etc. Automatic tolling systems (such as EZ-Pass) can also provide information on vehicles operating in an area that are not in the local registration database.

Use national databases for states not submitting data. Some MOVES inputs are available through national databases that compile data at the local level. Databases of vehicle registration data are compiled annually from state registration records by commercial entities, and available for purchase at the county level. FHWA publishes the annual Highway Statistics series which includes travel data for each state with additional detail for several major metropolitan areas in the U.S., and the National Household Travel Survey (NHTS) which compiles data that can feed MOVES inputs such as age distribution or starts per vehicle. Local fleet data on some specific source types such as transit and school buses are available through national databases as well, for example the National Transit Database compiled by the FTA. For these data sources, a national effort to purchase and disseminate local data (or directly populate MOVES CDBs for use in the NEI, etc.) is recommended to ensure that local data was being used in the NEI, even where state agencies did not have the resources to obtain the data themselves.

Overall, MOVES provides a tremendous amount of flexibility in customizing the model to local areas. The air quality community can continue to refine regional and national emission inventories to take advantage of new data and new approaches brought on by more sophisticated data collection technologies, and broader compilation of data sources. The 2011 NEI was an excellent first step in compiling data from many states; the A-84 project showed the variety in approaches and data, and the importance of these data on improving emission estimates. The recommendations discussed above are some ways that practitioners can continue to build on this success for subsequent NEIs.

7.0 Study Limitations

The scope of this study was inherently dictated by the number of states that submitted MOVES data, which was ultimately out of the control of the project sponsors, the authors, or EPA. The incompleteness of submitted data relative to full submission (i.e. where every state submits custom data for all MOVES inputs) defines a fundamental limitation of the study. While the results reported in this paper reflect a fairly broad distribution of input data, the sample size is still less than ½ of the counties in the U.S; and as noted, of the states that did submit data, many submitted MOVES defaults for specific fields. Ideally this analysis would have been conducted using actual local data for all MOVES inputs for every county in the U.S. outside California. This level of participation in the NEI is a good goal to aspire to, but is not a reality at this time as use of MOVES is still developing, particularly in states that aren't preparing SIP and Conformity analyses.

One reviewer of this work specifically questioned whether there were enough states submitting data to predict accurate results. Given the incompleteness of the state-submitted data, the results from this study cannot be used to fully quantify variability and uncertainty in the entire on-road U.S. NEI. The study is not making this inference, however, instead focusing only on data that was submitted, and reporting results (including emissions sensitivity) within this context. The sample size of data analyzed was sufficiently large to quantify the spread of the distribution within the states that submitted data, and could reasonably be used to conduct a sensitivity analysis of national on-road NEI emissions based on sampling the distribution of input data from submitted states.

The methodology for estimating emissions sensitivity is meant to show more extreme ends of the input distribution range, by using the 10th and 90th percentile inputs. The purpose of this was to help understand how real-world ranges of different inputs translate to emission differences in MOVES. These results can't be applied directly to other counties, where input data falls at other points in the distribution (30th percentile, etc.), because total daily emissions are not a linear function of the percentile level for any of the input fields analyzed. The cases run were meant to bound the real-world inputs, without modeling any specific real world case (no county has inputs in the 10th or 90th percentile for every MOVES input). Further work could make the results more directly applicable to specific counties by defining intermediate percentile point inputs and conducting additional MOVES runs with these.

Beyond the representativeness of the state-submitted data, this analysis did reveal some QA/QC issues in the data, requiring removal of data prior to analysis. The inputs that EPA used for the NEI underwent thorough QA prior to their use. However, the input data from Texas available for this project were not included in the EPA QA process, because they were not submitted through the EIS system (Texas submitted emissions directly in the EIS, rather than MOVES inputs). 66 Texas counties had outlier numbers for HPMS VMT and/or SouceTypeID populations. This was by design in the Texas inventory methodology, but the values could not be used in this analysis, so were not included. Incomplete submissions for certain MOVES fields provided by Oregon, Ohio and Knox County, TN required the removal of some data as well (in these cases, the data removed were MOVES default values).

8.0 Acknowledgements

The authors gratefully acknowledge several individuals who contributed to the success of this project:

- CRC's Atmospheric Impacts Committee for devising the project objectives and providing project funding, in particular the A-84 project leaders Susan Collet (Toyota) and Rory MacArthur (Chevron), and Brent Bailey and Betty Taylor of CRC;

- Marc Houyoux and Laurel Driver of EPA's Office of Air Quality Planning and Standards for granting permission to analyze submitted data prior to NEI release;

- Edward Nam of EPA's Office of Transportation and Air Quality for permitting analysis of their vehicle GPS dataset to be used in this work;

- Meredith Weatherby, Jeanette Alvis, Heather Perez and Anita White of ERG for assistance in data compilation, analysis, GIS mapping and formatting.

(this page left intentionally blank)

Appendix A

Figure A-1. Vector Angle vs. Mean Age, SourceTypeID=11

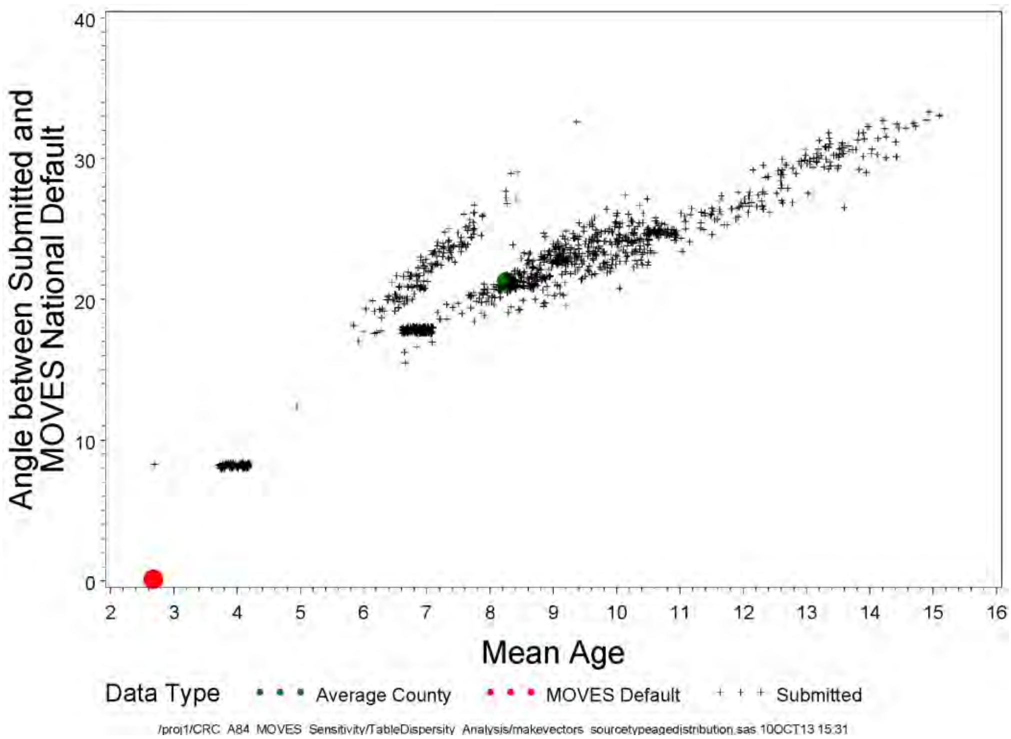


Figure A-2. Vector Angle vs. Mean Age, SourceTypeID=21

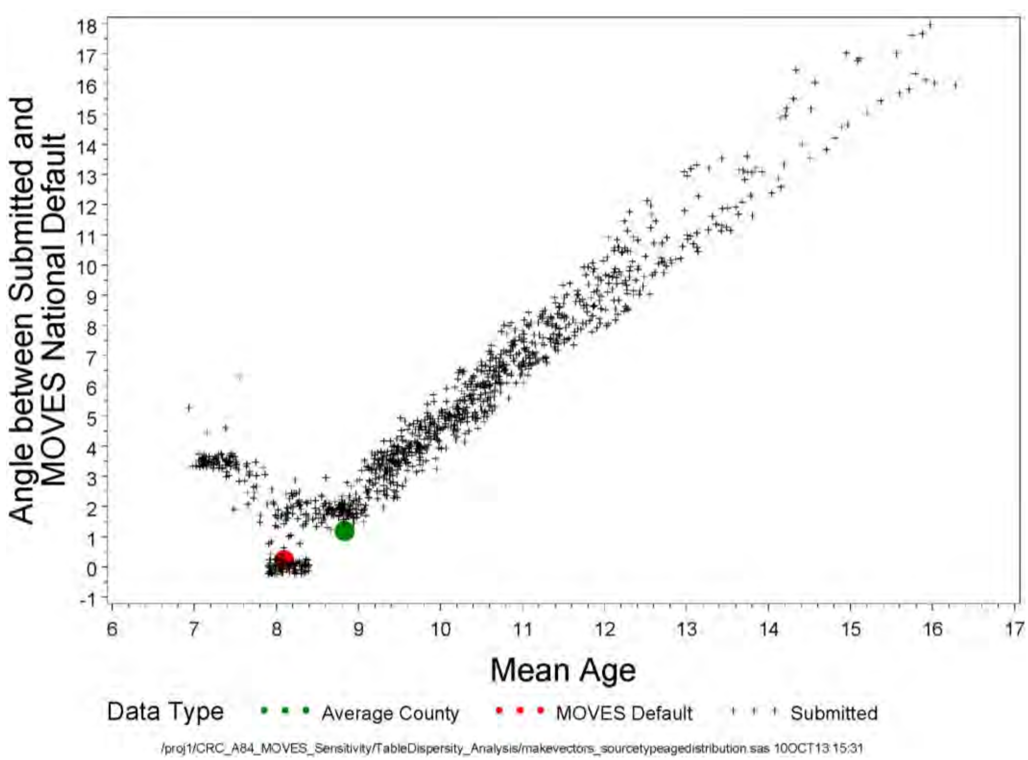


Figure A-3. Vector Angle vs. Mean Age, SourceTypeID=31

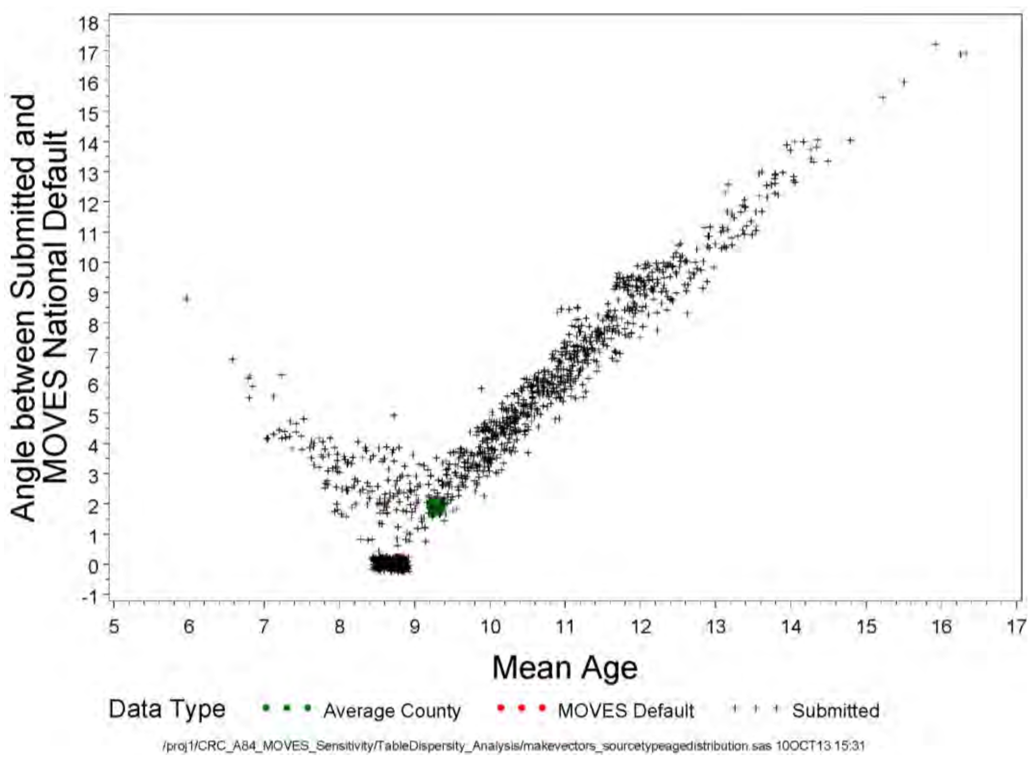


Figure A-4. Vector Angle vs. Mean Age, SourceTypeID=32

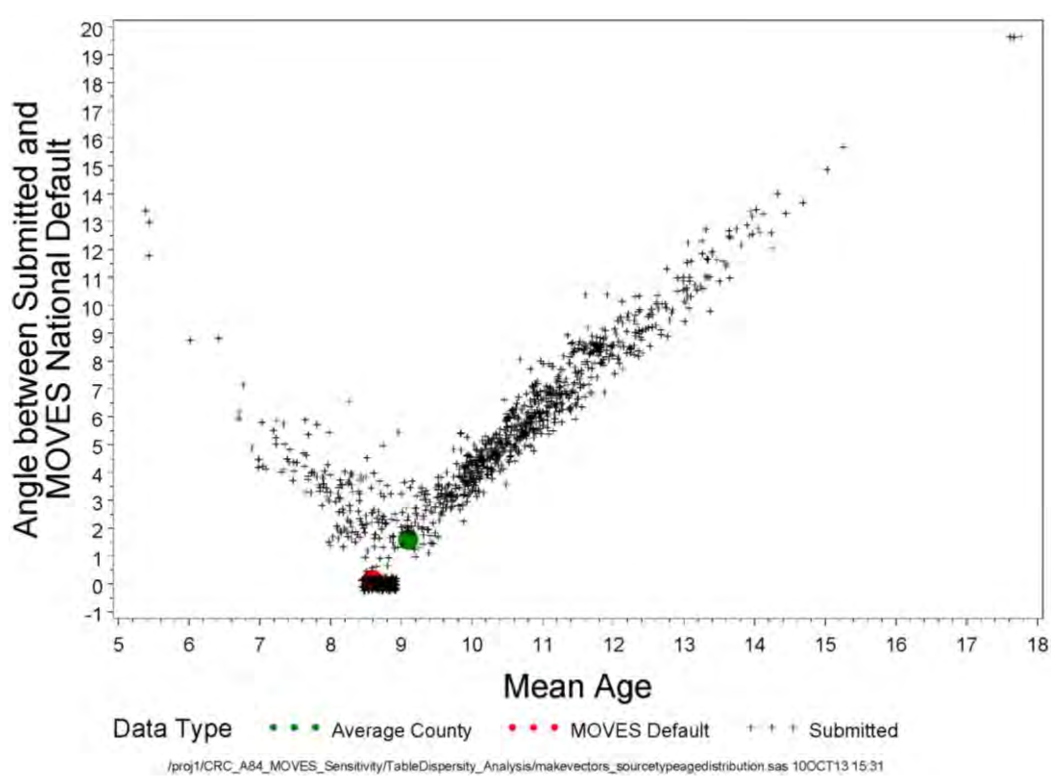


Figure A-5. Vector Angle vs. Mean Age, SourceTypeID=41

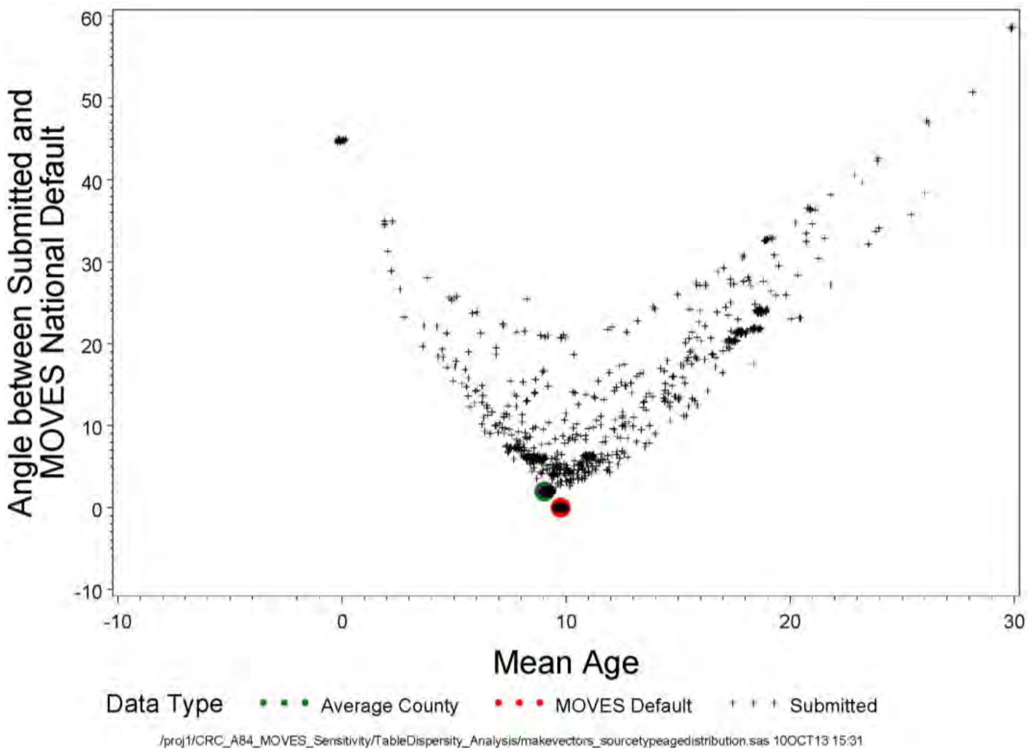


Figure A-6. Vector Angle vs. Mean Age, SourceTypeID=42

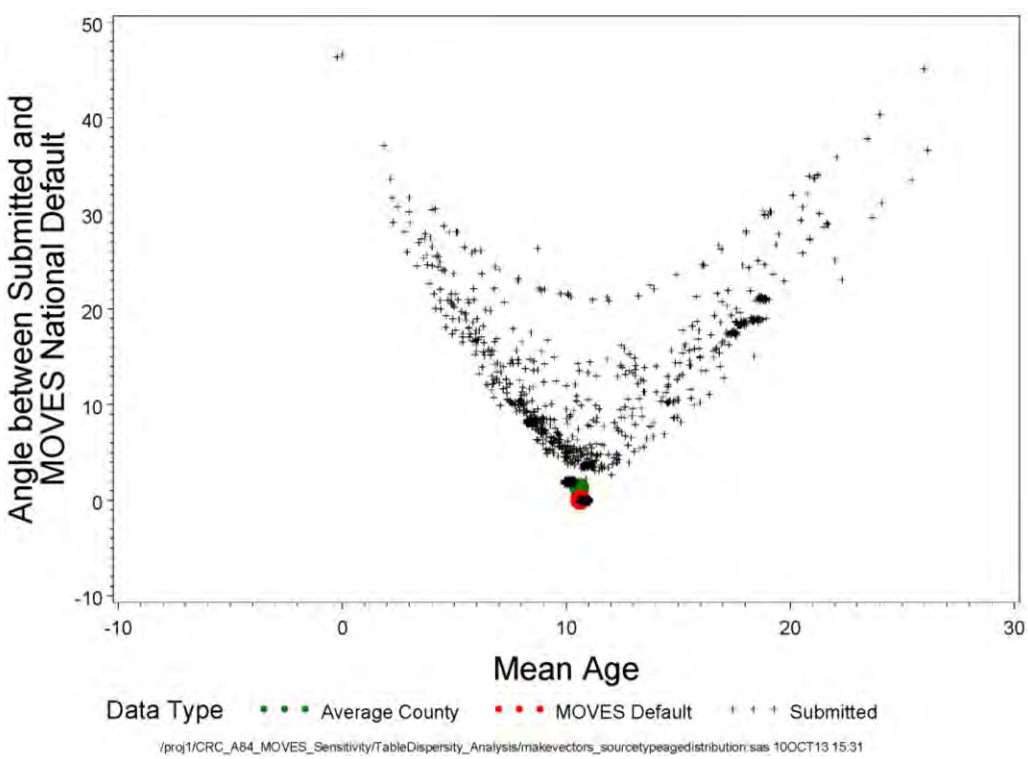


Figure A-7. Vector Angle vs. Mean Age, SourceTypeID=43

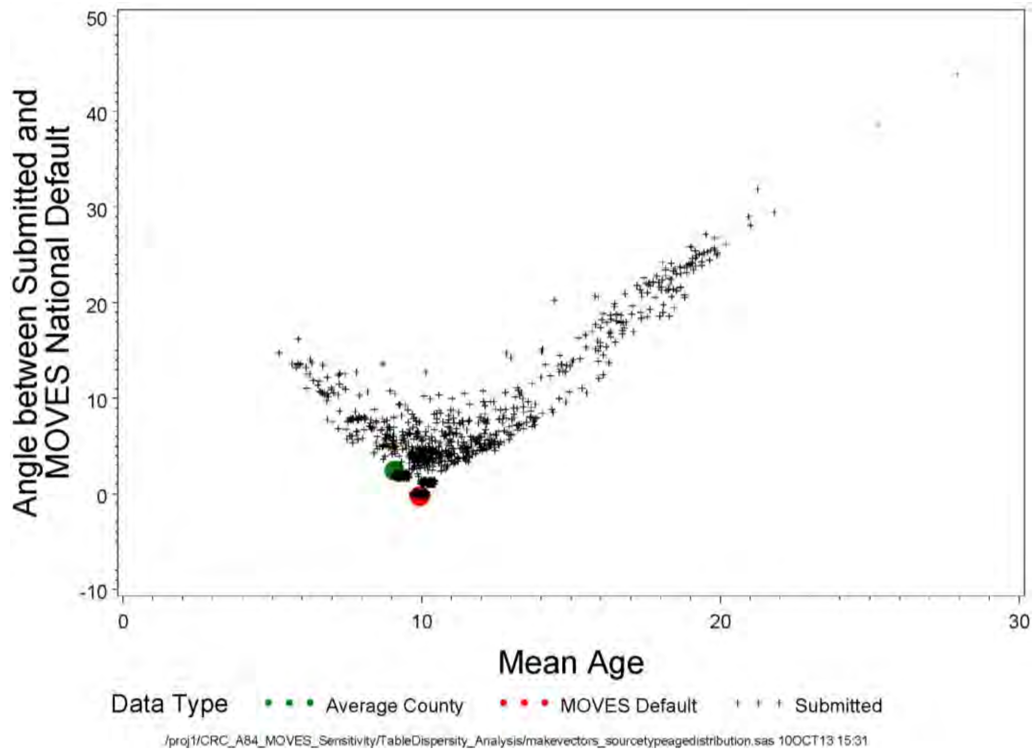


Figure A-8. Vector Angle vs. Mean Age, SourceTypeID=51

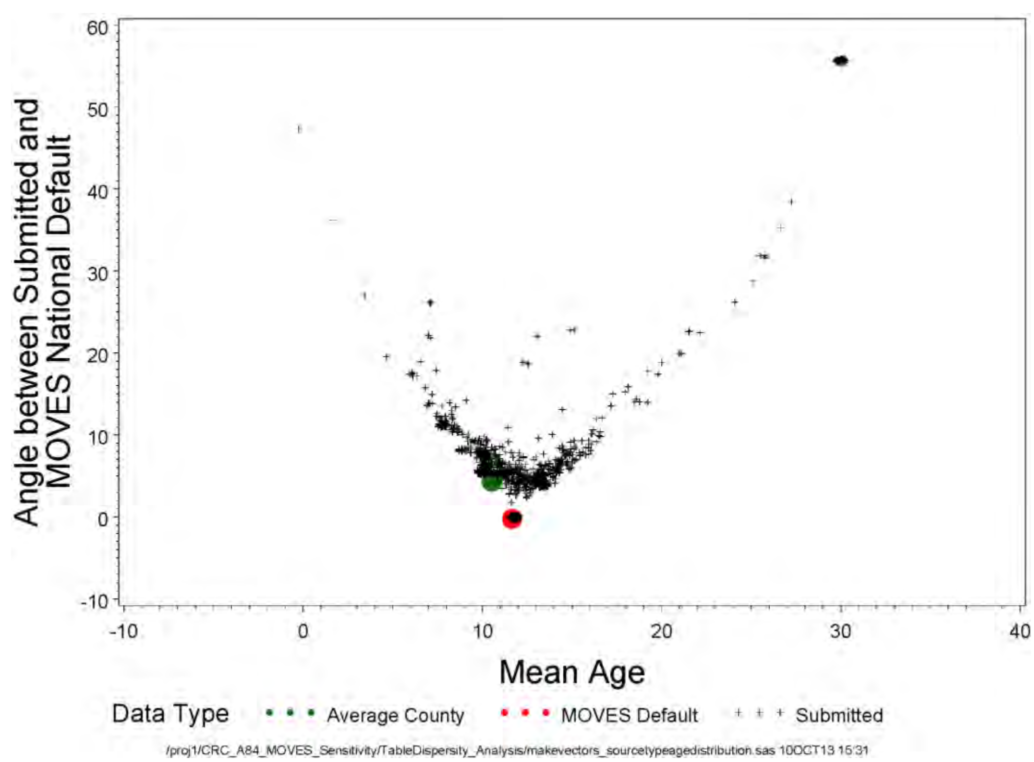


Figure A-9. Vector Angle vs. Mean Age, SourceTypeID=52

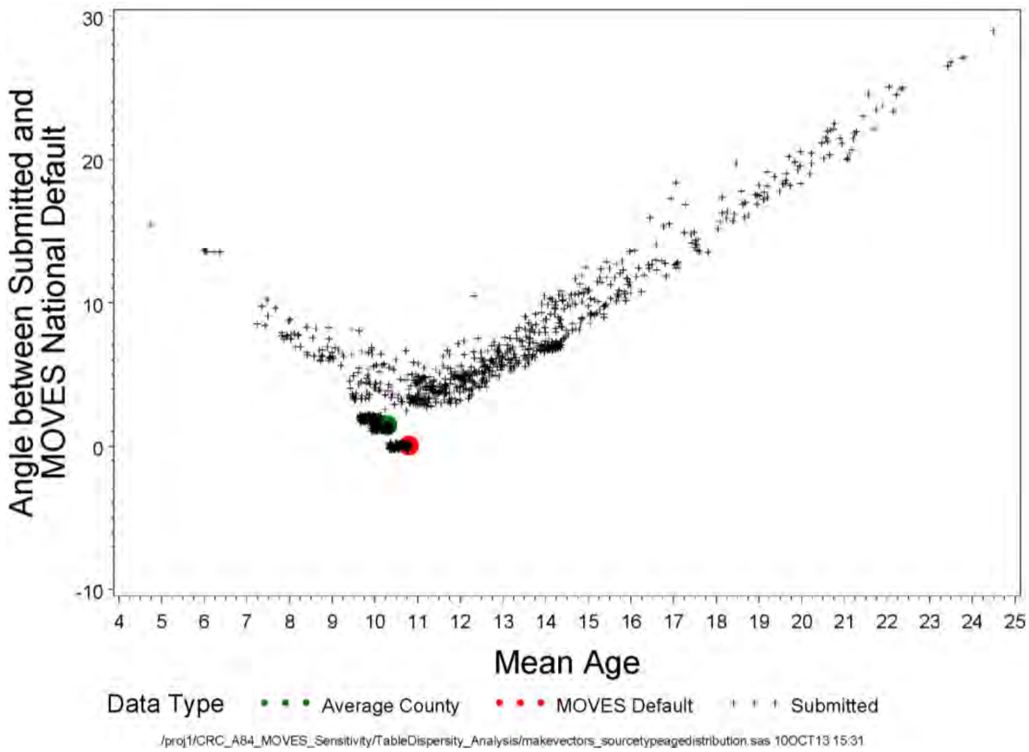


Figure A-10. Vector Angle vs. Mean Age, SourceTypeID=53

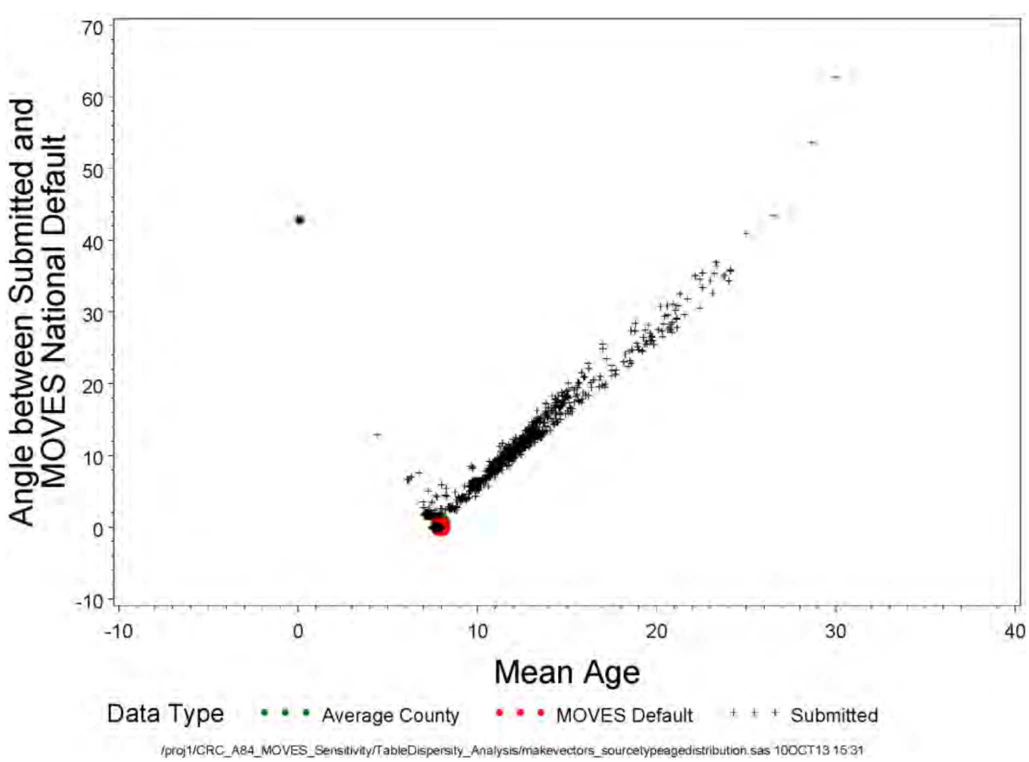


Figure A-11. Vector Angle vs. Mean Age, SourceTypeID=54

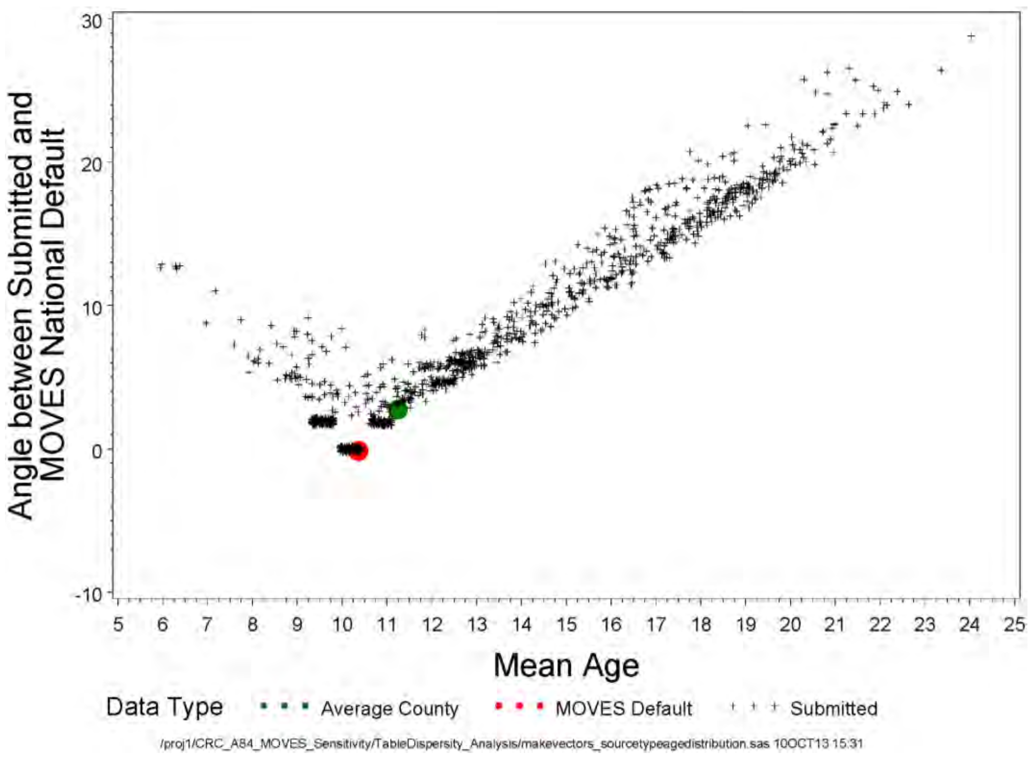


Figure A-12. Vector Angle vs. Mean Age, SourceTypeID=61

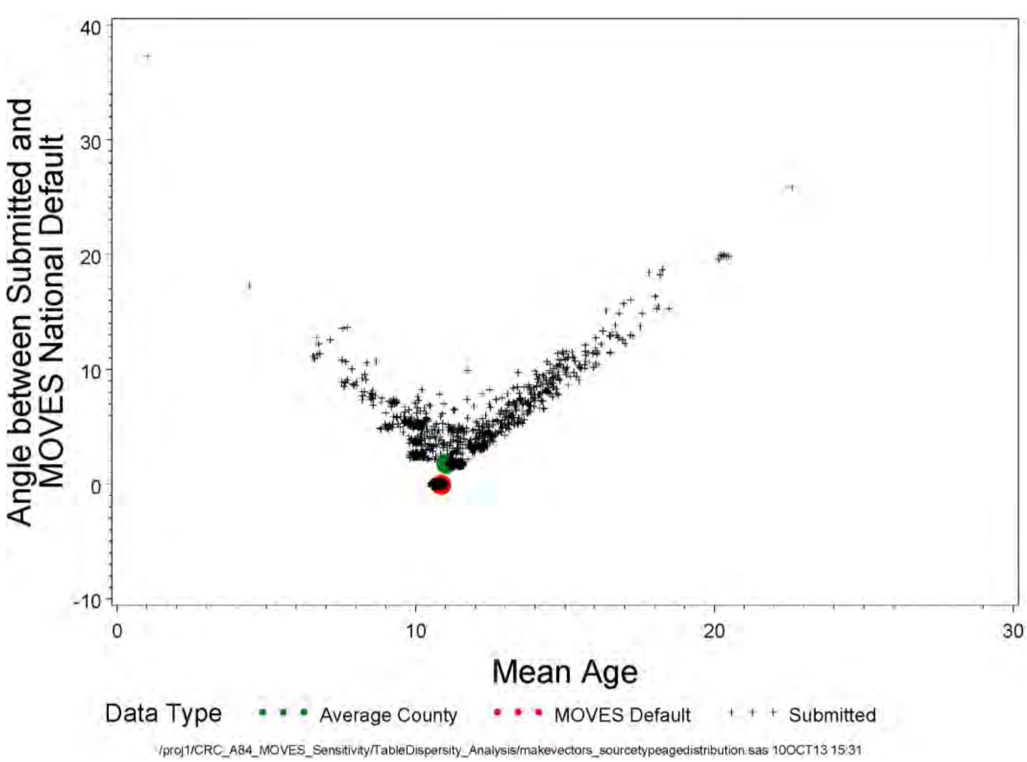
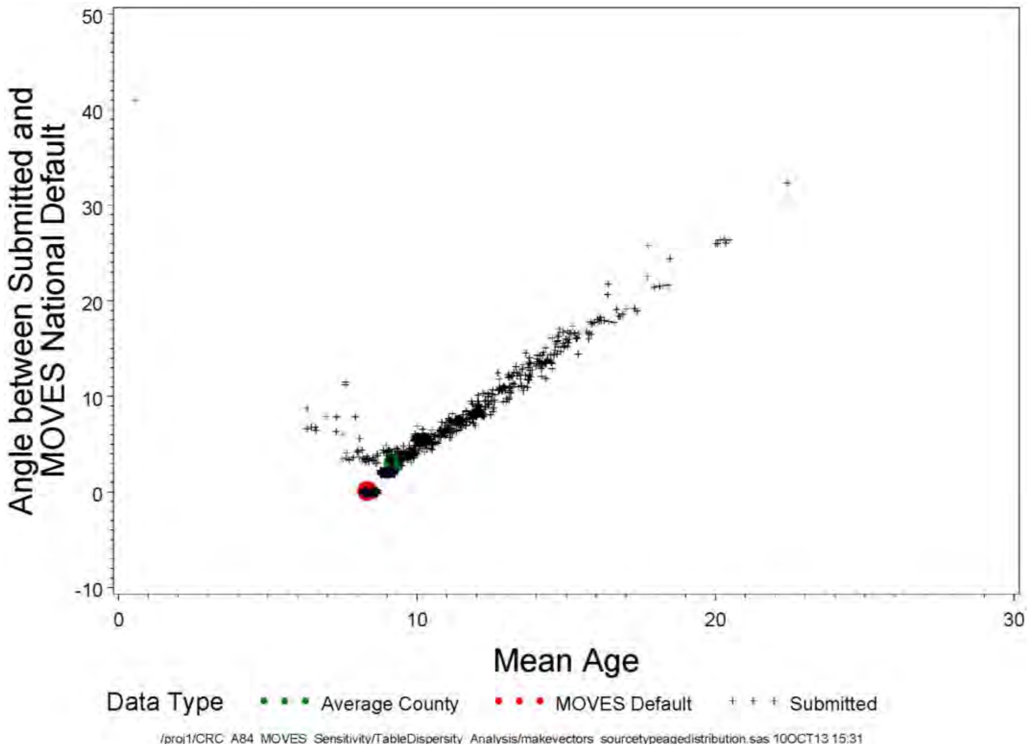


Figure A-13. Vector Angle vs. Mean Age, SourceTypeID=62



Appendix B

Figure B-1. Submitted and Default Populations for SourceTypeID=11

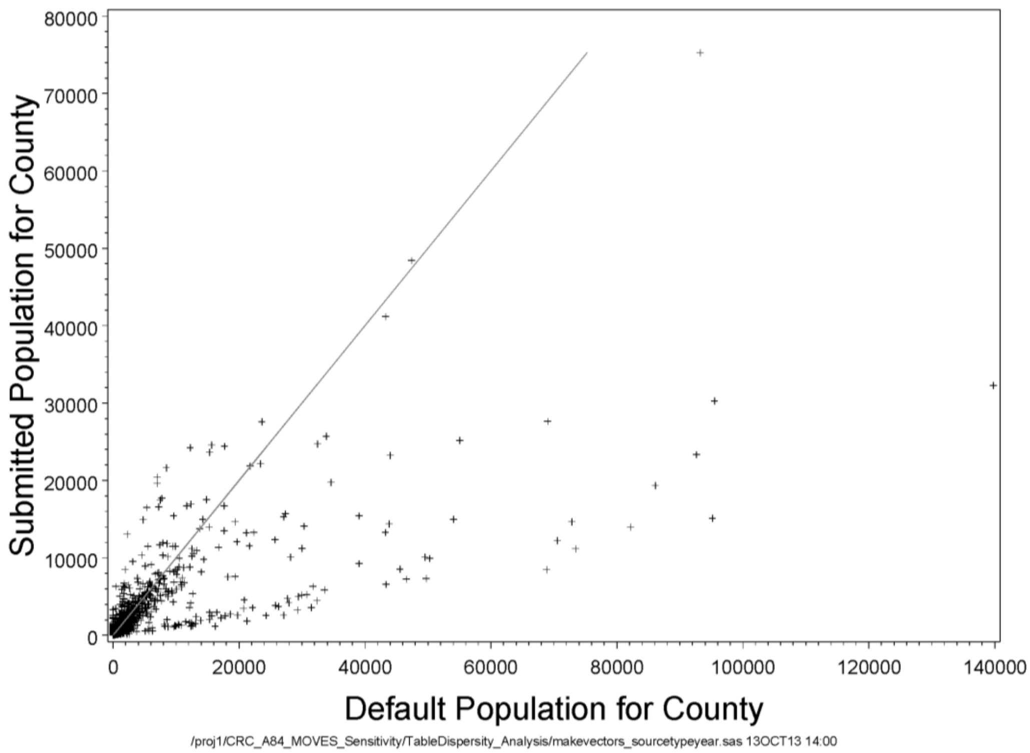


Figure B-2. Submitted and Default Populations for SourceTypeID=21

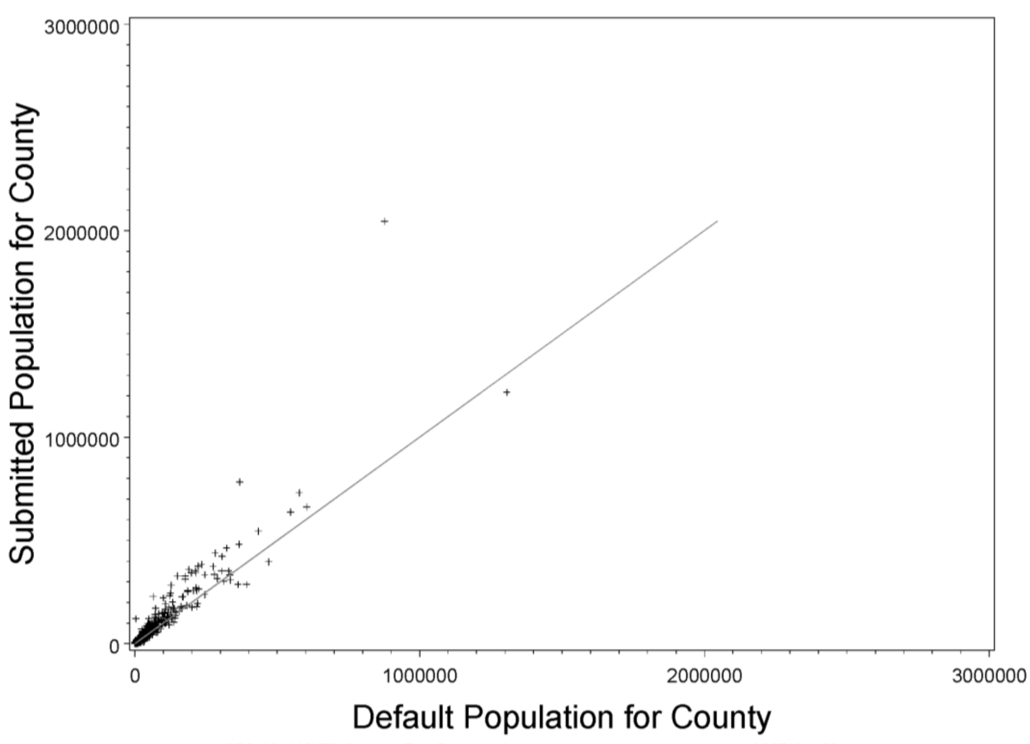


Figure B-3. Submitted and Default Populations for SourceTypeID=31

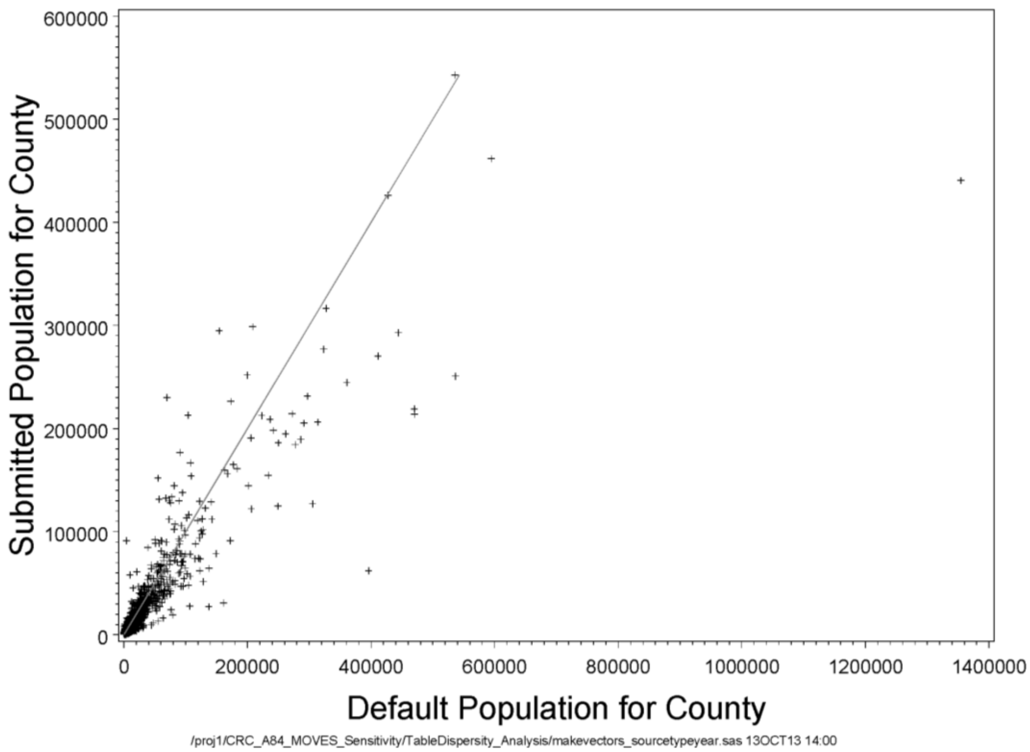


Figure B-4. Submitted and Default Populations for SourceTypeID=32

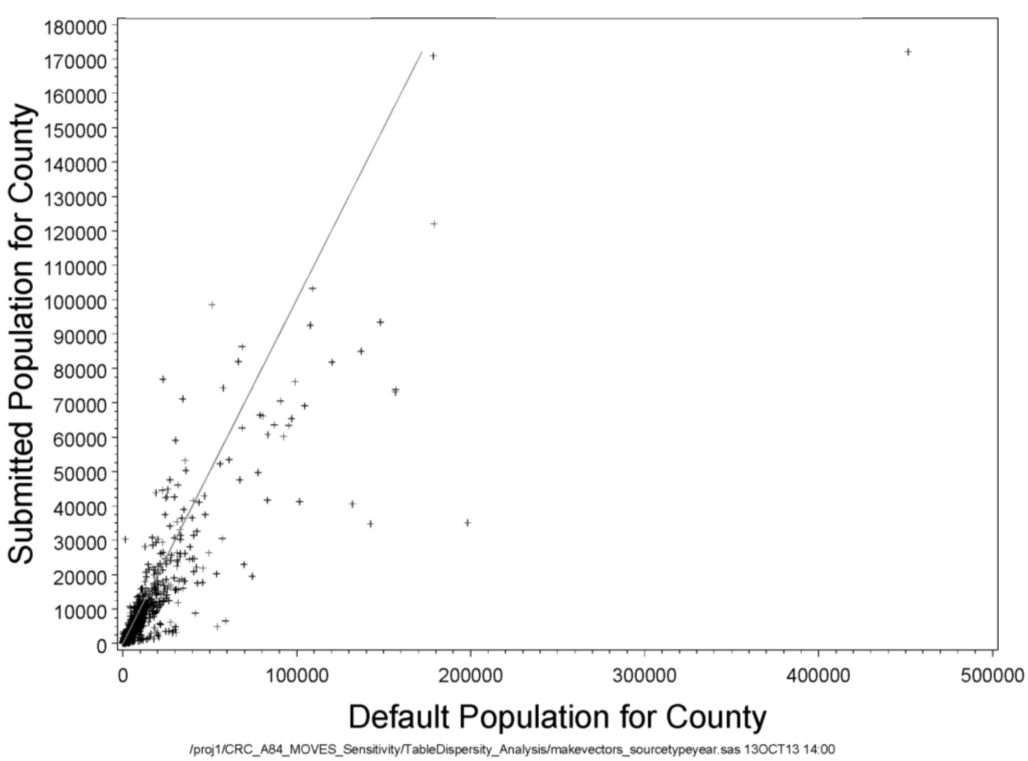


Figure B-5. Submitted and Default Populations for SourceTypeID=41

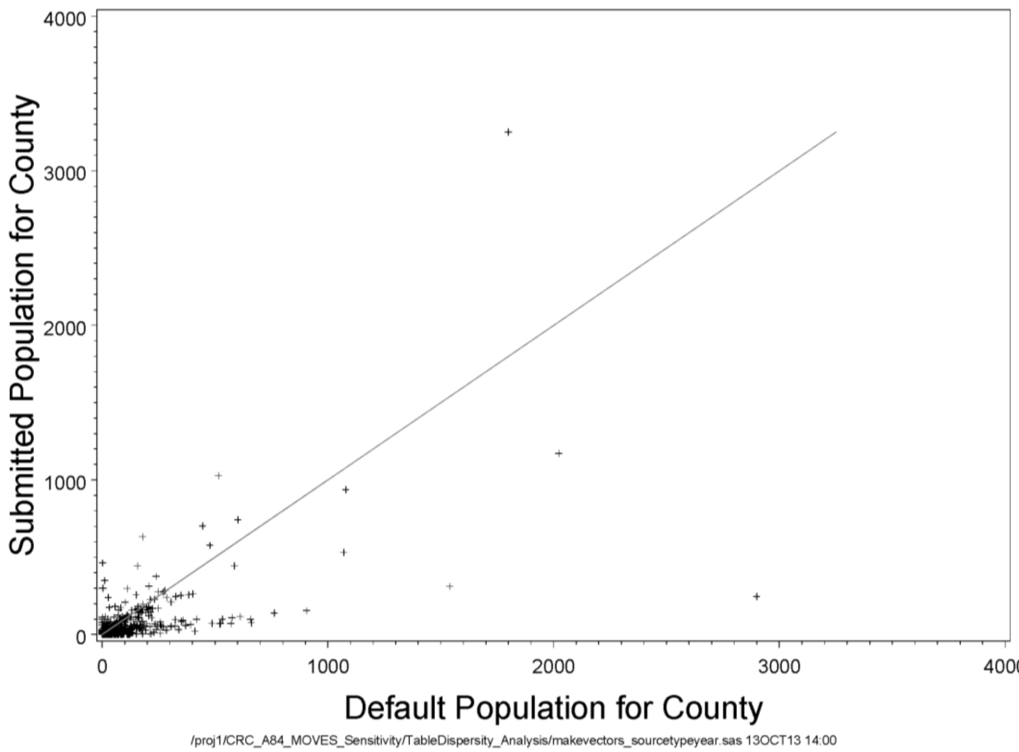


Figure B-6. Submitted and Default Populations for SourceTypeID=42

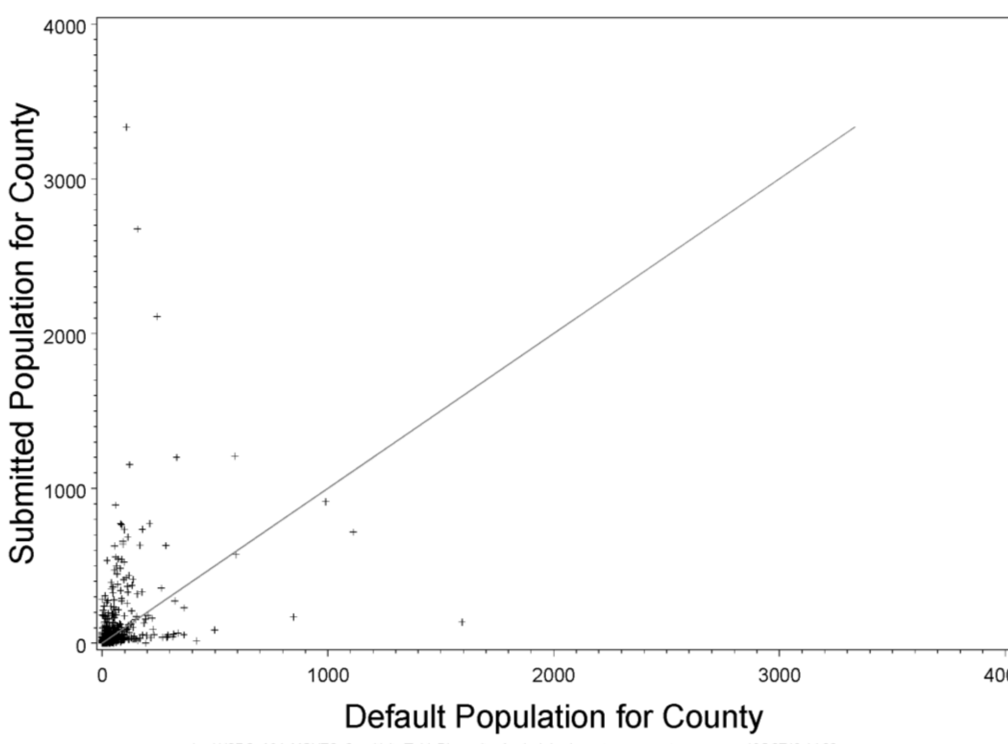


Figure B-7. Submitted and Default Populations for SourceTypeID=43

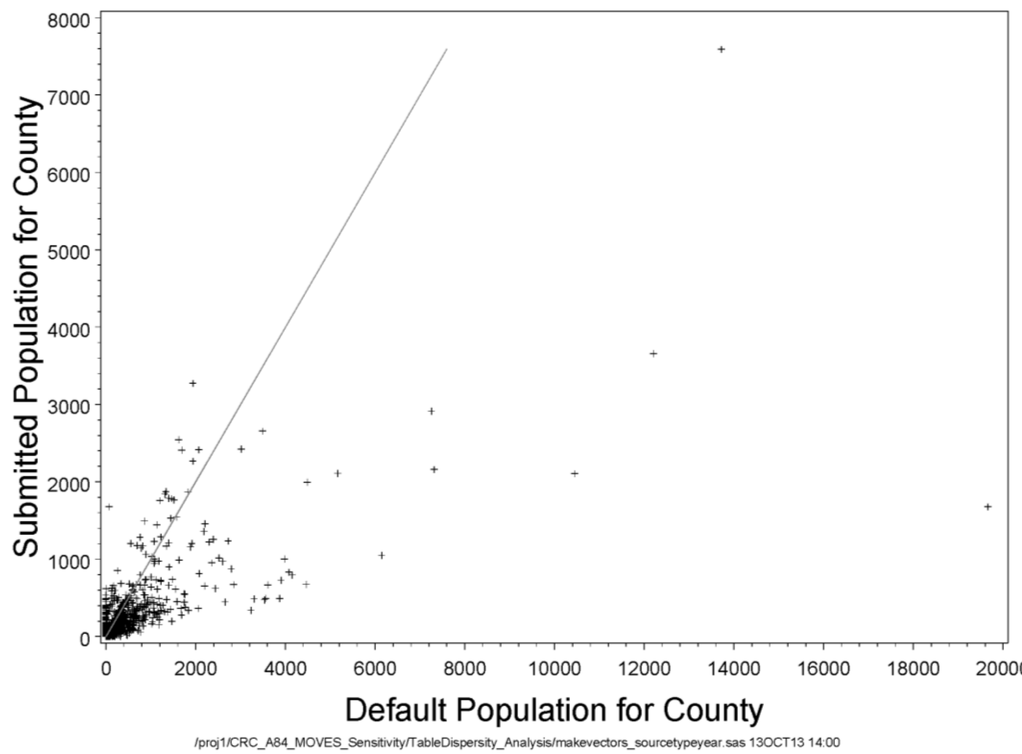


Figure B-8. Submitted and Default Populations for SourceTypeID=51

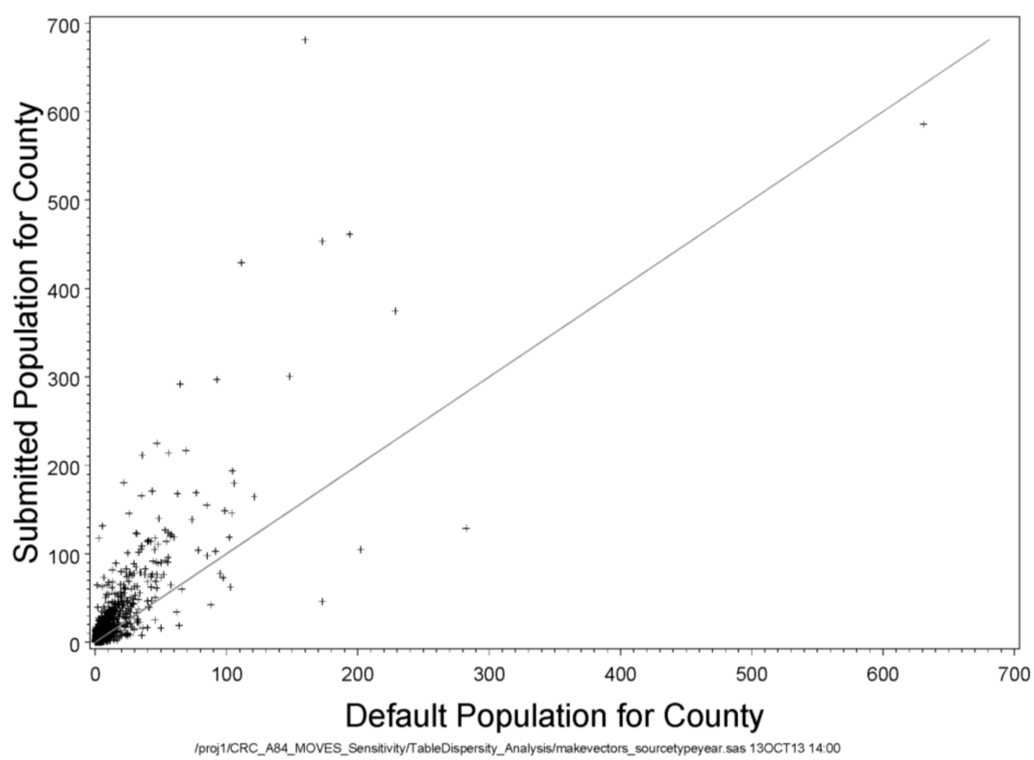


Figure B-9. Submitted and Default Populations for SourceTypeID=52

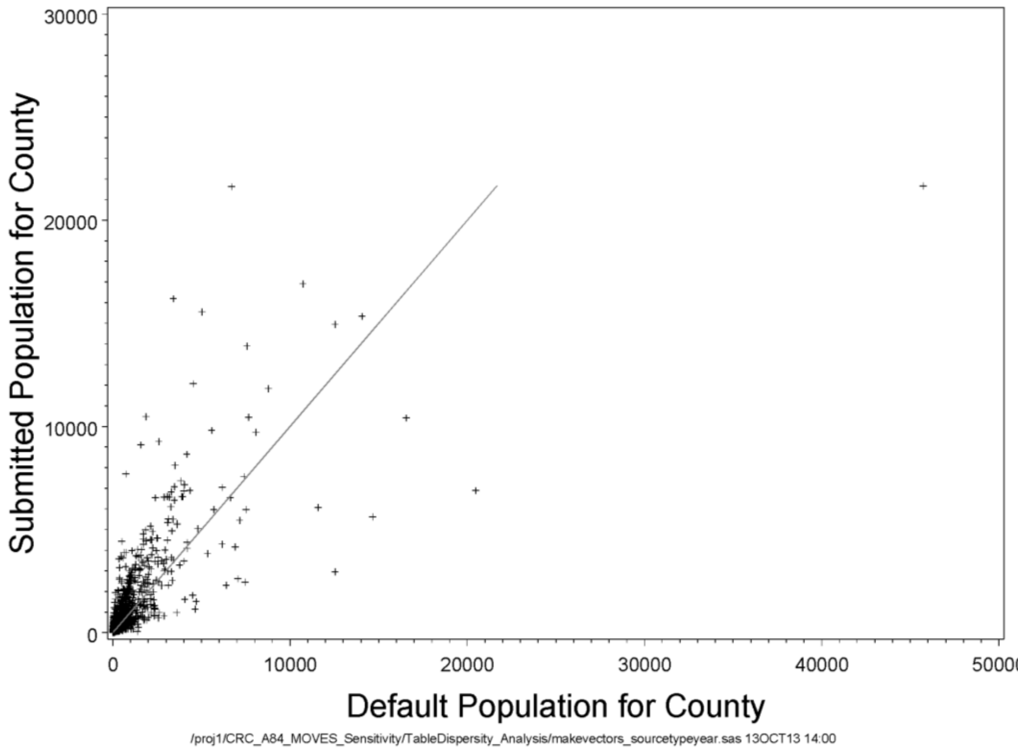


Figure B-10. Submitted and Default Populations for SourceTypeID=53

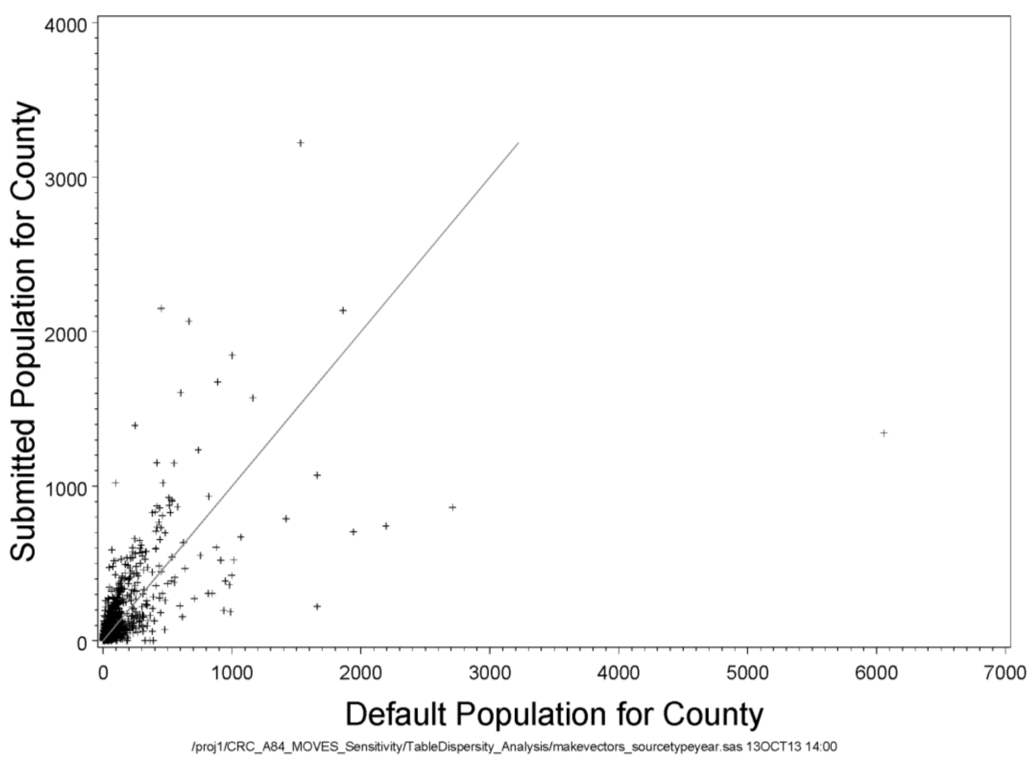


Figure B-11. Submitted and Default Populations for SourceTypeID=54

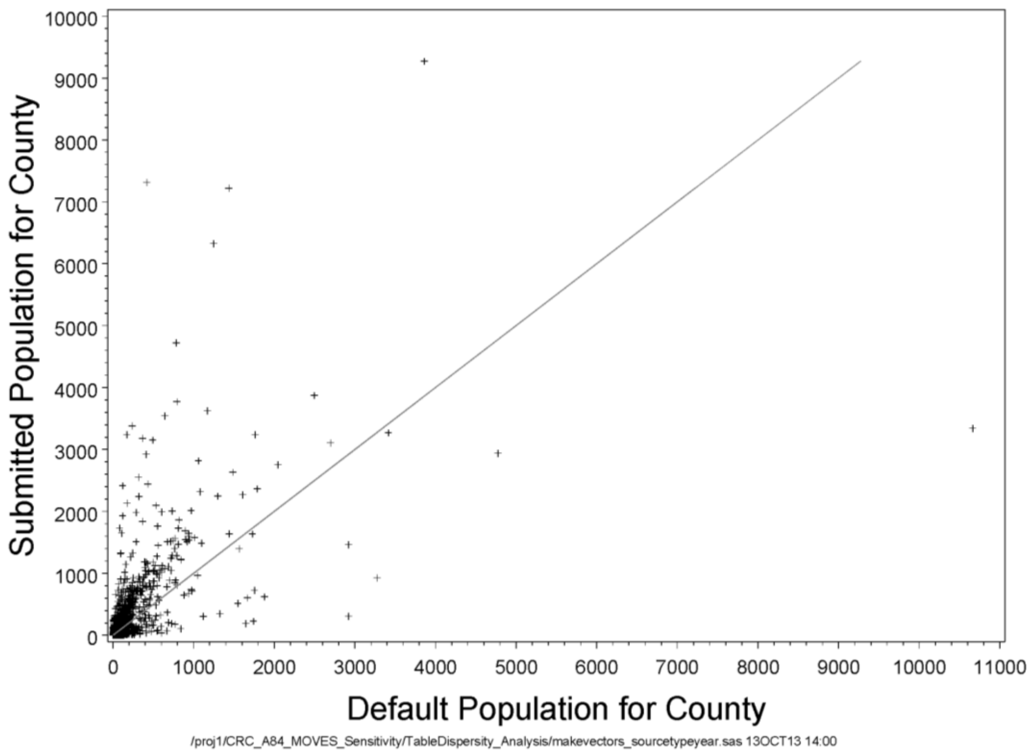


Figure B-12. Submitted and Default Populations for SourceTypeID=61

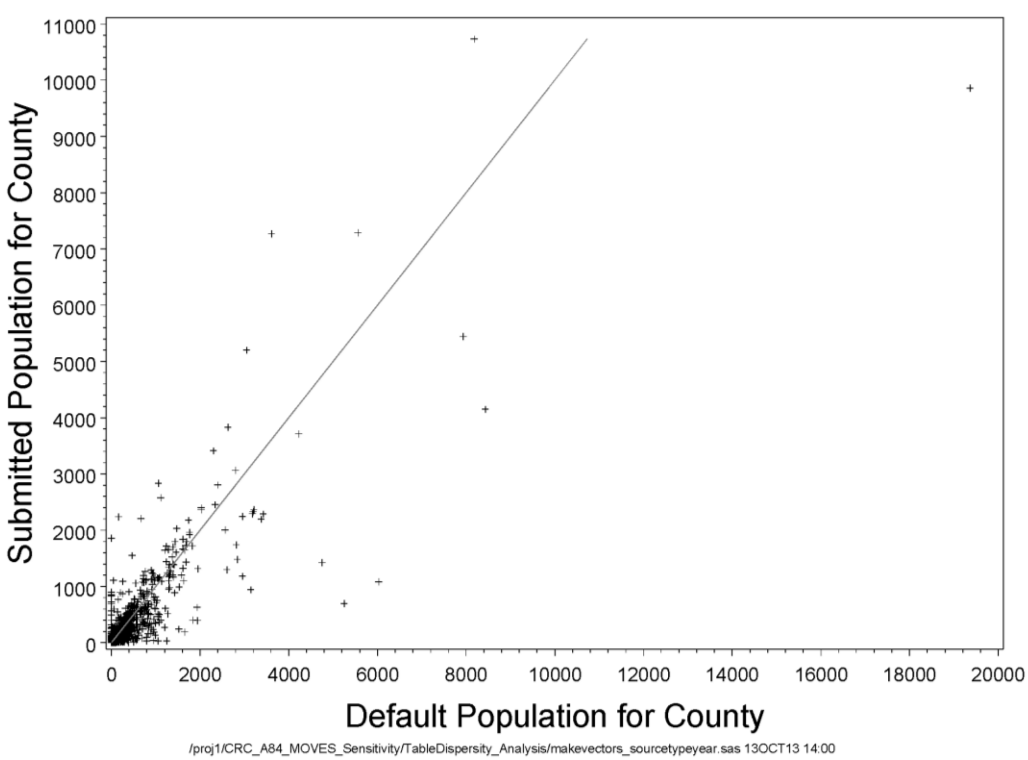
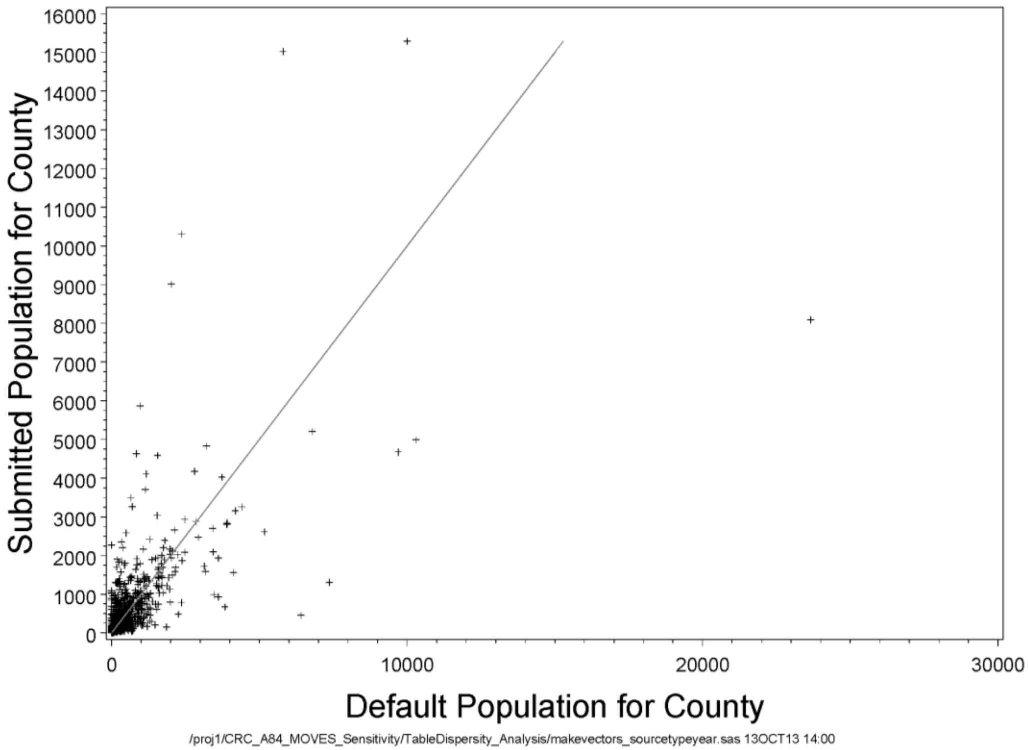


Figure B-13. Submitted and Default Populations for SourceTypeID=62



Appendix C

Figure C-1. Submitted and Default VMT for HPMSVtype=10

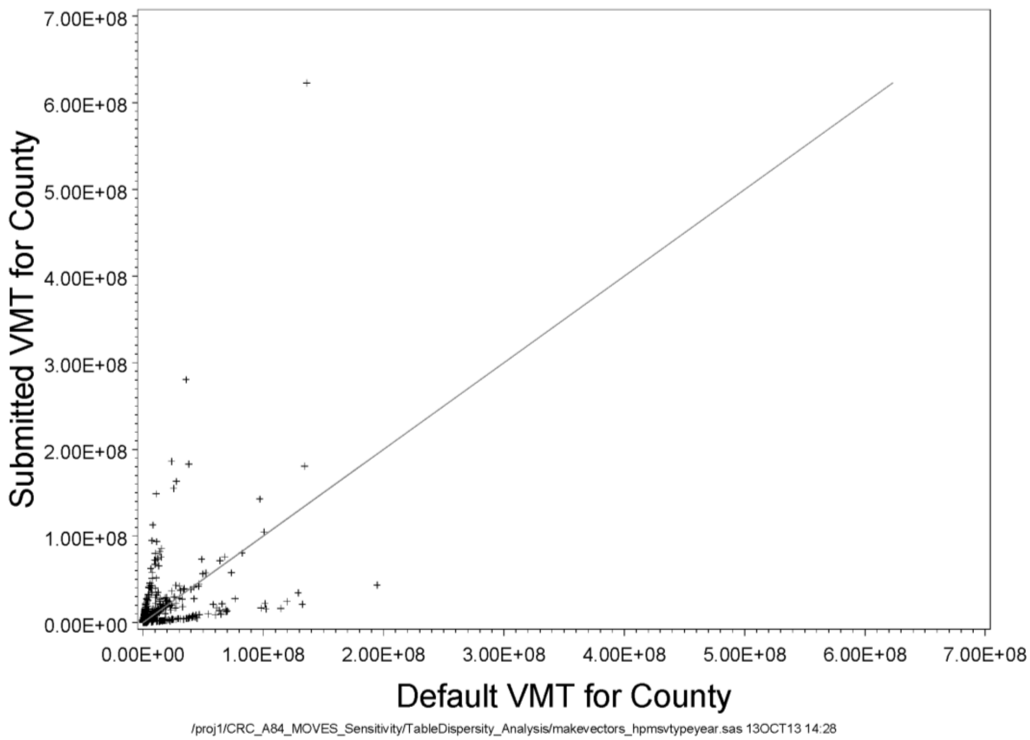


Figure C-2. Submitted and Default VMT for HPMSVtype=20

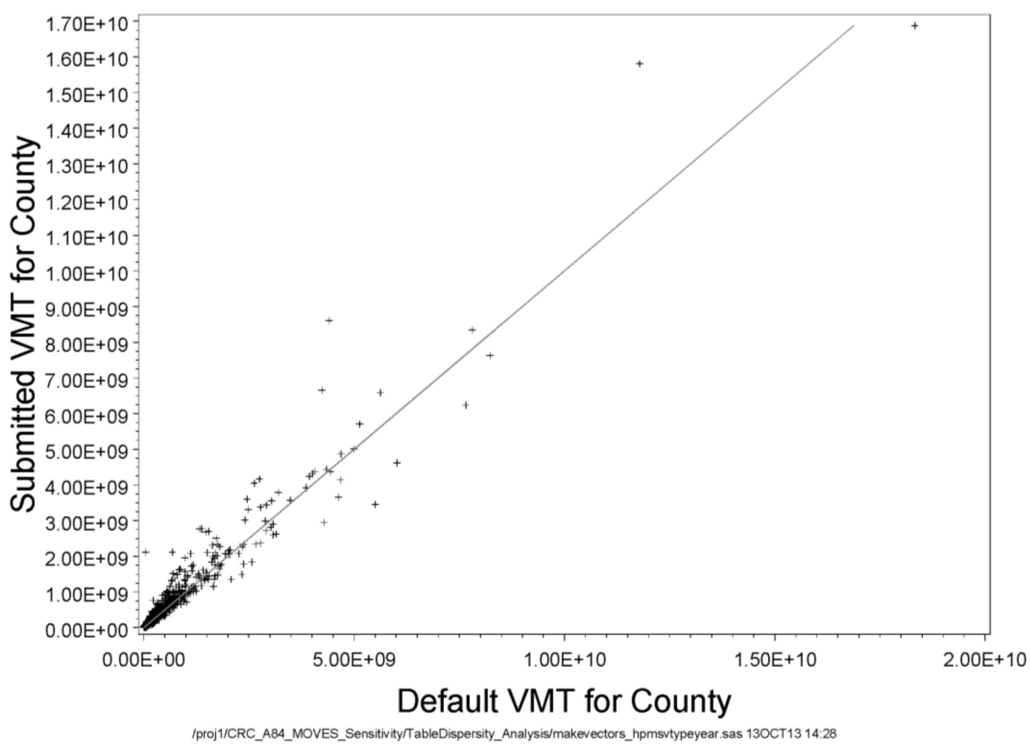


Figure C-3. Submitted and Default VMT for HPMSVtype=30

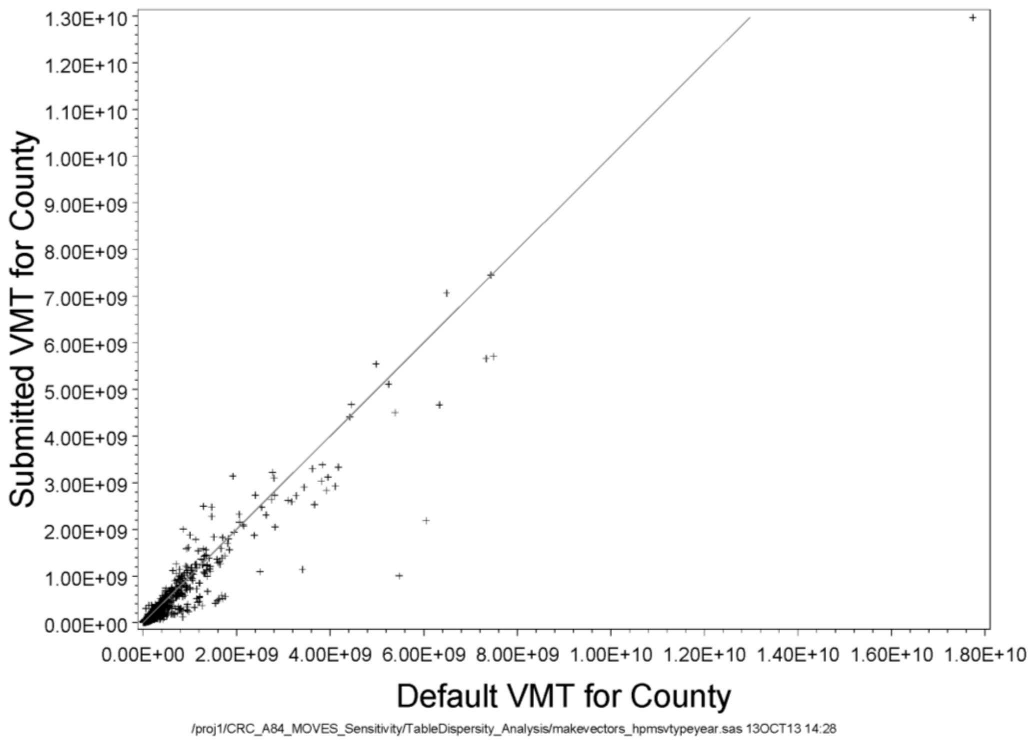


Figure C-4. Submitted and Default VMT for HPMSVtype=40

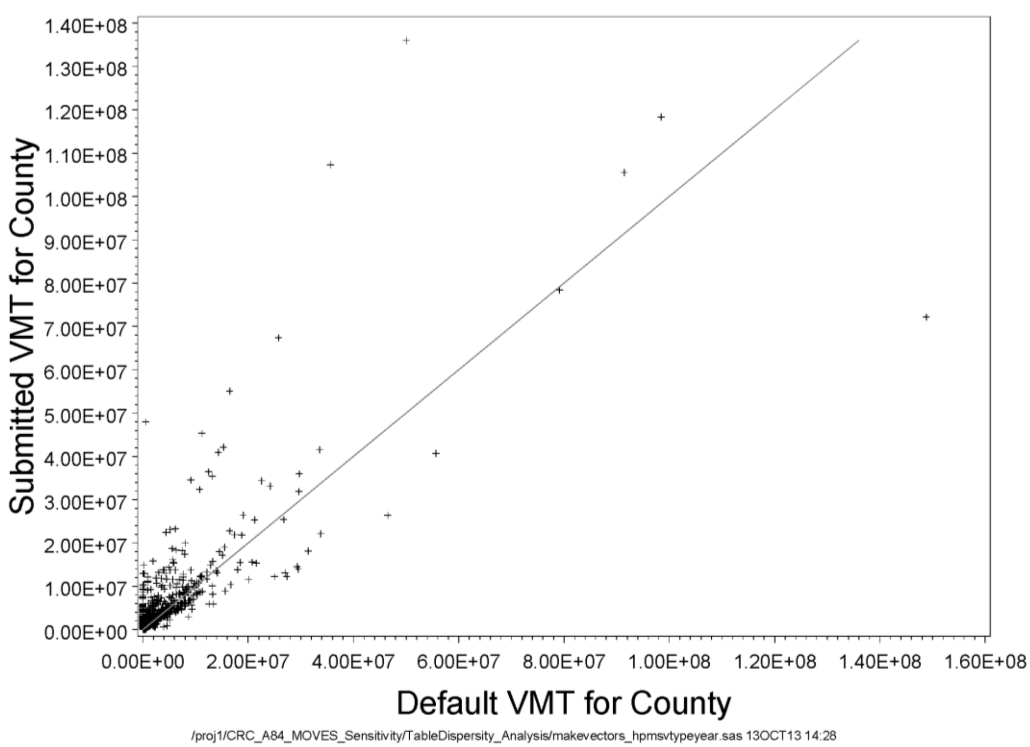


Figure C-5. Submitted and Default VMT for HPMSVtype=50

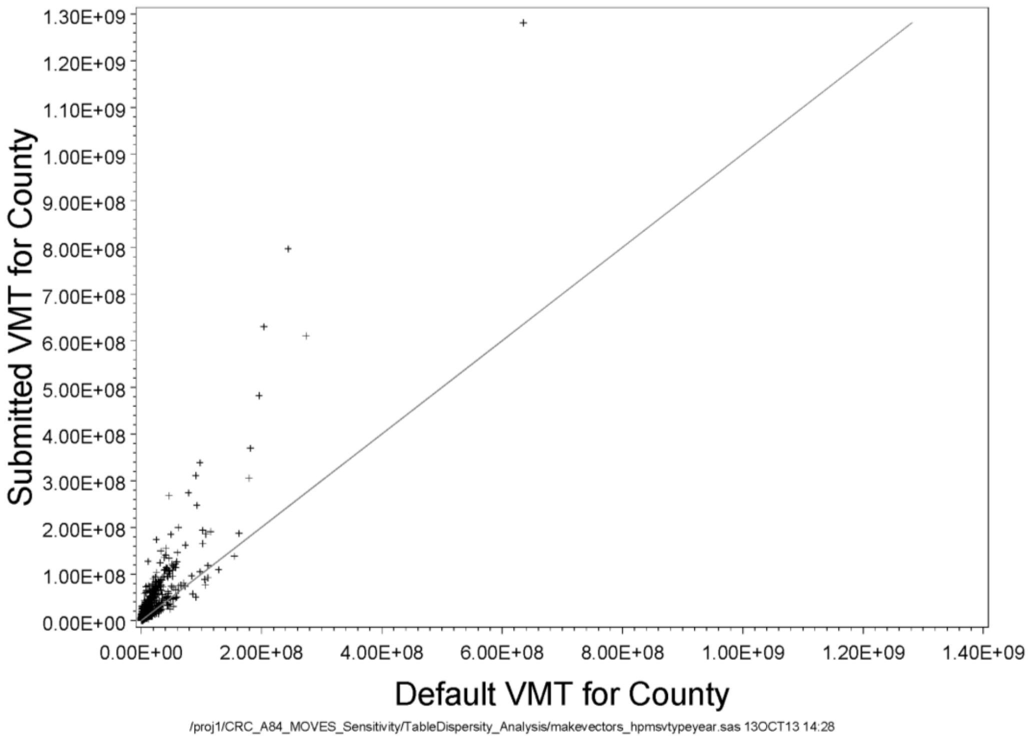
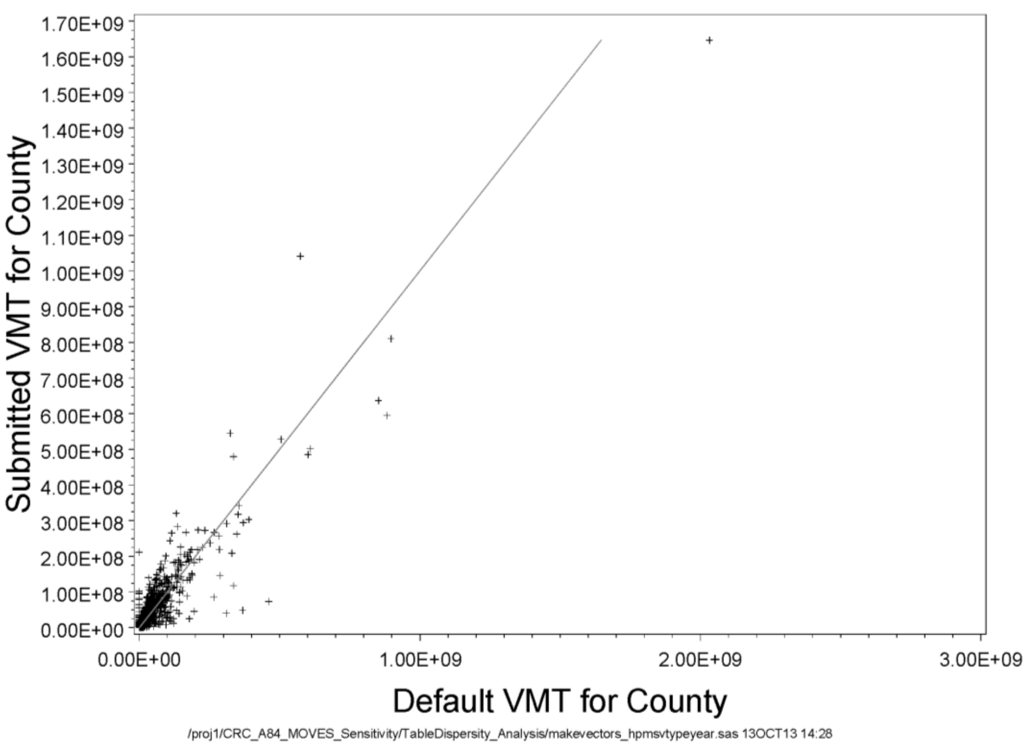


Figure C-6. Submitted and Default VMT for HPMSVtype=60



Appendix D

Figure D-1. Vector Angle vs. Mean RoadTypeID, for SourceTypeID=11

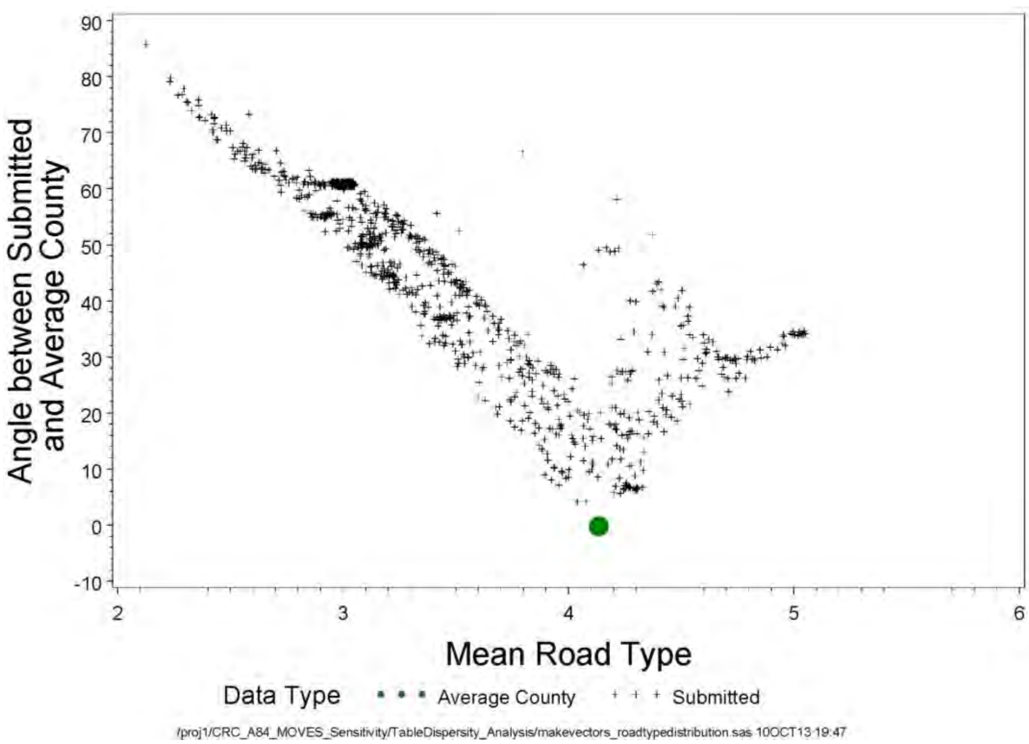


Figure D-2. Vector Angle vs. Mean RoadTypeID, for SourceTypeID=21

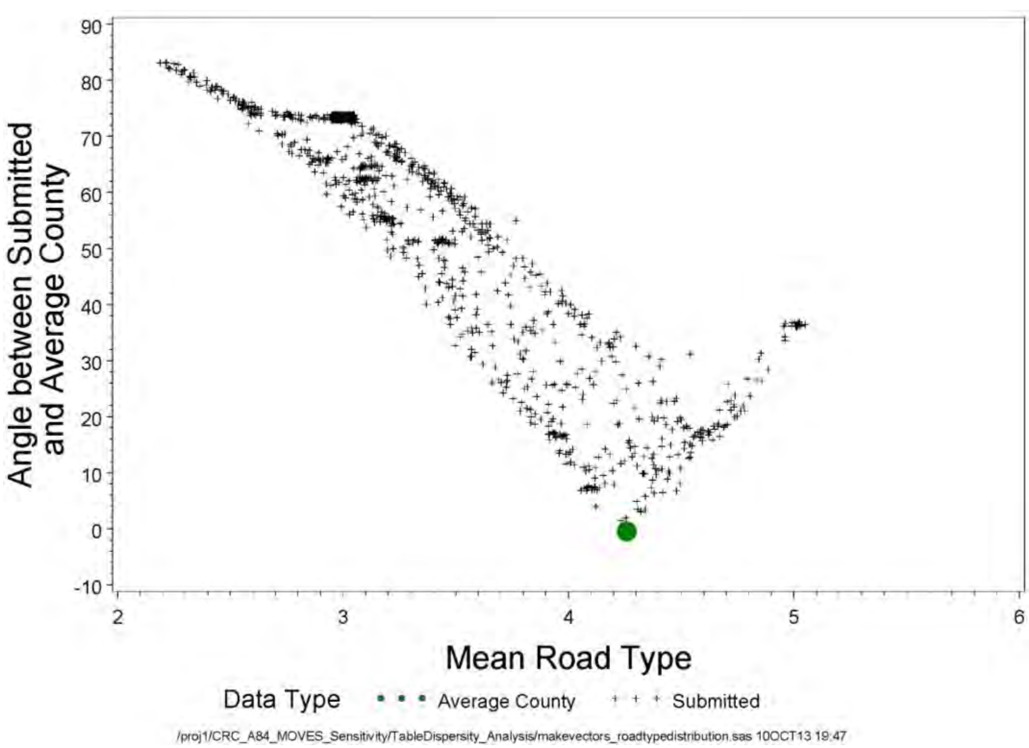


Figure D-3. Vector Angle vs. Mean RoadTypeID, for SourceTypeID=31

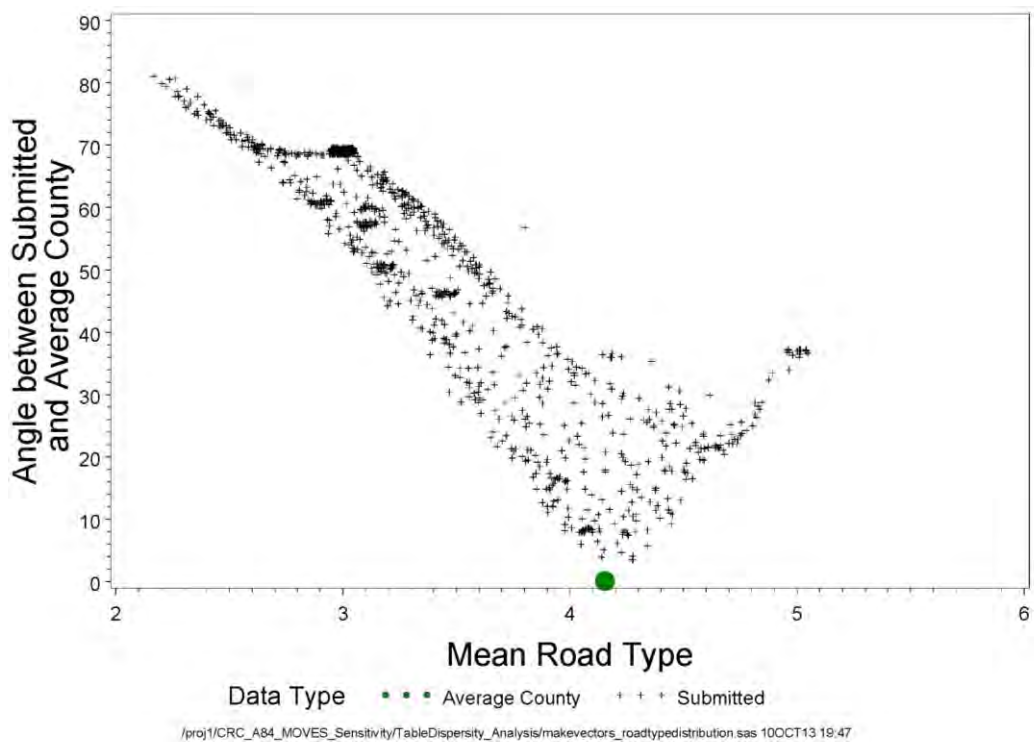


Figure D-4. Vector Angle vs. Mean RoadTypeID, for SourceTypeID=32

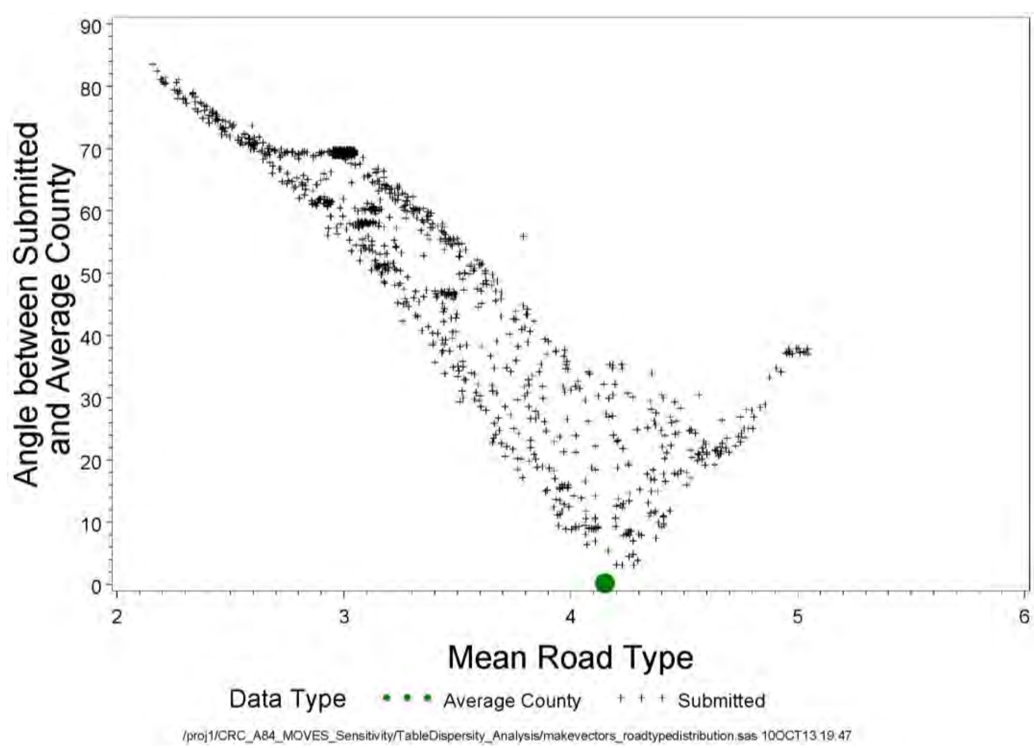


Figure D-5. Vector Angle vs. Mean RoadTypeID, for SourceTypeID=41

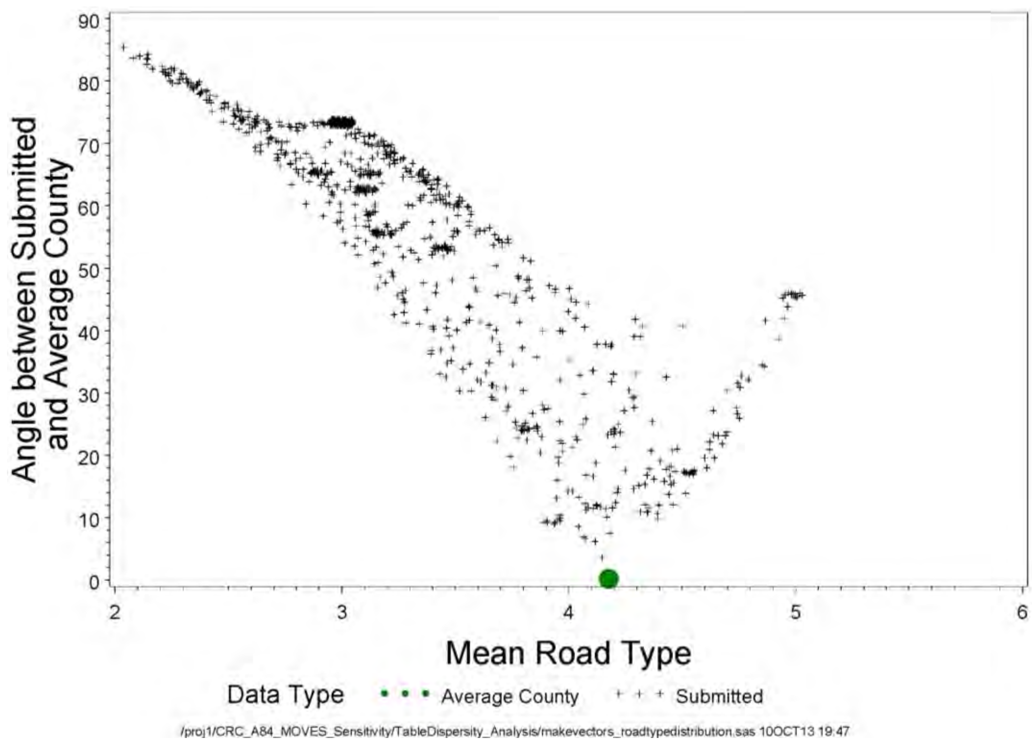


Figure D-6. Vector Angle vs. Mean RoadTypeID, for SourceTypeID=42

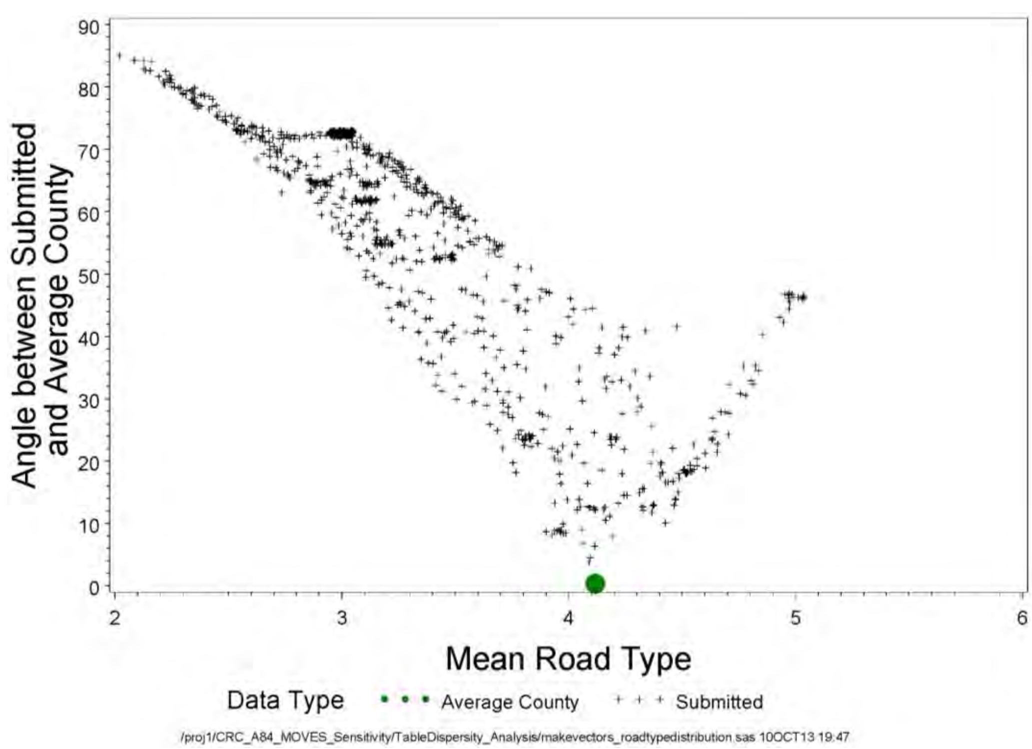


Figure D-7. Vector Angle vs. Mean RoadTypeID, for SourceTypeID=43

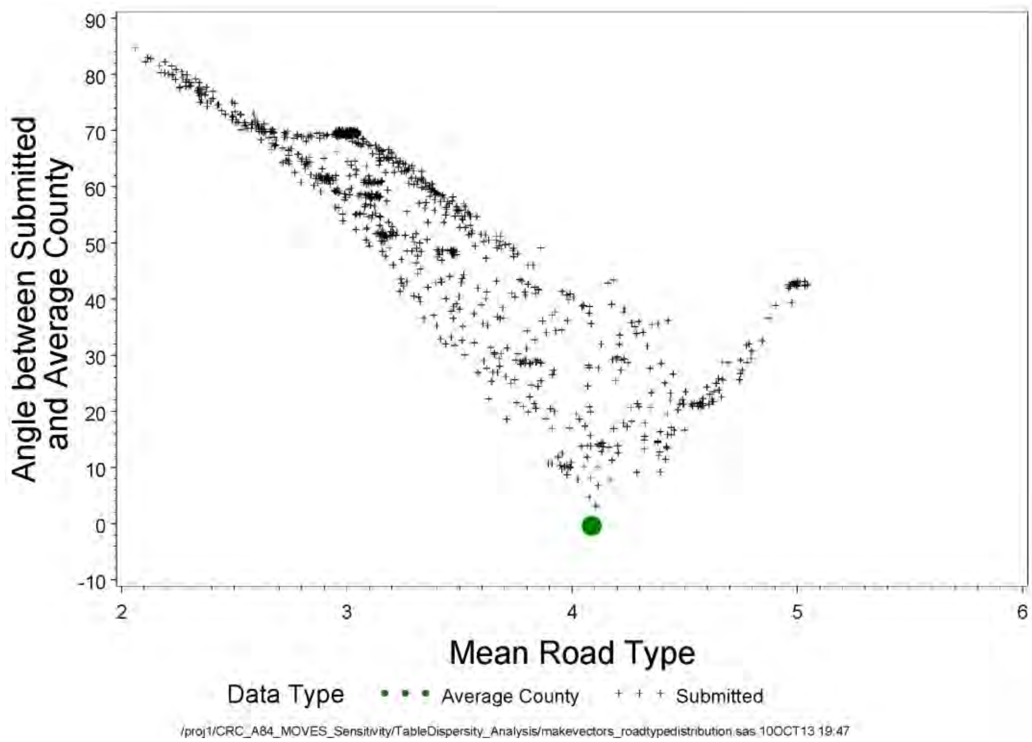


Figure D-8. Vector Angle vs. Mean RoadTypeID, for SourceTypeID=51

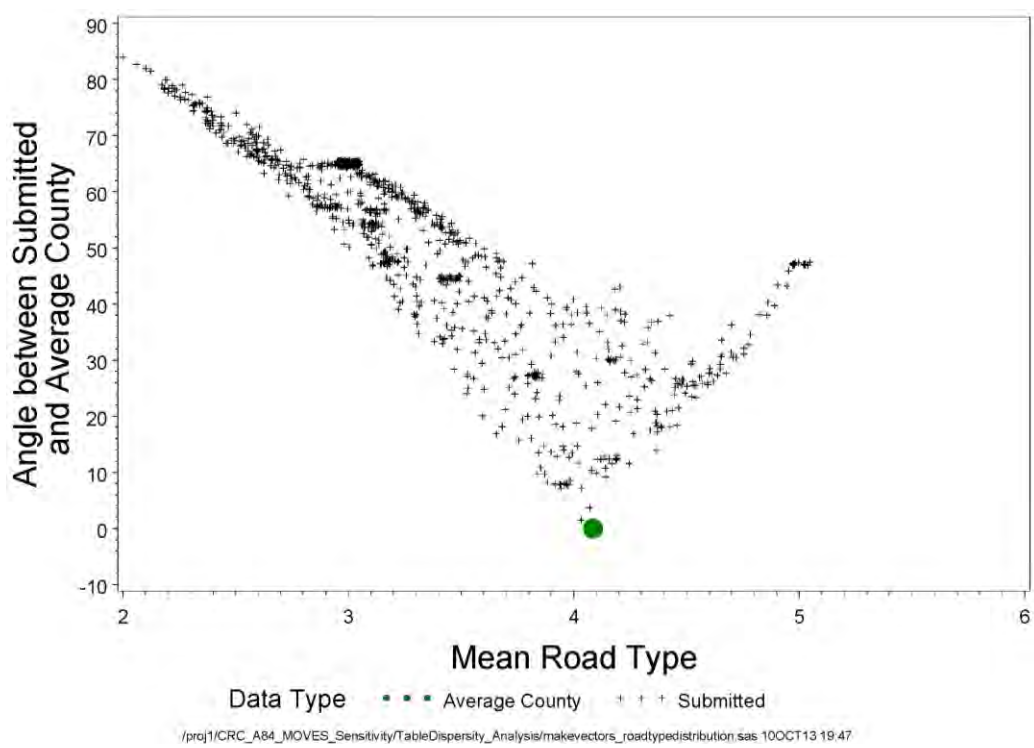


Figure D-9. Vector Angle vs. Mean RoadTypeID, for SourceTypeID=52

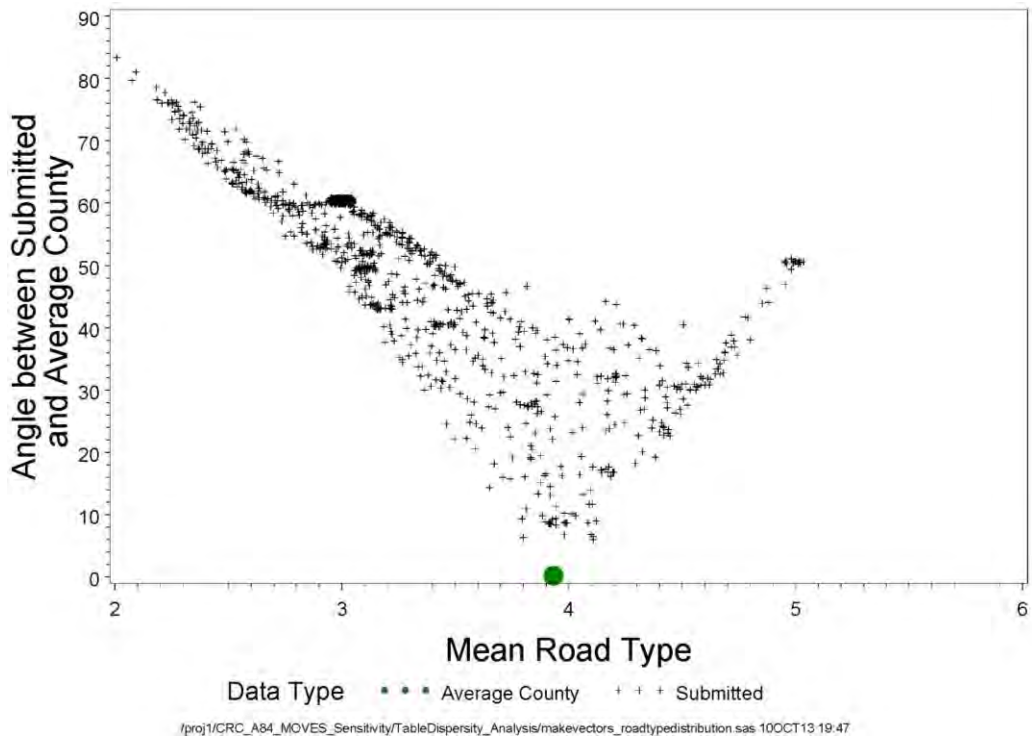


Figure D-10. Vector Angle vs. Mean RoadTypeID, for SourceTypeID=53

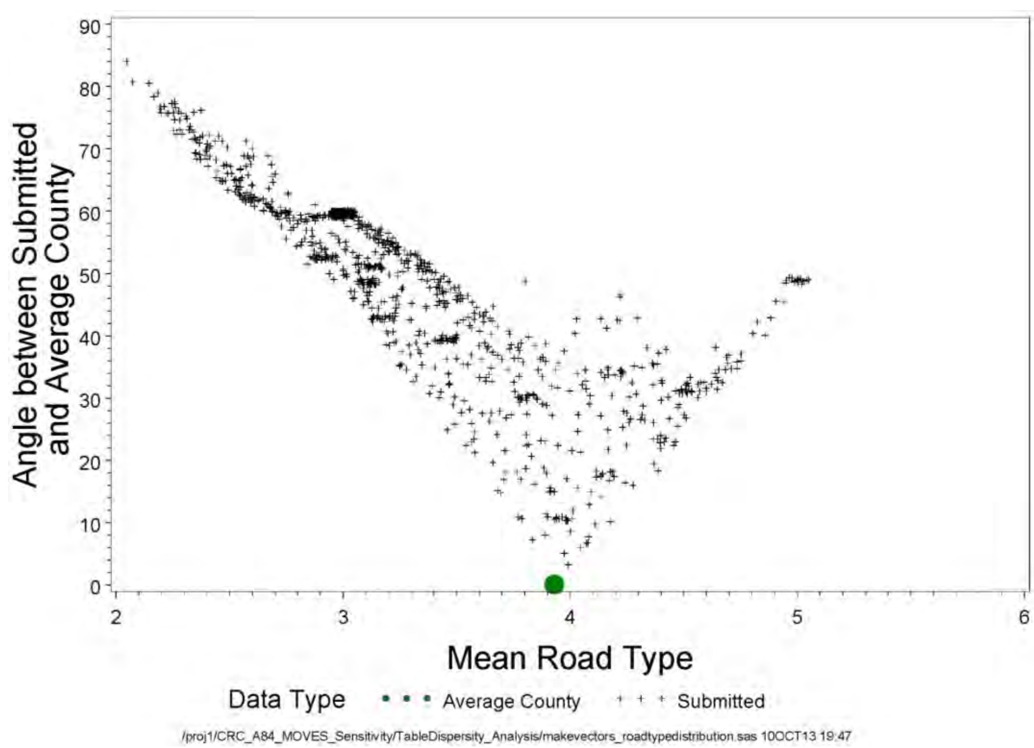


Figure D-11. Vector Angle vs. Mean RoadTypeID, for SourceTypeID=54

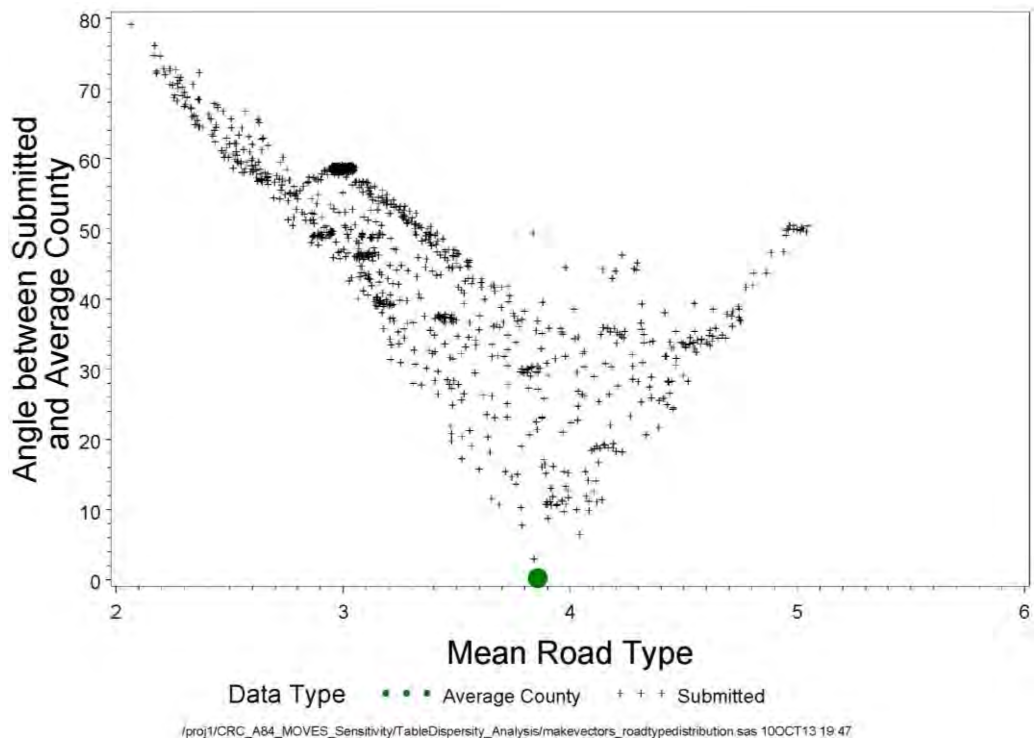


Figure D-12. Vector Angle vs. Mean RoadTypeID, for SourceTypeID=61

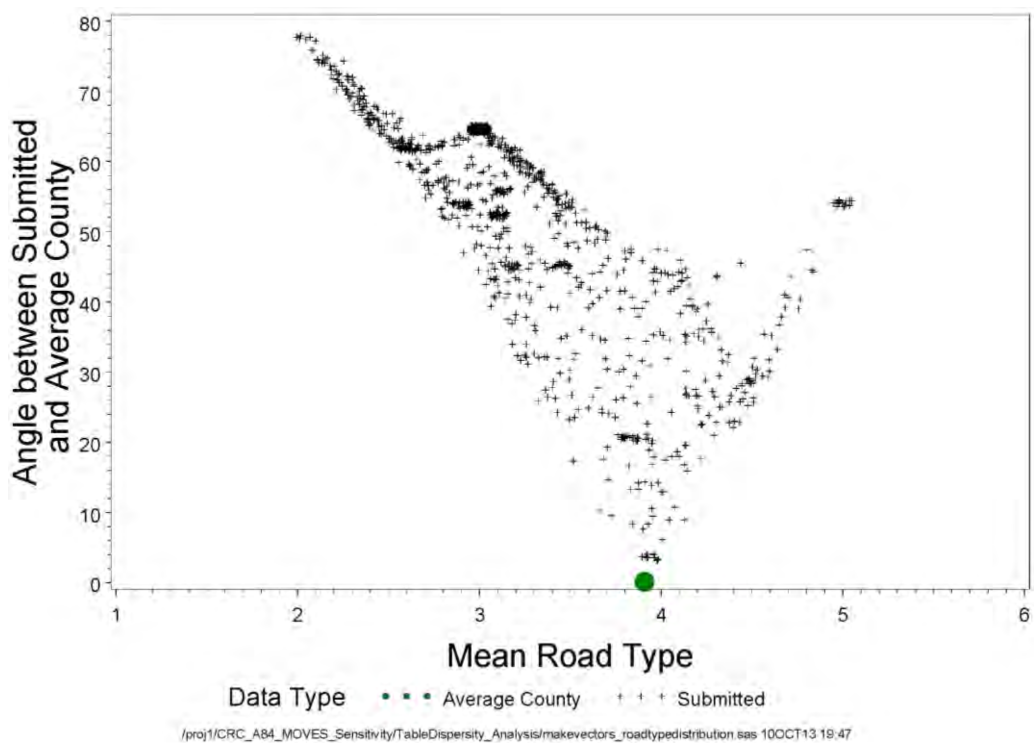
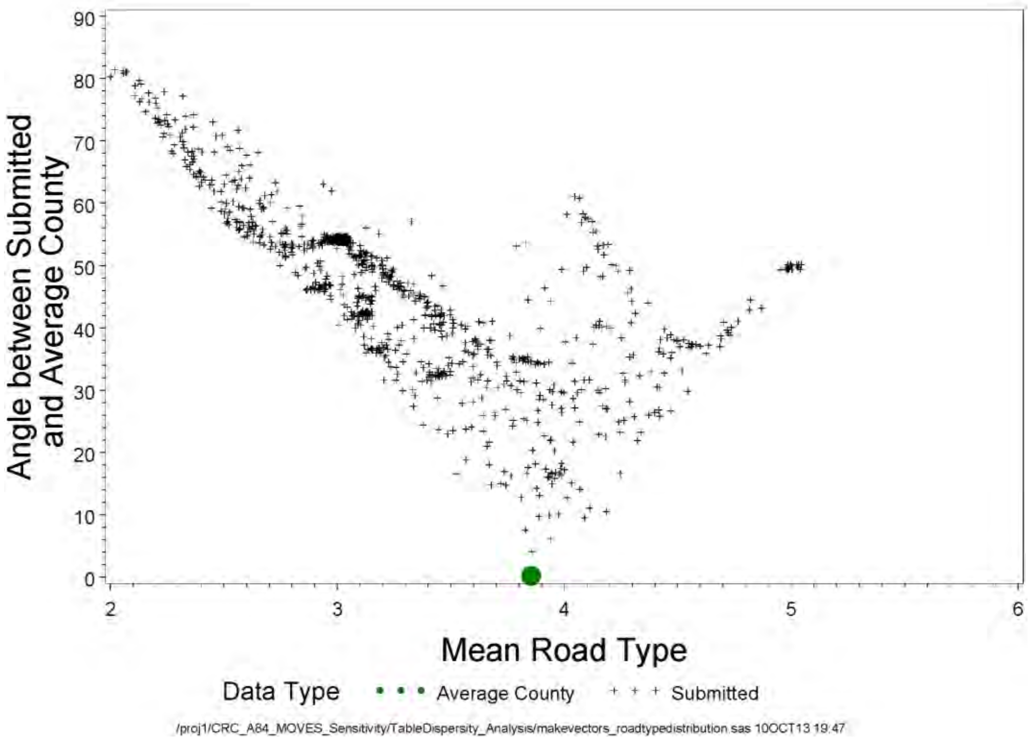


Figure D-13. Vector Angle vs. Mean RoadTypeID, for SourceTypeID=62



Appendix E

Figure E-1. Vector Angle vs. Mean Speed Bin, SpeedDistCluster 1 for SourceTypeID=21

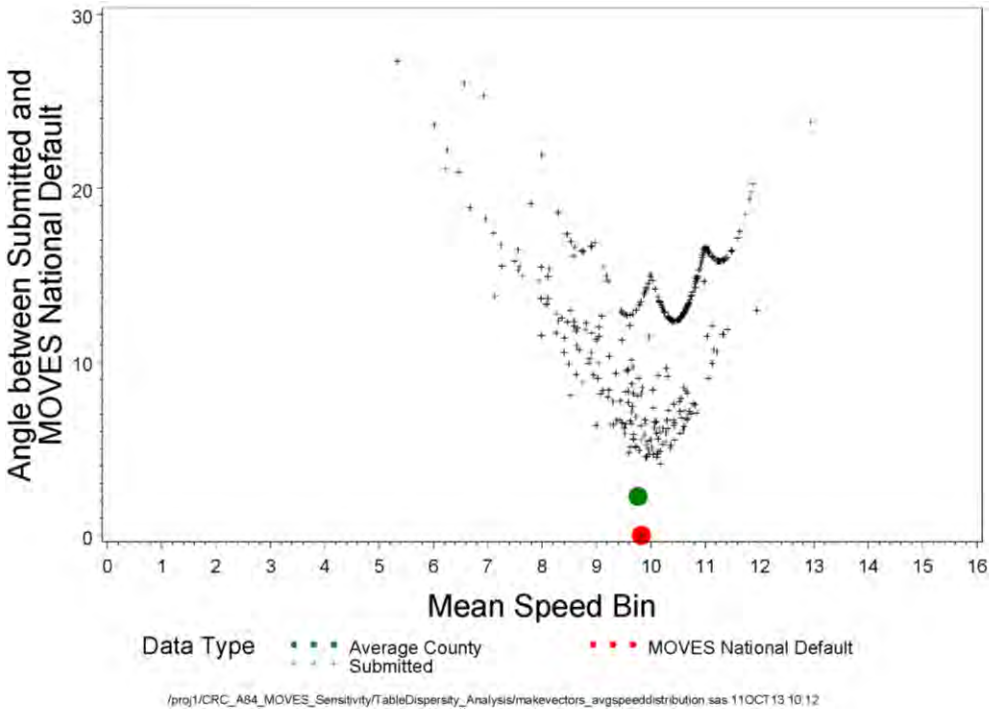


Figure E-2. Vector Angle vs. Mean Speed Bin, SpeedDistCluster 2 for SourceTypeID=21

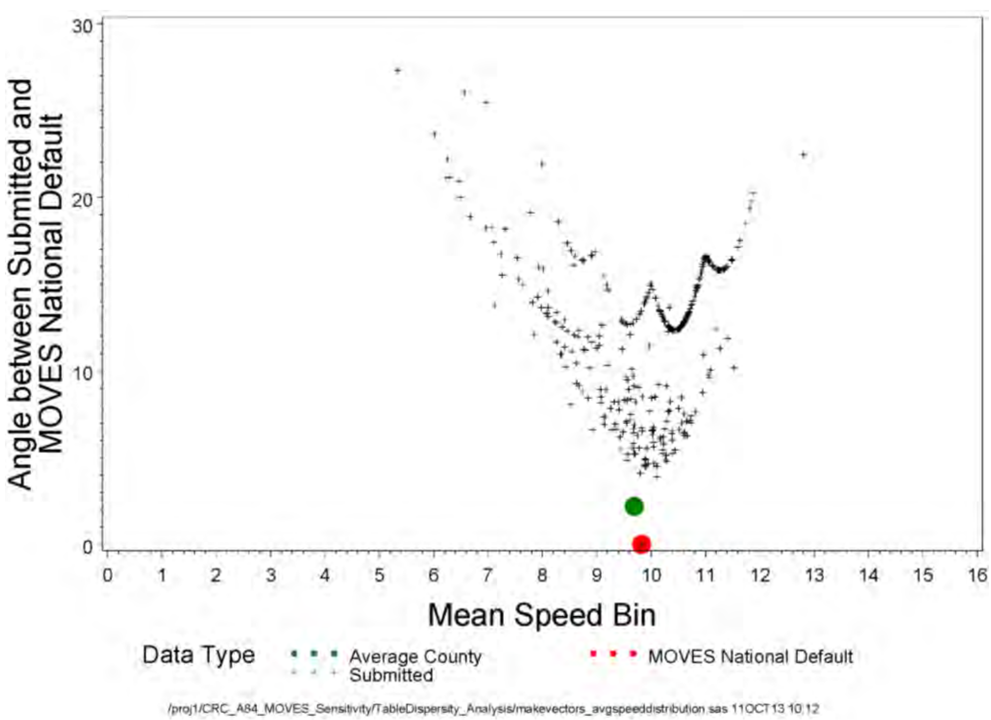


Figure E-3. Vector Angle vs. Mean Speed Bin, SpeedDistCluster 3 for SourceTypeID=21

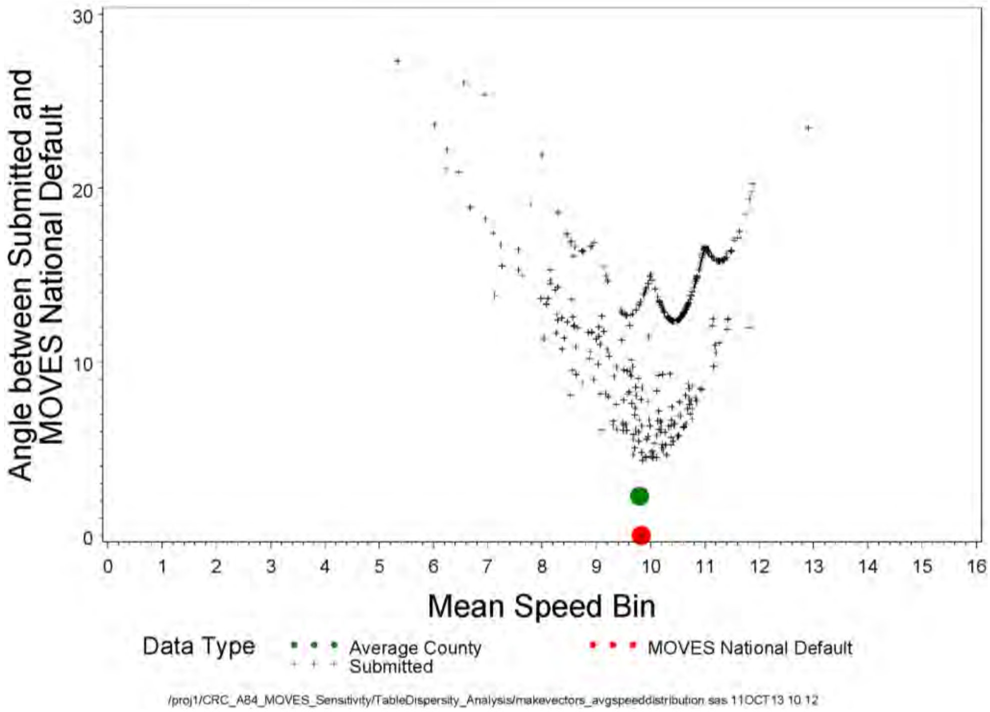


Figure E-4. Vector Angle vs. Mean Speed Bin, SpeedDistCluster 4 for SourceTypeID=21

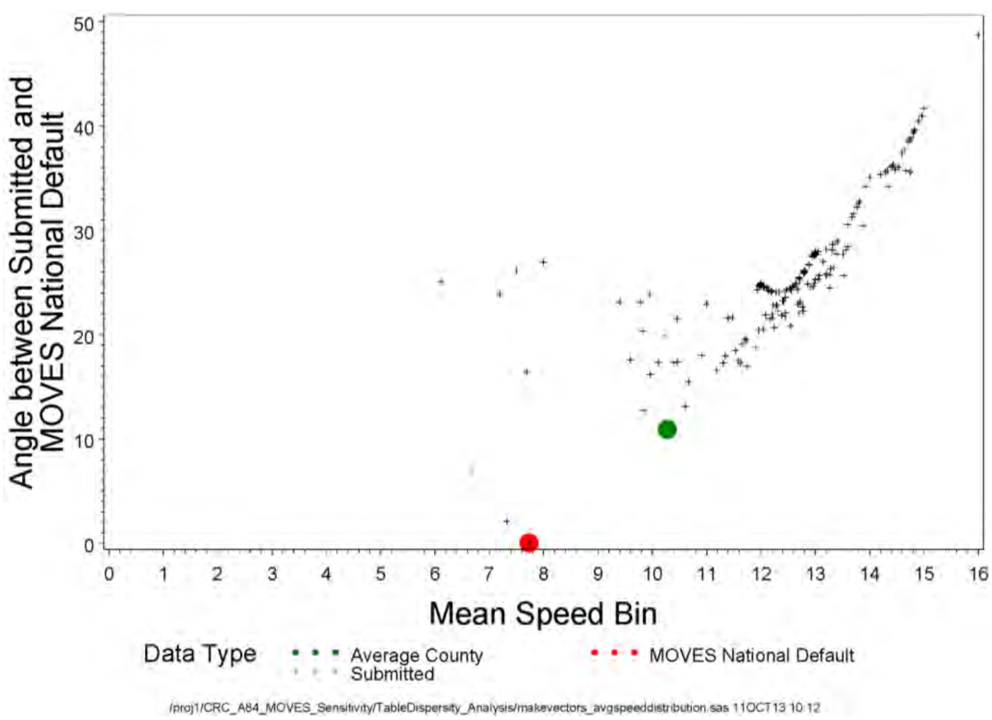


Figure E-5. Vector Angle vs. Mean Speed Bin, SpeedDistCluster 5 for SourceTypeID=21

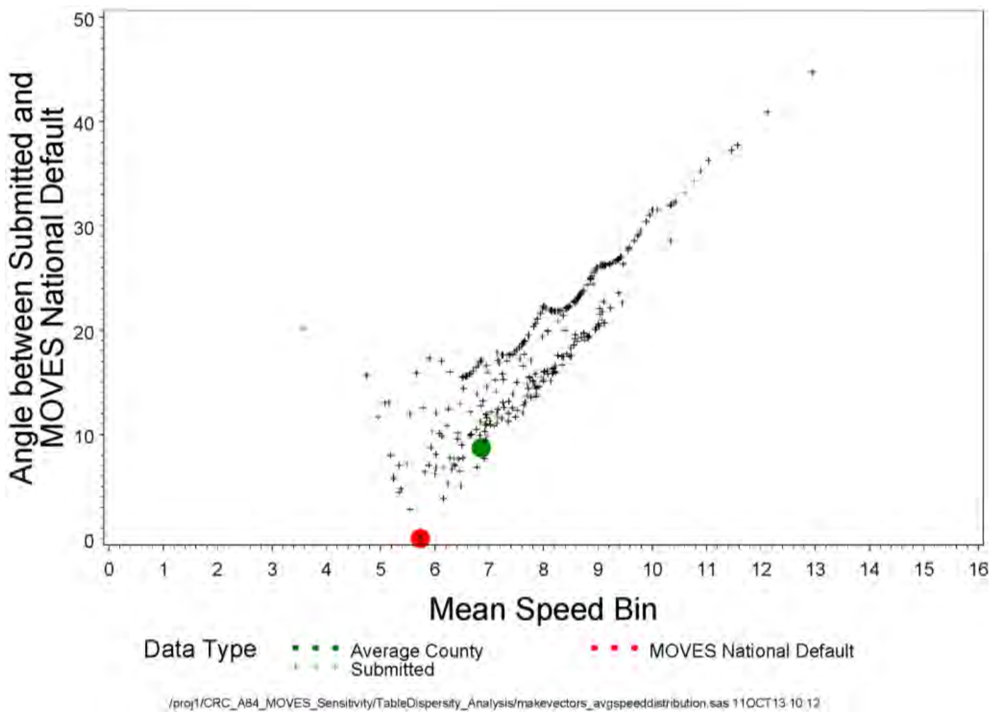


Figure E-6. Vector Angle vs. Mean Speed Bin, SpeedDistCluster 6 for SourceTypeID=21

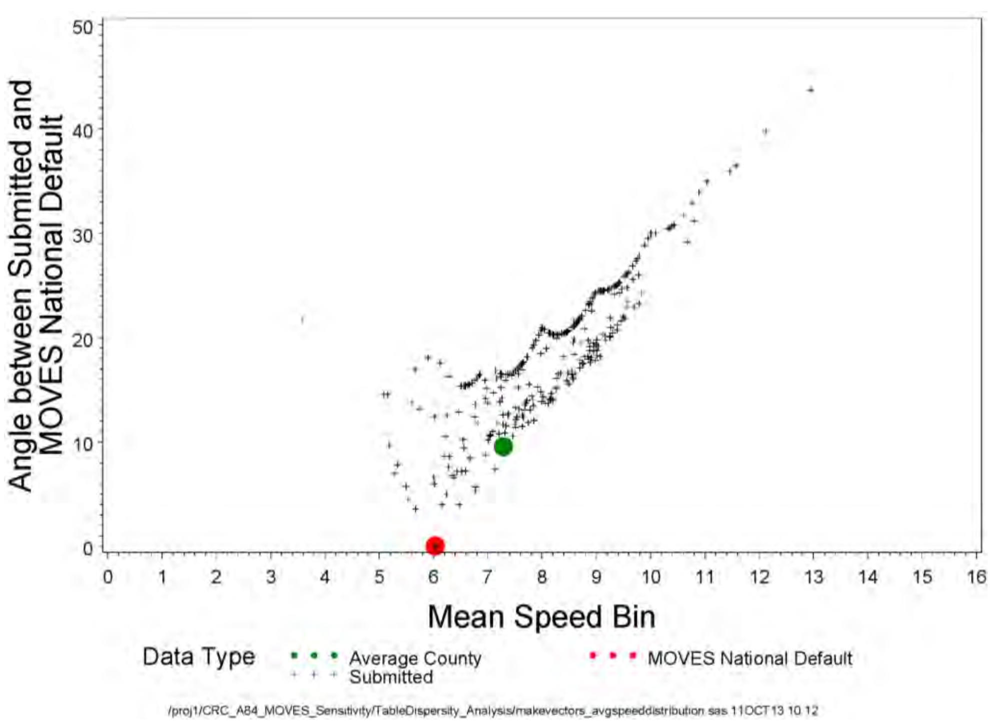


Figure E-7. Vector Angle vs. Mean Speed Bin, SpeedDistCluster 7 for SourceTypeID=21

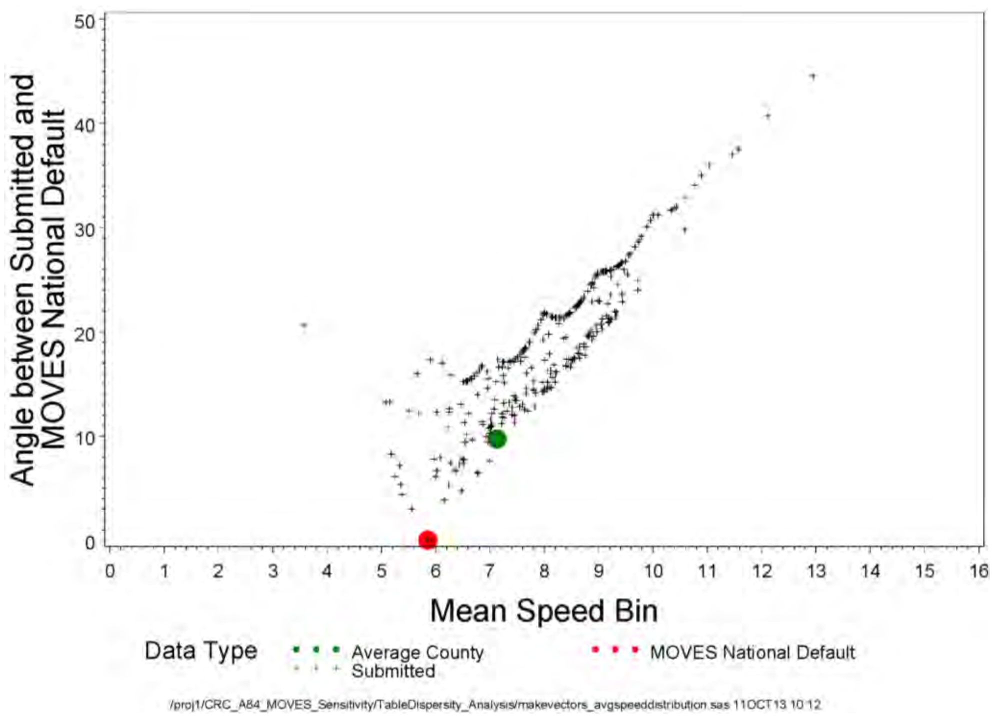


Figure E-8. Vector Angle vs. Mean Speed Bin, SpeedDistCluster 8 for SourceTypeID=21

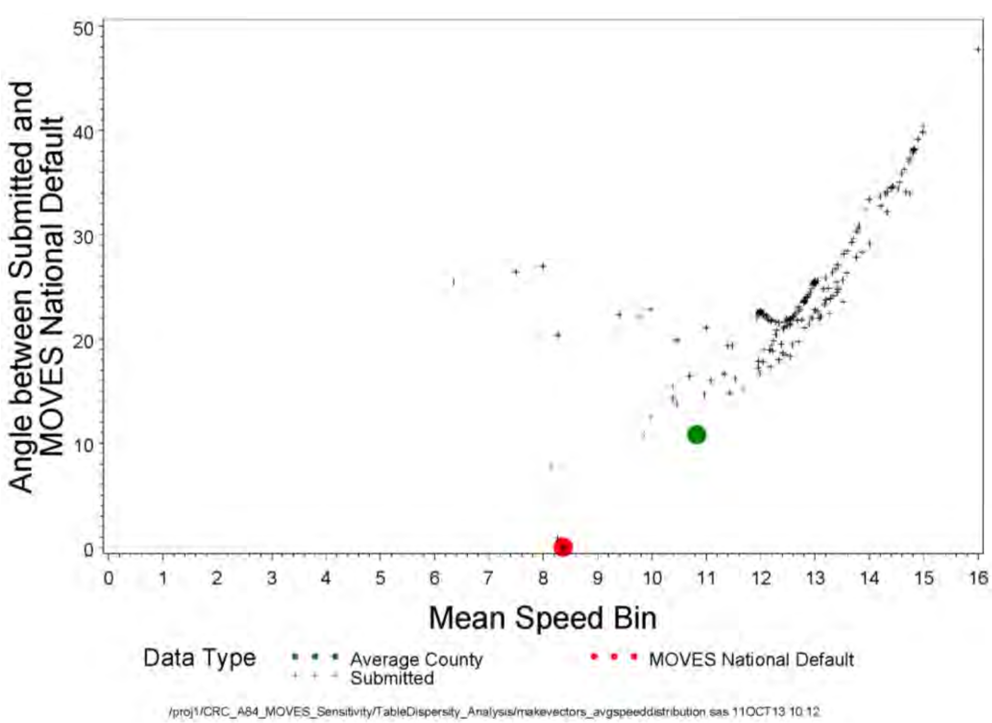


Figure E-9. Vector Angle vs. Mean Speed Bin, SpeedDistCluster 9 for SourceTypeID=21

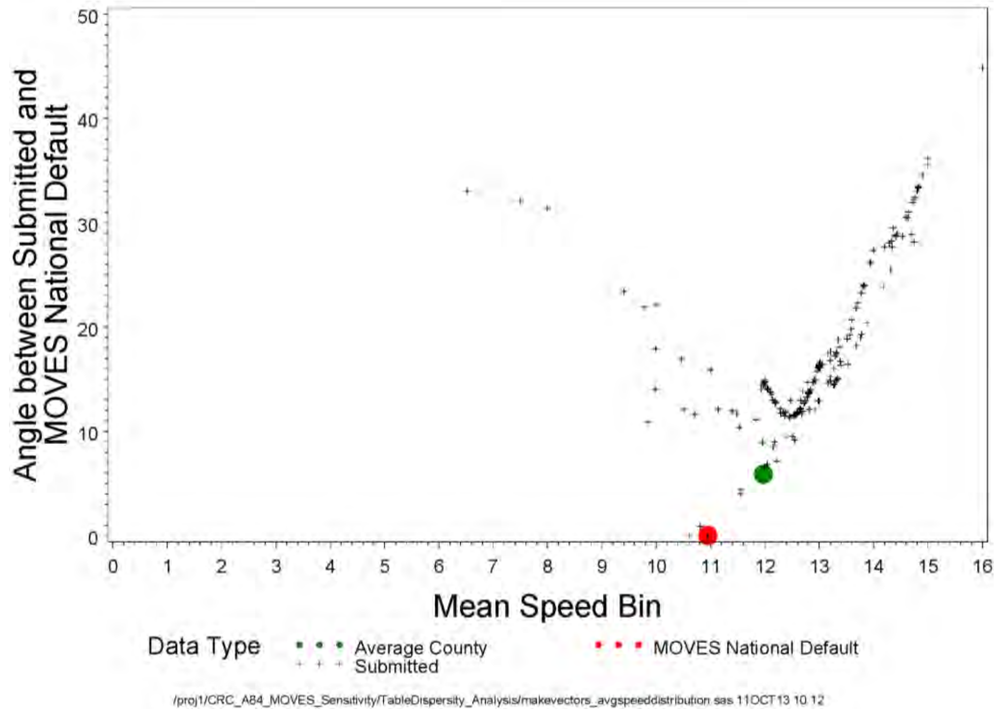


Figure E-10. Vector Angle vs. Mean Speed Bin, SpeedDistCluster 10 for SourceTypeID=21

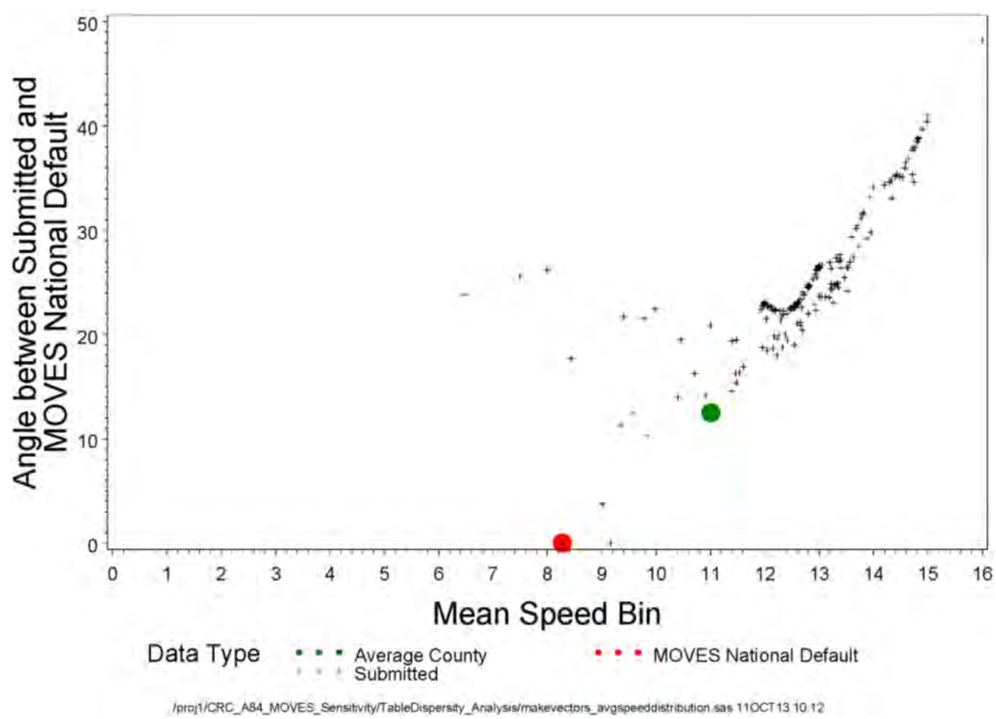


Figure E-11. Vector Angle vs. Mean Speed Bin, SpeedDistCluster 11 for SourceTypeID=21

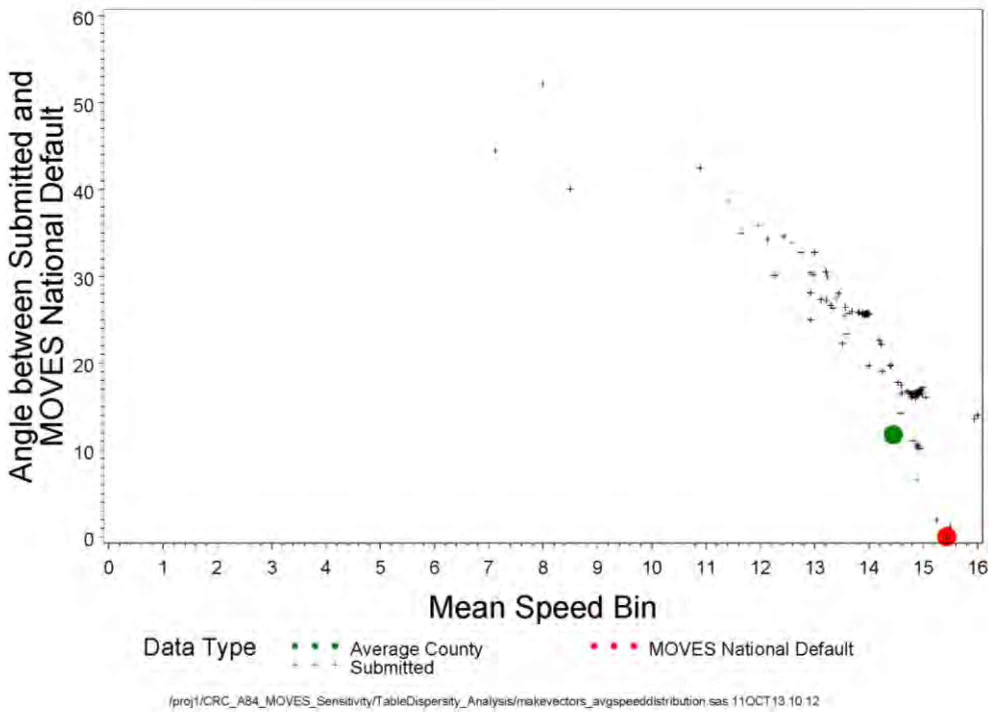


Figure E-12. Vector Angle vs. Mean Speed Bin, SpeedDistCluster 12 for SourceTypeID=21

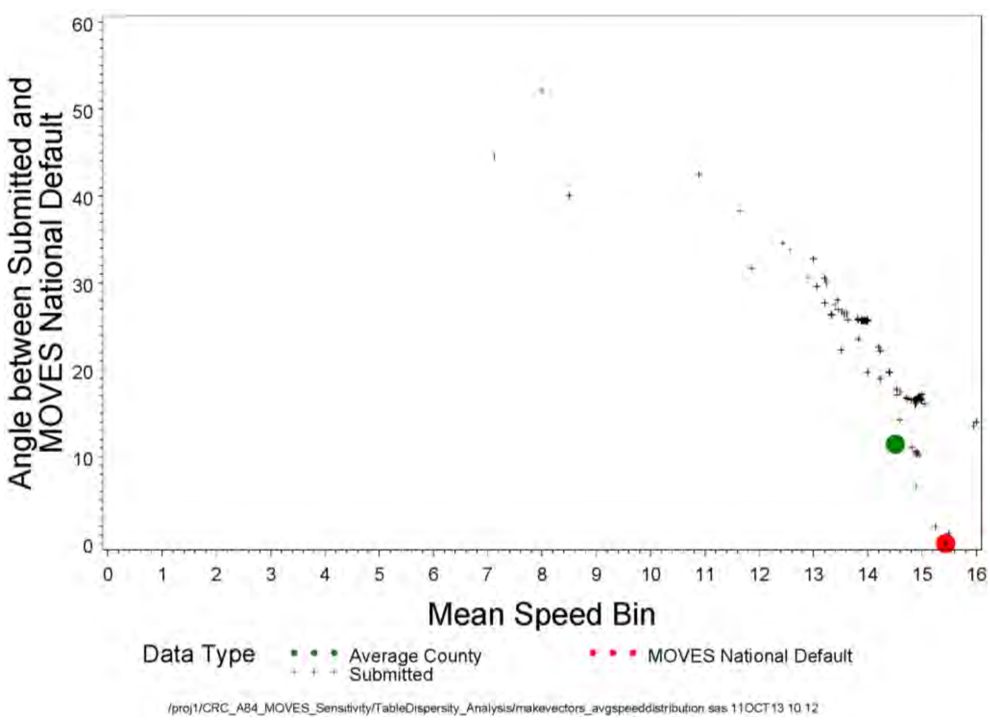


Figure E-13. Vector Angle vs. Mean Speed Bin, SpeedDistCluster 13 for SourceTypeID=21