

MOVES2004 Energy and Emission Inputs

Draft Report

MOVES2004 Energy and Emission Inputs

Draft Report

John Koupal
Larry Landman
Edward Nam
James Warila
Carl Scarbro
Edward Glover
Robert Giannelli

Assessment and Standards Division
Office of Transportation and Air Quality
U.S. Environmental Protection Agency

NOTICE

*This Technical Report does not necessarily represent final EPA decisions or positions.
It is intended to present technical analysis of issues using data that are currently available.*

*The purpose in the release of such reports is to facilitate an exchange of
technical information and to inform the public of technical developments.*

Table of Contents

1. Introduction.....	3
2. Overview of MOVES2004 Energy & Emission Inputs	5
3. Overview of Data Used in MOVES2004.....	5
3.1 Data Gathering and Processing Steps	6
3.2 Data Sources	7
4. Total Energy Rates	9
4.1 Source Bin Definitions	9
4.2 Operating Mode Definitions.....	11
4.3 Running Energy Rate Development.....	12
4.4 Start Energy Rate Development	36
4.5 Extended Idle.....	44
5. Petroleum and Fossil Energy Calculations	46
6. Carbon Dioxide (CO₂) Calculations.....	47
7. Methane (CH₄) and Nitrous Oxide (N₂O) Rates	49
7.1 Source Bin Definitions	49
7.2 Operating Mode Definitions.....	49
7.3 Data Sources	50
7.4 Running and Start Rate Development	50
7.5 Advanced Technologies & Alternative Fuels	55
8. CO₂ Equivalent Calculation	56
9. Adjustments.....	57
9.1 Temperature.....	57
9.2 Air Conditioning	64

10. Well-To-Pump Energy & Emission Rates.....	65
11. Data Source Identification	67
12. Acknowledgments	69
Appendix A: Binning Methodology Proof-Of-Concept	70
Appendix B: Calculation of Running Energy Consumption Rates Using “The Binner Program”	82
Appendix C: Proof-Of-Concept Assessment of Hole Filling Methods	104
Appendix D: Algorithm for Running Energy Hole Filling Using Interpolation / Copying	120
Appendix E: Start Energy Rates	135
Appendix F: CH₄ & N₂O Rates by Model Year.....	139
Appendix G: Comments on Emission Analysis Plan.....	147
Appendix H: Pre-Publication Peer Review Comments	159

1. Introduction

The basis of energy and emission estimation in MOVES2004 is the energy and emission rates associated with each pollutant (including energy)^a and emission process estimated by MOVES. This report serves to document the development of these inputs. Included in this report are the structure of emissions rates in terms of vehicle and operating mode characterization, including supporting analyses; data sources used in the development of emission rates; methods for producing emission rates, including the “binning” process, and “hole filling” where adequate data were not available for a particular rate; and advanced technology modeling using the Physical Emission Rate Estimator (PERE) and other sources.

Energy and emission rates for all pollutants and processes are stored in a table in the MOVES Default database named EmissionRate, with the exception of well-to-pump energy and emissions, which are stored the table GREETWellToPump. All of the work described in this report pertains to the development of entries in EmissionRate or GREETWellToPump. These tables are therefore considered the “results” of the work presented in this document. The default EmissionRate table has over 20,000 records, and the process to generate advanced technology and future rates adds over 90,000 records, in total accounting for the full range of energy and emission rates by factors including pollutant, process, operating mode, fuel, engine technology, model year group, regulatory class, loaded weight and engine displacement.

MOVES2004 implements a number of “firsts” with regard to EPA’s mobile source emission models, including: modeling energy consumption, N₂O and CH₄ explicitly; defining rates by speed and power-based operating modes; modeling a broad array of advanced technology vehicles, including several permutations of hybrid applications; modeling periods of extended idling (e.g. heavy-duty “hoteling”) as a separate emission process; relying primarily on second-by-second data to develop emission rates; and including well-to-pump energy emission estimates to enable well-to-wheel (e.g. life cycle) analysis. Several of these features will be carried over and expanded in the continued development of MOVES over the next few years. Thus, in addition to documenting the development of specific inputs used in MOVES2004, this report presents the foundation of methodologies which will be applied to future versions of MOVES.

A report entitled “Draft Emission Analysis Plan for MOVES GHG”, published in Fall 2002, contains much of the background for the studies and methods documented here. This report, referred to herein as the “emission analysis plan”, underwent stakeholder and formal peer review in accordance with EPA Peer Review guidelines. Appendix G of this report contains a summary of comments from these review processes along with responses to these comments. Because much of the background for the

^a For brevity, within this report energy consumption is sometimes grouped within the term “emission” or “pollutant”.

MOVES emission analysis work is covered in the emission analysis plan, the reader is encouraged to consult with the emission analysis plan to provide a broader context for the work presented here. A copy of the emission analysis plan may be requested directly via email to mobile@epa.gov.

A pre-publication version of this report underwent formal peer review by Professor Lawrence Caretto of California State University Northridge. The resulting comments and our responses to these comments are contained in Appendix H.

This report is one of a series which documents the design, technical inputs and use of MOVES2004. More detail on some of the concepts and methodologies presenting in this report are found in the following reports:

- “MOVES2004 Software Design Reference Manual” (referenced herein as the Design Manual), which provides detail on MOVES design, mathematical formulation and database structure.
- “MOVES2004 Highway Vehicle Population and Activity Data” (referenced herein as the Fleet and Activity Report), which provides detail on methodology and data used to generated inputs related to parameters such as vehicle population, vehicle miles traveled, average speed distributions, etc.
- “Fuel Consumption Modeling of Diesel, Motorcycle, and Advanced Technology Vehicles in the Physical Emission Rate Estimator (PERE)” (referenced herein as the PERE report), which provides detail on the PERE model used to fill data holes and generate advanced technology rates.
- “Update of Methane and Nitrous Oxide Emission Factors for On-Highway Vehicles”, which provides detail on the analysis of CH₄ and N₂O test data used to generate MOVES2004 emission rates.
- “GREET User Manual and Technical Input for MOVES Integration” (referenced herein as the GREET report), which provides detail on the use of and default assumptions in the version of Argonne National Laboratories’ GREET model developed to to provide well-to-pump energy and emissions in MOVES2004

A more general overview of the MOVES model and documentation is contained in the document “A Roadmap to MOVES2004”. These reports will be available on the MOVES website, <http://www.epa.gov/otaq/ngm.htm>, or may be requested directly via email to mobile@epa.gov

2. Overview of MOVES2004 Energy & Emission Inputs

Emission rates in MOVES2004 are categorized according to pollutant and emission process. The pollutants in MOVES2004 are: total energy, fossil-based energy, petroleum-based energy, methane (CH₄) and nitrous oxide (N₂O).^b Of these, however, most are calculated in MOVES2004 from estimates of three “root” pollutants: total energy (from which fossil energy and petroleum energy are calculated), CH₄ and N₂O. Thus, the EmissionRate table only contains rates for total energy, CH₄ and N₂O.

The emission “processes” in MOVES2004 are: running operation, i.e. emissions from on-road vehicle operation once the vehicle has warmed up; start operation, i.e. excess emissions due to vehicle start-up; extended idle operation, i.e. emissions due to long periods of off-network idling; and well-to-pump, i.e. emission produced during the process of getting fuel from raw material to the pump. In MOVES2004, all emission processes employ all pollutants, except extended idling, which applies only to energy consumption.

Pollutant and process combinations vary significantly in terms of data availability, which dictate the level of sophistication for which different combinations could be modeled. Specifically, running energy consumption rates were developed with a great level of detail, broken down by speed and power-based operating modes based on analysis of millions of seconds worth of in-use data covering all vehicle types. Conversely, CH₄ and particularly N₂O were based on relatively small samples of laboratory bag (e.g. standard Federal Test Procedure) data, and were not split into operating modes. Start energy consumption and emission rates were based largely on FTP bag data, and hence are expressed as aggregate per-start quantities without further breakdown. Very limited data are available on extended idle energy consumption rates for heavy-duty trucks, hence these rates were developed as adjustments to in-traffic idle rates based on a recent EPA study.

Cold start temperature effects and air conditioning usage were modeled as adjustment factors in MOVES2004, and are documented in this report.

3. Overview of Data Used in MOVES2004

The MOVES model was designed to generate emission rates primarily from second-by-second emissions data. This represents a large shift from MOBILE, which generated emission rates from Federal Test Procedure (FTP) bag emission, with additional corrections to account for “off-cycle” driving (although this shift was only employed for running energy consumption in MOVES2004, while start energy consumption, CH₄ and N₂O rates were developed from traditional bag data). Using second-by-second data

^b Also presented in this report are inputs for the calculation of Atmospheric CO₂ and CO₂ equivalent. However, these emissions are not included in MOVES2004 at this time. Atmospheric CO₂ includes CO₂ emitted directly from the tailpipe, plus HC and CO emitted from the tailpipe which is oxidized into CO₂ in the atmosphere.

allows a much broader range of data to be used in developing emission rates, but also opens up a host of issues regarding data quality assurance.

No new testing was performed specifically to provide data for MOVES2004. Instead, data were compiled from previous EPA test programs and from several external sources, quality assured and corrected, and (if not already in the database) entered into EPA's Mobile Source Observation Database (MSOD).¹ EPA contracted with Eastern Research Group Inc., (ERG) to assist in the acquisition, quality checking and compilation of data collected by outside parties; the reports resulting from this work contain the specifics of the test programs collected, and should be consulted for additional detail.^{2,3,4} Data checking and loading software was developed by the contractor Britton Information Systems.^{5,6} This data acquisition process was carried out over the course of a year and resulted in changes in EPA software, hardware, and processes for obtaining, documenting, and reformatting emissions data from non-agency data sources. This section discusses the processes used to obtain, check, and distribute the data to the MOVES modeling team followed with a description of the data sets used for MOVES.

3.1 Data Gathering and Processing Steps

The overall data acquisition and management process for all MOVES data followed these steps (some steps were not necessary for data already in the possession of EPA, or if the data were bagged emissions rather than second-by-second):

1. Identify candidate emission test programs based on literature search and contact with non-EPA parties conducting emission research.
2. Determine suitability of emissions data and vehicle information data from these test programs according to quality ranking guidelines for data and test program documentation
3. Obtain the data from the data source if not currently with EPA.
4. Reformat the data for processing by EPA quality checking software
5. Check the data for errors and reasonableness (e.g. out-of-range, data discontinuity, and analyzer "freezing").
6. Load the data into EPA's Mobile Source Observation Database (MSOD).
7. Check and (if needed) shift time alignment for second-by-second data based on pre-determined algorithm correlating CO₂ emissions and vehicle power.
8. Deliver aligned emissions data and vehicle information in a suitable format to the MOVES modeling team for processing.

The process on a dataset-by-dataset basis came to a halt for a variety of reasons. Many data sets were incomplete in content and description. They could not physically be loaded into MSOD. EPA and ERG had to contact the organization that originated the data many times to make amendments or corrections. This rework was significant and necessary to make the data useable. In the end some of the supplied data was not used.

3.2 Data Sources

The measurement and collection of second-by-second vehicle emissions data is not a new practice. It has been done for the certification of diesel engine vehicles for total hydrocarbon and oxides of nitrogen (NO_x) for a decade. The practice is done for all pollutants as a tool in the development of emissions control systems by manufacturers for years and done in volume daily by emission and inspection programs across the United States and Canada. However, in-use emission assessment programs are relatively new to the practice.

EPA collected a great deal of second-by-second emission data in the development of the I/M240 test procedure in the early 1990's and began to collect it on a routine basis across all emissions factor test programs in 1997; in particular, a large effort was made to collect second-by-second data on many vehicles over many inventory drive schedules to support the development of MOBILE6, although the second-by-second data wasn't used in MOBILE6 itself. These data are stored in MSOD and was available to MOVES immediately.

The Agency's in-use emissions test program during that period was small in comparison to past efforts, however, and the volume of data on-hand was considered insufficient for MOVES. To supplement the MOVES2004 dataset, a list of potential data sources was compiled by MOVES team members and presented to ERG, who was given the task of contacting the principle investigators and determining the suitability of the data based on quality ranking criteria developed by EPA. ERG then obtained documentation on the test programs the data represented.

Although second-by-second data was the focus of the data gathering effort, bag emission data was also sought where available, particularly in some cases where we knew second-by-second data wouldn't be used – primarily for N₂O emissions.

Sources for test data outside of EPA were:

- **California Air Resources Board (CARB)**
 - Development of Unified Correction cycles (UCC) in 1996
 - 16th Vehicle Surveillance Program with N₂O bag data
- **Coordinating Research Council (CRC)**
 - Study in 1997 to determine the effects of sulfur levels in fuel on vehicles (E-42 & -47)
 - Auto/Oil Air Quality Improvement Research Program (early 1990's)

- Study in 2002 on Heavy -duty Vehicle Chassis Dynamometer Testing for Emissions Inventory (E-55)
- **The New York State Instrumentation/Protocol Assessment Study** which compared the standard IM240 test procedure and instrumentation with the New York Transient Emissions Short Test (NYTEST)
- **North Carolina State University** study to determine the emissions savings that could be achieved through better traffic management
- **University of California Riverside College of Engineering Center for Environmental Research and Technology (CE-CERT)**
 - NCHRP 25-11 Comprehensive Modal Emission Model and Vehicle Emissions Database, Version 2.02
 - Ammonia from Light-duty vehicles
 - Heavy-Duty Diesel Truck study
- **Environment Canada's** study on the Effects of Aged Catalysts and Cold Ambient Temperatures on Nitrous Oxide Emissions
- **Texas Department of Transportation and the University of Texas** study in 2002-03 on the use of new fuels in heavy -duty diesel vehicles
- **West Virginia University** testing of heavy duty vehicles using their portable dynamometer (four distinct datasets were obtained)

The above-referenced ERG reports cover the efforts to obtain, format, and deliver the data to EPA. These reports cover the number of tests, the number and vehicle types tested, test procedures followed, test conditions, and test fuels. The reports indicated the limitations of the test data and quality of the documentation for the test programs that the data represented.

In addition to the programs listed above, this effort also gathered second-by-second data on tens of thousands of vehicles from Inspection / Maintenance (I/M) Programs:

- Arizona I/M program (Car Care). Data from January 1, 2002 to June 30, 2002
- British Columbia I/M program (Air Care). Data from January 1, 2001 to June 3, 2002
- Colorado (I/M) program (Air Care). Data from January 1, 1999, to September 1, 2002

These data were not used in MOVES2004, however, since I/M program data introduces another layer of data concerns which couldn't be adequately addressed in time for inclusion in the model. We will be considering the use of these data, along with additional I/M programs and remote sensing device (RSD) data, for the criteria pollutant version of MOVES.

All data used in MOVES2004, and/or available in the MSOD, can be obtained via request to EPA. Interested parties may request the data via email to the mobile@epa.gov inbox.

4. Total Energy Rates

4.1 Source Bin Definitions

Source bins are groupings of parameters which distinguish differences in energy and emission rates according to physical differences in the source, e.g. weight, engine size, model year group etc. In MOVES the entire on-road fleet is characterized by source bins, which are mapped to the activity-based Source Use Types within the model. In the MOVES design, source bin definitions are able to vary by pollutant, since in the real world the vehicle characteristics which are most influential to energy consumption and emissions do vary depending on the pollutant. The reader should consult the Design Manual for additional background on these concepts.

Source bins fields and the definitions within each field were based on quantitative analysis of available data to determine the most important variables to consider, and important breakpoints to account for within each field; and qualitative assessment of important categories to define for full coverage of the current and future fleet, or consistency with existing methodologies.

Source bin fields for energy consumption were developed based on a quantitative assessment of the most important factors influencing CO₂ emissions (as a surrogate for fuel and energy consumption) performed by ERG.⁷ This work analyzed several thousand chassis dynamometer emission tests on light-duty gasoline and diesel vehicles and several hundred chassis dynamometer emission tests on heavy-duty diesel vehicles. A ranking of the most important factors to consider in modeling CO₂ and CH₄ emissions was developed. The analysis was performed using ANOVA, with the pollutant as the dependent variable and numerous vehicle, fuel, and operating parameters as the independent variables. The analysis examined both the partial r^2 (the contribution of a particular factor in explaining overall variability) and the relative size of the parameter estimate of a particular factor. Based on this analysis, ERG ranked the factors as “very important,” “somewhat important,” “not important,” and “not significant.”

For CO₂, the conclusion we derived from the ERG analysis is that power (for the ERG analysis characterized by average total power, since emission results were only available at the bag level) is by far the most critical factor for both the light-duty gasoline and heavy-duty diesel vehicles, followed by which driving cycle the vehicle was run on. The most important vehicle characteristic was the ratio of vehicle displacement to weight, which was a surrogate for vehicle power/weight ratio. Many variables one would assume are influential turned out not to be, such as start versus running or vehicle mileage. One anomaly is the importance of air conditioning operation, which was classified as “not important” in the ERG analysis. This result was likely driven by the

fact only a small portion of the dataset was collected when the air conditioning was on. Another study, which focused specifically on this subset of data, found significant differences in fuel consumption when the air conditioning was engaged.⁸

The ERG analysis led us to conclude that it was important to define source bins by vehicle (loaded) weight and engine displacement (the most important parameters of average total power and driving cycle are accounted for with the use of operating mode bins described in Section 4.2). We further expanded source bins to make sure that MOVES had the flexibility to model different mixes of vehicle technologies – adding fuel type, engine technology and model year group. The source bin fields and the bin definitions are shown in Table 4-1.

Table 4-1 – MOVES Source Bin Definitions for Total Energy

Fuel Type	Engine Technology	Model Year Group	Loaded Weight	Engine Size
Gas	Conventional IC (CIC)	1980 and earlier	Null	Null
Diesel	Advanced IC (AIC)	1981-85	<= 2000 lbs	< 2.0 liters
CNG	Moderate Hybrid - CIC	1986-90	2001-2500	2.1-2.5 liters
LPG	Full Hybrid - CIC	1991-2000	2501-3000	2.6-3.0 liters
Ethanol (E85)	Moderate Hybrid - AIC	2001-2010	3001-3500	3.1-3.5 liters
Methanol (E85)	Full Hybrid - AIC	2011-2020	3501-4000	3.6-4.0 liters
Gas H ₂	Fuel Cell	2021 and later	4001-4500	4.1-5.0 liters
Liquid H ₂	Hybrid - Fuel Cell		4501-5000	> 5.0 liters
Electric	(See Table 4-14 for combinations of fuel type and engine technology used in MOVES2004)		5001-6000	
			6001-7000	
			7001-8000	
			8001-9000	
			9001-10,000	
			10,001-14,000	
			14,001-16,000	
			16,001-19,500	
			19,501-26,000	
			26,001-33,000	
			33,001-40,000	
			40,001-50,000	
			50,001-60,000	
			60,001-80,000	
			80,001-100,000	
			100,001-130,000	
			>=130,001	

A source bin is a unique combination of values across each category. For example, all gasoline vehicles, conventional internal combustion engines, model year 1991 – 2000, loaded weight 2501 – 3000 lbs, engine displacement <2.0 liters would define a single source bin with a unique set of emission rates. It is important to note that, for total energy, traditional definitions of vehicle regulatory class (e.g. LDV, LDT1, LDT2 etc.) are not part of the source bin definition for total energy. While vehicle class inherently influences tailpipe emission results, energy consumption is primarily a function of fuel, technology, vehicle weight, engine size, and model year regardless of whether it is a car or a truck.

To keep track of the myriad of source bins in the EmissionRate data table, a 19-digit Source Bin ID schema has been developed. While seemingly Byzantine at first, we're confident that MOVES users will learn to recognize Source Bin IDs as quickly as

their own phone numbers. Table 4-2 shows the source bin schema, which differs for total energy, CH₄ and N₂O.

Table 4-2: Source Bin ID Schema

Digit	Field
1	Leading digit
2-3	Fuel type
4-5	Engine Technology
6-7	Regulatory Class (CH ₄ , N ₂ O only)
8-9	Model Year Group
10-13	Engine Size
14-17	Loaded Weight
18-19	Trailing zeros

The key values (IDs) for each field can be located in the Design Manual, or by directly querying the FuelType, EngineTech, ModelYearGroup, EngineSize and WeightClass tables in the MOVES default database.

4.2 Operating Mode Definitions

We subdivide total activity into categories that differentiate emissions, known as operating mode bins. Operating mode bin definitions are allowed to vary by pollutant and emission process, to accommodate variability in the important activity-based effects on energy and emissions. 17 operating mode bins have been defined for running total energy; braking, idle, and 15 subdivisions of cruise and acceleration defined by instantaneous vehicle speed and vehicle specific power (VSP). A detailed account of how these bins were established, plus proof-of-concept validation analysis, is found in Appendix A. The specific bin definitions for total energy are shown in Table 4-3.

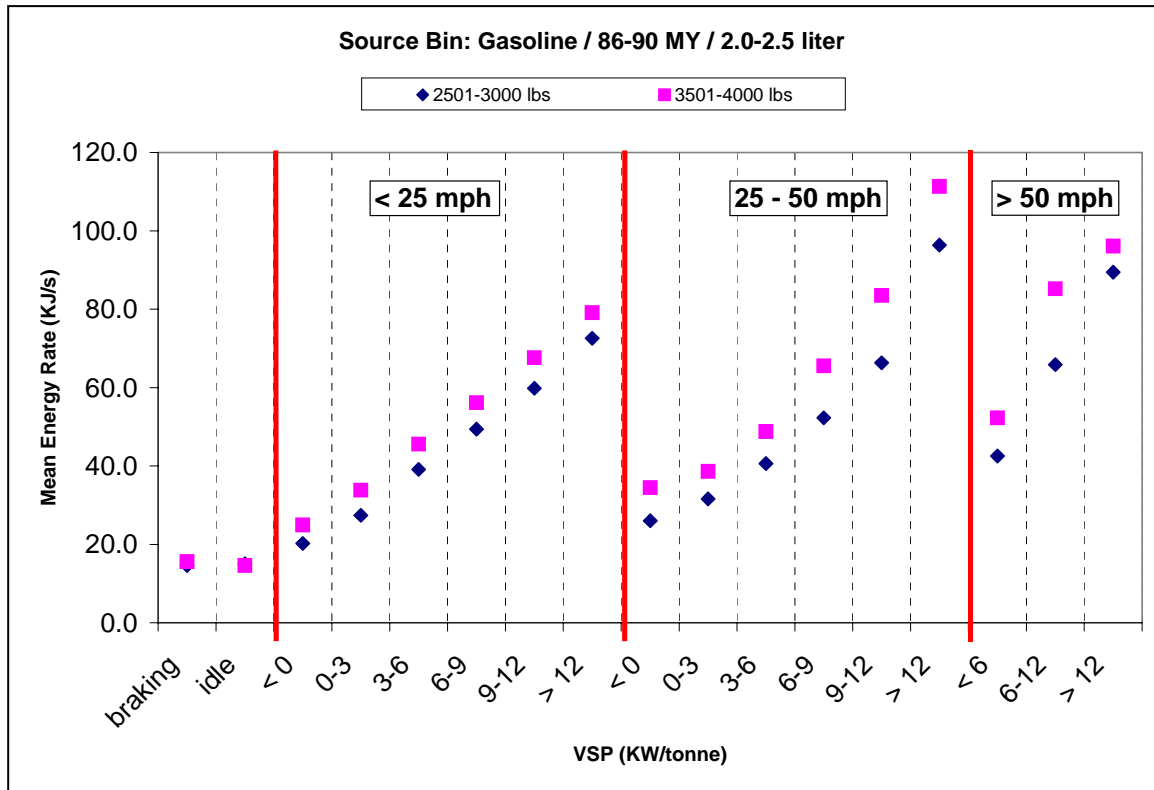
Table 4-3 Operating Mode Definitions for Running Total Energy

Braking (Bin 0)			
Idle (Bin 1)			
VSP \ Speed	0-25mph	25-50	>50
< 0 kw/tonne	Bin 11	Bin 21	
0 to 3	Bin 12	Bin 22	
3 to 6	Bin 13	Bin 23	
6 to 9	Bin 14	Bin 24	
9 to 12	Bin 15	Bin 25	
12 and greater	Bin 16	Bin 26	Bin 36
6 to 12			Bin 35
< 6			Bin 33

The Bin ID numbers shown in the tables are the key field IDs used in the EmissionRate table to identify the bins. Emission rates in the EmissionRate table are keyed by a) source bin, b) pollutant / process, and c) operating mode bin.

Figure 4-1 provides a graphical example of energy rates by operating mode bin for two different source bins differentiated only by weight class: conventional gasoline 1986-1990 model year vehicles with engine size in 2.0 – 2.5 liter range, with a loaded weight of 2501 – 3000 pounds, and 3501 – 4000 pounds.

Figure 4-1: Energy Rates by Operating Mode for Two Source Bins



4.3 Running Energy Rate Development

4.3.1 Data Used

4.3.1.1 EPA Test Programs

Data used to generate total energy rates were from the EPA and non-EPA test programs discussed in Section 3. All data used for generating total energy rates were second-by-second (1 Hz) resolution data. The EPA data sources used for total energy rates in MOVES are found summarized in Table 4-4. Many of the driving cycles listed under “test cycles” are the facility-based cycles developed for use in MOBILE6 (e.g. FWY or ART, which stand for freeway or arterial, along with a Level of Service designation); more detail on these cycles can be found in MOBILE6 reports documenting

their development.⁹ Under the “seconds of data” heading, “raw” is the number of seconds of data from each test program in the MSOD, before any processing or filtering. “After QA/QC” is what remained after the process discussed in Section 3.1; this was the sample passed to the binning program discussed in the next section. “After filtering” is what remained after the filtering performed by the binner program, discussed in Appendix B; this is the sample used directly in developing the total energy rates.

Table 4-4: EPA-sponsored test programs supplying data to MOVES2004

MSOD Program ID	Description	Vehicles Used	MY Range	Test Cycles	Seconds of Data		
					Raw	After QA/QC	After filtering
98N2OA	Nitrous Oxide (N2O) Study on Tier 1 LDVs, LDTs, and LEVs (LDVs) at various mileages.	23	1995-99	F505 FTP FWY-AC US06	171,240	170,579	77,255
CDHOT_PM_A	Determining the Relationship of Opacity and Exhaust Emissions (Including Total PM) in In-use Gasoline Powered Vehicles During an IM240	107	1969-98	3IM240	74,673	74,488	74,176
CYCLES_A	Determining Basic Exhaust Emission Rates for Light Duty Cars and Trucks using Multiple Drive Schedules and with the Air Conditioning On and Off	43	1990-96	ART-AB ART-CD ART-EF FWY-AC FWY-D FWY-E FWY-F FWY-G FWY-HI LA4 LA92 LOCAL NONFRW NYCC RAMP ST01 US06	610,671	606,385	349,767
LDV_AC_A	Determining Basic Exhaust Emission Rates for Light Duty Cars and Trucks using Multiple Drive Schedules	62	1983-96	ART-AB ART-CD ART-EF FWY-AC FWY-D FWY-E FWY-F FWY-G FWY-HI LA92 LOCAL NONFRW NYCC RAMP ST01	671,704	670,367	552,492
LHDDT_A	Investigation on diesel LHDT exhaust emissions on various cycles, play-loads including measuring for toxics, PM and unregulated pollutants.	6	1998-99	ART-CD FTP FWY-E FWY-G FWY-HI LA92 NYCC ST01 US06	74,902	74,828	73,304
LHDT_A	Investigation of gasoline 10 LHDTs on 8 driving cycles at different loads, start emissions, IMs, FTPs and different fuel sulfur levels	10	1991-97	ART-EF F505 FTP FWY-D FWY-F	399,258	397,611	387,097

				FWY-G FWY-HI IM240 LA92 NYCC RAMP ST01 US06			
LHDT_B	Investigation on gasoline LHDT exhaust emission on various driving cycles, payloads, and fuels (sulfur)	2	1993-95	ART-CD FTP FWY-D FWY-E FWY-F FWY-G FWY-HI IM240 LA92 NYCC ST01 US06	73,489	73,196	71,814
LHDT_C	Determining Basic Exhaust Emission Rates for Light Heavy Duty Trucks using Multiple Drive Schedules and Payloads	18	1989-97	ART-CD F505 FTP FWY-D FWY-E FWY-F FWY-G FWY-HI LA92 NYCC ST01 US06	365,572	365,164	357,842
LHDT_EVAP	LHDTs with evaporative emission tests at 3 different temp. ranges and two different fuels (6.3 and 9.0) including FTPs. Gas cap on/off followed .	2	1993-97	F505 FTP IM240 US06	7,949	7,877	7,647
LHDT_LDT	Inventory Cycles/LA92 Exhaust Emissions Data Collection and Amendment #1 Sulfur Fuel Testing	46	1988-97	ART-EF F505 FTP FWY-D FWY-F FWY-G FWY-HI IM240 LA92 NYCC RAMP ST01 US06	792,700	791,659	773,893
OBD_A	Determining the Effectiveness of Onboard Diagnostic Systems in Identifying Vehicles that Fail the FTP	47	1996-98	FTP IM240	105,573	105,416	100,852
OEM_2100	An Investigation of OEM 2100's Capabilities to Accurately Measure Emissions of Late-Model Gasoline Vehicles	2	1996-98	FTP FWY-HI NYCC US06	15,962	15,950	15,512
ROVER_A	Determining the Viability of Gathering with ROVER Exhaust and Vehicle Information for Diesel and Gasoline Powered Light-Heavy Duty Trucks	8	1998-99	FTP FWY-HI NYCC US06	59,031	59,016	57,272
SHOOT_OUTA	On-Road Emission Test Data from 15 Light-Duty Vehicles and 15 Heavy Duty Diesel Trucks for On-Board Emission Data Analysis and Collection for the New Generation Model	35	1992-00	FTP LA4 US06	40,509	40,479	38,716
TIER_1	Determining Basic Exhaust Emission Rates for Tier 1 Light Duty Cars using Multiple Drive Schedules	38	1993-97	ART-EF F505 FTP FWY-F FWY-HI IM240	289,768	287,161	278,764

				LA92 NYCC ST01 US06			
TIER_1_B	Investigation on gasoline vehicles/trucks/SUVs over different driving cycles including SFTP	15	1996-00	FTP FWY-HI LA92 NYCC SC03 ST01 US06	93,585	93,359	90,159

The EPA programs were performed in the Ann Arbor, Michigan EPA laboratory or at one of two contractor labs: Automotive Testing Laboratory or Southwest Research Institute. The vast majority of vehicles in the program are in-use vehicles recruited from the public. Many of the LHDTs and both vehicles in the ‘OEM_2100’ test program were leases from vehicle dealers or vehicle rental agencies. Several programs were assessments of new methodologies to measure emissions with on-board instrumentation (“ROVER_A”, “OEM_2100”, and “SHOOT_OUTA”) – although the data used in MOVES was from the laboratory measurements from those programs. Two of the programs were focused on the characterization of two emission components of interest in SI vehicles N₂O (“98N2O_A”) and particulate matter (‘CDHOT_PM_A’) where the principal and typical emission components were collected. All the tests followed typical EPA test documentation procedure, resulting in complete and extensive documentation.

The EPA programs were not subject to the ERG data acquisition process, but followed the data checking, loading, alignment, and delivery parts of the MOVES data process. An EPA test’s second by second data was excluded from the MOVES modeling team delivery if any of the aggregated second by second constituents failed to be within 10 percent of the bag analysis values. The various programs reported as being delivered to MOVES have statements of work describing how EPA thinks the work should be done, a work plan delivered by testing contractor that describes how the work will be done, and a final report describing what work was done. These reports are all available in electronic form from the Data Acquisition and Management Team of the Assessment and Standards Division, and can be obtained upon request to mobile@epa.gov.

4.3.1.2 Non-EPA Test Programs

There were sixteen groups of data either delivered to EPA from ERG, or already in the MSOD, that were originally targeted for MOVES. Thirteen were emission characterization programs and the three were emission tests from state or regional inspection and maintenance programs (these data were not used in MOVES2004). In all there were 10,760 vehicles and 35,489 tests in the characterization programs, of which 9,161 vehicles and 34,901 tests loaded in MSOD; the excluded characterization data was not considered to be of sufficient quality to be useful.

A summary of non-EPA programs used in MOVES2004 are shown in Table 4-5.

Table 4-5: Test programs supplying data to MOVES2004 not conducted by EPA

MSOD Program ID	Description	No. of Vehicles Used	MY Range	Test Cycles	Seconds of Data		
					Raw	After QA/QC	After filtering
CARB_UCC96	California Air Resources Board (CARB) development of Unified Correction Cycles (UCC) in 1996	42	1973-94	FTP LA92 UCC15 UCC20 UCC25 UCC30 UCC35 UCC40 UCC45 UCC50	496,827	495,825	487,112
CECERT_NH3	University of California Riverside College of Engineering Center for Environmental Research and Technology (CE_CERT) Emissions of Ammonia for Light-duty vehicles	35	1983-01	FTP FWY NYCC US06	82,262	82,262	78,278
CRC_E55_59	Coordinating Research Council (CRC) study in 2002 on Heavy-duty Vehicle Chassis Dynamometer Testing for Emissions Inventory.	25	1973-00	AC5080 CARB-C CARB-I CARB-R CARB-T CARBCL	631,380	473,589	472,155
CRC_S_LDV1	Coordinating Research Council (CRC) study in 1997 to determine the effects of sulfur levels in fuel on vehicles.	12	1997	FTP	618,567	616,719	582,122
GRANT97_NY	New York/DEC-Characterization and Control of HD Diesel Vehicle Emissions in the New York Metropolitan Area	35	1966-99	5PEAK UDDS_D WVUCBD	257,101	257,101	256,303
NCHRP	University of California Riverside College of Engineering Center for Environmental Research and Technology (CE_CERT) NCHRP 25-11 Comprehensive Modal Emissions Model and Vehicle Emissions Database, Version 2.02	337	1965-99	FTP MEC5 MEC6 MEC7 SMEC6 SMEC7 US06	1,379,200	1,378,101	720,911
NYIPA/ NYIPA2002	Second Iteration of the New York State Instrumentation/Protocol Assessment Study which compares the standard IM240 test procedure and instrumentation with the New York Transient Emissions Short Test (NYTEST).	9900	1992-01	IM240	5,945,603	5,902,463	5,720,688
WVU1-4	West Virginia University testing of heavy duty vehicles using their portable dynamometer.	149	1991-99	2-5MIL 2CSHVR 2TESTD 3CBD 5MILE CBD CSHVR DRT HVDUTY KERN NYBUS NYCCT NYGT2 RT22 RT77 TEST_D UDDS_W VFAC WHM	1,634,576	1,558,382	1,554,281

The “MEC” and “SMEC” cycles run as part of the NCHRP program were ultimately removed from the data used for MOVES2004. They are cycles engineered for populating UC Riverside’s Comprehensive Modal Emission Model (CMEM). These cycle contain a high proportion of very aggressive driving (e.g. wide open throttle) which

skewed the bin results. A similar argument could be made for US06, but the same analysis determined that the US06 was not as extreme an outlier as the MEC and SMEC cycles – and the US06 is based on real in-use driving. This analysis is documented in a presentation made at the 2004 CRC On-Road Emission Workshop.¹⁰

4.3.2 Generating Rates from Available Data

Total Energy Rates were developed by binning second-by-second data from the test programs described in Section 4.3.1 for each intersection of source bin and operating mode bin (termed “cells”). “Binning” means computing the average energy rate (energy use per time – e.g. KJ per hour) for each cell across all the seconds of data falling within that cell.

To do this in an automated way, a MySQL script termed the “binner program” was written that performed the following steps, starting with an input of second-by-second data compiled from the data sources listed in Section 4.3.1: 1) calculate vehicle specific power (VSP) for each second of data; 2) determine the source bin and operating mode “cell” for each second of data; 3) calculated second-by-second energy consumption from HC, CO and CO₂ mass emissions based on carbon balance and energy content calculations, and 4) calculate, for each cell, the average energy rate, coefficient of variation, and a number of diagnostic statistics. The output of the binner program was a MySQL table containing energy rates (in KJ per hour) and coefficient of variation (CV) by source bin and operating mode bin, in the same format of the MOVES emission rate database table. Detailed documentation on the binner program are in Appendix B, including the algorithms used for performing these three steps.

4.3.3 Hole Filling

The binner program directly populated a significant number of source bin and operating mode cells with data. However, a number of cells were not populated because either no data existed to populate them with, or the data which did exist failed the data quality objectives set out in the binner process - i.e. a cell must have data from at least 3 tests and have a coefficient of variation (CV) of less than 0.50. The cells left unpopulated are dubbed “holes” which required supplemental methods to populate.

The two methods used for hole-filling were 1) employing the Physical Emission Rate Estimator (PERE) to fill holes directly, and 2) interpolation or copying of neighboring cells populated with data. These methods were chosen based on a proof-of-concept evaluation which showed the efficacy of these methods (Appendix C). As discussed in Appendix C, an additional method initially evaluated for hole-filling but not ultimately used for MOVES2004 was to derive modal rates from aggregate bag data.

After the initial binner run, an analysis was performed which identified the fraction of the fleet covered by the binned data, and the relative importance of the remaining holes. Table 4-6 shows the estimated fraction of the fleet covered by just the binned second-by-second data, by source use type.

Table 4-6: Fraction of 1999 on-road fleet covered by MSOD data

Source Use Type	Fraction Covered
Motorcycle	0
Passenger Car	0.98
Passenger Truck	0.93
Light Commercial Truck	0.87
Intercity Bus	1.0
Transit Bus	0.99
School Bus	0.84
Refuse Truck	0.86
Single-Unit Short Haul	0.65
Single-Unit Long Haul	0.65
Motorhome	0.58
Combination Short-Haul	0.36
Combination Long-Haul	0.24

As shown, the light-duty source types are covered fairly well (over 90 percent) by the existing data, with coverage dropping off for the heavier-duty classes – single-unit and particularly combination trucks, where for long-haul trucks about a quarter of the fleet is covered.

From this analysis a list of source bins holes was created, ranked according to their importance in terms of the relative fraction of the fleet – in essence a ranking of the most important holes to fill. The highest priority holes were defined as those which a) were needed to bring the fleet coverage up to 95 percent within each source use type, and b) represented at least 2 percent of the fleet within a source use type. These priority holes were generally filled by PERE, with remaining holes left for interpolation. As expected from the results in Table 4-6, the most important holes to fill were the heaviest weight bins making up the combination trucks. In particular, source bins representing loaded weights of 80,000 pounds and higher are a significant portion of heavy-duty activity (according to the U.S. Census Bureau’s Vehicle Inventory and Use Survey, or VIUS, as discussed in the Fleet and Activity report); but because available heavy-truck emission data was limited to approximately 60,000 lbs, no in-use data were available for the heaviest trucks.

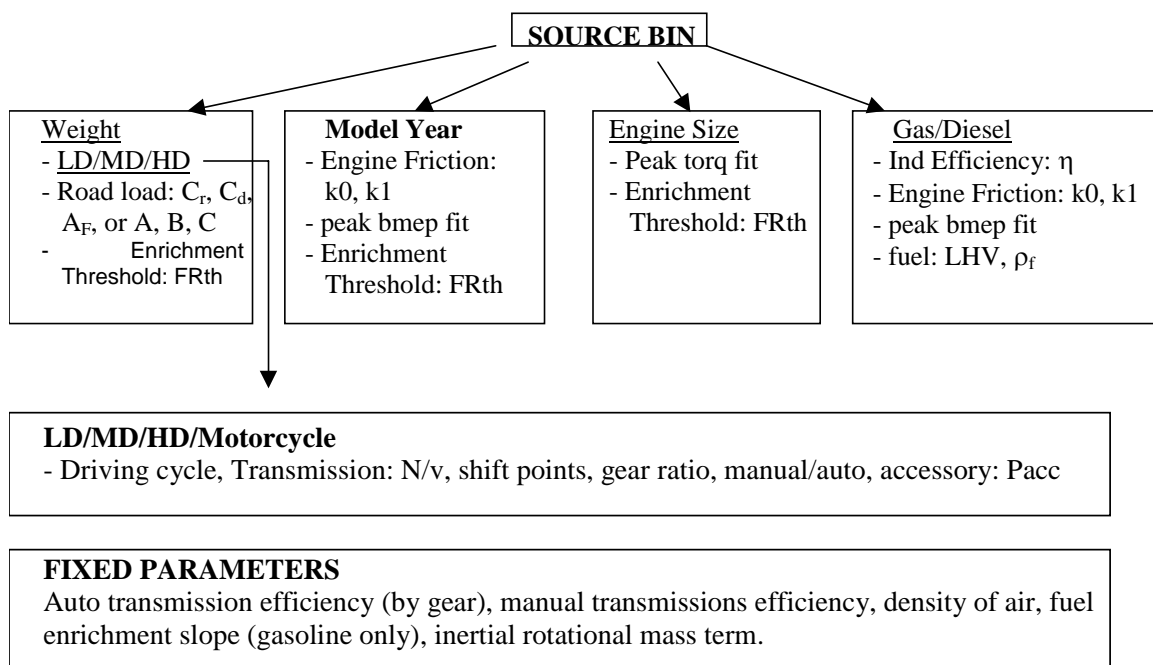
4.3.3.1 Hole Filling with PERE

The Physical Emission Rate Estimator (PERE) is a stand-alone spreadsheet model developed to fill data gaps in MOVES and to help it extrapolate to future projections of energy and emissions. The details of the PERE model are in a separate document.¹¹ Its inputs are vehicle parameters and second-by-second driving traces, and it outputs second-by-second fuel consumption rates for the running operating (e.g. it does not yet include starts). PERE uses physical principles to model propulsion systems in the vehicle. The model is based on a simple model for the internal combustion engine. Simulation of

hybridization is achieved by inserting a secondary power source and energy storage device (usually a battery/motor combination).

Aside from advanced technologies, PERE was used to fill several holes in the current fleet for MOVES2004, including most of the top-priority holes where test data was lacking or deemed insufficient. To fill a source bin hole, PERE was run once for each source bin over a series of representative driving schedules. This required choosing vehicle parameter inputs which represented an entire source bin. The parameter selection for PERE is a somewhat involved process. The decision tree is shown in Figure 4-2. The parent variables (underlined) are above the dependent variables. The parent variables define the decision source of the parameters. The dependent variables are the parameters that are determined from the parent variable. Some dependent parameters vary with multiple parents (e.g. engine friction). Some other parameters are fixed.

Figure 4-2: Decision Tree for Determine PERE Inputs for a Source Bin



The determination of general vehicle weight classification is the very first step in the parsing process. Since MOVES does not have explicit splits for light, medium and heavy duty, PERE assumes that the light to medium split occurs at weighclassID 80 (7,000-8,000 lbs), where above this weight bin is medium duty (single unit delivery trucks, buses, etc). The medium to heavy duty split occurs at weighclassID 400 (33,000-40,000 lbs). This in turn defines whether the transmission will be a 5-speed automatic (LD), 6-speed automatic (MD), 12-speed manual (HD), or 5-speed manual (motorcycle). Motorcycles have their own weight and engine size categories as well as a different accessory loading (0.25 vs 0.75kW).

The weight determines the road-load coefficients for the vehicles directly. For model years 2000 to present, A, B, and C track coefficients are provided by

manufacturers. Older vehicles only have the single Tractive Road Load Horsepower (TRLHP) figure. The TRLHP for a typical light-duty vehicle (and truck) in the source bin can be approximated from weight using the relation used in the binner program, detailed in Appendix B.

The weight class also defines the driving cycles that PERE will run. Before model year 1999, weight also is a term in the enrichment equation. After 1999, the vehicles are assumed to go into enrichment so rarely that it is ignored by PERE.

The model year also affects the performance of the engine, both in terms of friction and peak bmep (which mainly influences power downshifting). The engine/fuel type determines the efficiency and friction characteristics of the engine, and it also defines the fuel parameters since gasoline and diesel have different physical properties.

Table 4-7 show an example source bin filled by PERE

Table 4-7: Example Source Bin Filled by PERE

SourceBinID: 1020100045099080000				
Digit	Field	Code	Meaning	PERE input
1	Leading digit	1	-	
2-3	Fuel type	02	Diesel	
4-5	Engine Technology	01	Conventional Internal Combustion	
6-7	Regulatory Class	00	- (not used for energy)	
8-9	Model Year Group	04	1991-1999	1995
10-13	Engine Size	5099	> 5.0 Liters	6.9 Liters
14-17	Loaded Weight	0800	60,000 – 80,000 lbs	70,000 lbs
18-19	Trailing zeros	00	-	

Using the information above and the descriptions in the PERE documentation, the PERE input selected to fill this source bin are shown in Table 4-8.

Table 4-8: PERE Inputs For Filling Source Bin 1020100045099080000

Vehicle	
Model Year	1995
Vehicle wgt (kg)	31752
Cr0 (rolling resistance)	N/A
Cd (drag coeff)	N/A
A (frontal area m^2)	N/A
A (N)	2098.81
B (N/mps)	0.000
C (N/mps^2)	4.2268
Pacc (accessory - kW)	0.75
Engine	
Engine Displ (L)	6.9
k0 (N indep friction kJ/Lrev)	0.0605
k1 (N dependent fric)	0.00333
P/T indicated eff (eta)	0.48
Transmission	
N/v (rpm/mpg)	26.7
Nidle (rpm)	700
trans eff	0.95
Shift point 1-2 (mph)	2.48
Shift point 2-3	4.75
Shift point 3-4	7.75
Shift point 4-5	13.5
Shift point 5-6	17.5
Shift point 6-7	23
Shift point 7-8	34
Shift point 8-9	50
Shift point 9-10	56
Shift point 10-11	57
Shift point 11-12	64
Shift point 12-13	64
g/gtop 1	13
g/gtop 2	9.4
g/gtop 3	6.9
g/gtop 4	5
g/gtop 5	3.7
g/gtop 6	2.7
g/gtop 7	2
g/gtop 8	1.5
g/gtop 9	1.2
g/gtop 10	1.1
g/gtop 11	0.93
g/gtop 12	0.86
g/gtop 13	0.86
Fuel	
LHV (kJ/g)	43.2
density gas (kg/L)	0.8114

Once the vehicle inputs were determined, PERE was run over a series of driving schedules (cycles) to generate second-by-second energy consumption. Binned energy consumption rates for MOVES were generated from the second-by-second PERE results using the same process the binner program used with test data, as described in Section 4.3.2.

The binned energy consumption and emissions rates are dependent on the driving cycle input. For PERE, the driving cycle is (in turn) dependent on the weight of the vehicle. It is important to capture a representative sampling of real-world driving. Fourteen driving cycles for light duty applications have been selected for this purpose from MOVES, shown in Table 4-9. These are a subset of the schedules used by MOVES to generate operating mode distributions for the calculation of total energy consumption (the cycles themselves can be queried in the DriveScheduleSecond table of the MOVES default database). These cycles represent a broad spectrum of driving, from very low to very high speeds. When merged, the cycles run 6,981 seconds. In the future it is likely that similar binned rates could be obtained from a smaller sample of driving cycles, thus making the execution of PERE easier to manage.

Table 4-9: Light-duty driving cycles used as input to PERE

MOVES ScheduleID	Cycle	Avg Spd (mph)
156	FWYHI1	63.2
155	FWYAC	59.7
154	FWYD	52.9
153	FWYE	30.5
152	FWYF	18.6
151	FWYG	13.1
199	RAMP	N/A
105	ARTAB	24.8
104	ARTCD	19.2
103	ARTEF	11.6
102	NYCC	7.1
157	FWYHI2	68.2
158	FWYHI3	76.0
101	LOWSPEED1	2.5

For medium and heavy-duty applications, the driving cycles were drawn from the medium and heavy-duty schedules used to populate the MOVES default database, described in the Fleet and Activity report (and also contained in the DriveScheduleSecond table). The medium duty drive cycles are listed in Table 4-10 and totaled 6,050 seconds.

Table 4-10: MOVES Medium duty driving cycles used as input to PERE

MOVES ScheduleID	Cycle	Avg Spd (mph)
201	MD 5mph Non-Freeway	1.8
202	MD 10mph Non-Freeway	10.5
203	MD 15mph Non-Freeway	15.6
204	MD 20mph Non-Freeway	20.4
205	MD 25mph Non-Freeway	24.4
206	MD 30mph Non-Freeway	30.8
251	MD 30mph Freeway	37.4
252	MD 40mph Freeway	45.3
253	MD 50mph Freeway	55.5
254	MD 60mph Freeway	60.1

The heavy-duty cycles (Table 4-11) are also a subset from the MOVES drive schedules. Together they total 28,313 seconds of driving, weighted mainly with freeway driving. The length of this file makes the execution of PERE cumbersome for heavy-duty, and in the future, the number of cycles will be reduced.

Table 4-11: Heavy-duty driving schedules used as input to PERE

MOVES ScheduleID	Cycle	Avg Spd (mph)
301	HD 5mph Non-Freeway	1.2
302	HD 10mph Non-Freeway	10.8
303	HD 15mph Non-Freeway	15.2
304	HD 20mph Non-Freeway	19.8
305	HD 25mph Non-Freeway	24.9
306	HD 30mph Non-Freeway	30.8
351	HD 30mph Freeway	34.9
352	HD 40mph Freeway	46.9
353	HD 50mph Freeway	54.3
354	HD 60mph Freeway	59.5

The motorcycles are run only on the FTP and HWY driving cycles; small engine scooters (under 500 pounds) were run on only the reduced speed FTP scooter cycle.

With the driving cycles entered, a binning macro (on the PERE spreadsheet) is run in order to bin the fuel consumption rates according the 17 VSP / speed bins used by MOVES, which were in turn converted to MOVES energy rates.

Tables 4-12 and 4-13 show the source bin holes filled by PERE for all vehicle classes except motorcycles. The source bins in Table 4-12 were chosen based on their relative importance in the fleet (the “rank” category signifies the rank of importance in

terms of fraction of the overall in-use fleet). Table 4-13 are less important “leftover” bins which were not filled by the interpolation / copying process described in Section 4.3.3.2, generally because there were enough neighbor bins present to satisfy the criteria for this method.

Table 4-12 Priority Holes Filled by PERE

Rank	SourceBin ID	Fuel	MY Group	Engine Size Class (L)	Loaded Weight Class (lbs)
1	1020100045099080000	Diesel	91-00	>5.0	60-80K
2	1020100045099100000	Diesel	91-00	>5.0	80-100K
3	1020100055099080000	Diesel	<90	>5.0	60-80K
4	1010100045099016000	Gas	91-00	>5.0	14-16K
5	1010100055099016000	Gas	01-10	>5.0	14-16K
6	1020100055099100000	Diesel	<90	>5.0	80-100K
7	1010100045099019500	Gas	91-00	>5.0	16-19.5K
8	1020100055099014000	Diesel	<90	>5.0	10-14K
9	1020100055099016000	Diesel	<90	>5.0	14-16K
10	1020100045099016000	Diesel	91-00	>5.0	14-16K
11	1020100045099004500	Diesel	91-00	>5.0	4-4.5K
12	1020100045099005000	Diesel	91-00	>5.0	4.5-5K
13	1020100045099006000	Diesel	91-00	>5.0	5-6K
14	1010100055099026000	Gas	01-10	>5.0	19.5-26K
15	1010100032025006000	Gas	86-90	2.0-2.5	5-6K
16	1020100043540014000	Diesel	91-00	3.5-4.0	10-14K
17	1020100055099009000	Diesel	<90	>5.0	8-9K
18	1010100045099026000	Gas	91-00	>5.0	19.5-26K
19	1020100045099003500	Diesel	91-00	>5.0	3-3.5K
20	1010100055099019500	Gas	01-10	>5.0	16-19.5K
21	1020100045099004000	Diesel	91-00	>5.0	3.5-4K
25	1010100055099033000	Gas	01-10	>5.0	26-33K
26	1020100055099008000	Diesel	<90	>5.0	7-8K
27	1020100055099007000	Diesel	<90	>5.0	6-7K
28	1020100045099130000	Diesel	91-00	>5.0	100-130K
29	1020100055099010000	Diesel	<90	>5.0	80-100K
32	1010100044050002500	Gas	91-00	4.0-5.0	2-25K
35	1010100053035014000	Gas	01-10	3.0-3.5	10-14K
36	1010100044050009000	Gas	91-00	4.0-5.0	8-9K
37	1010100044050007000	Gas	91-00	4.0-5.0	6-7K
38	1010100045099033000	Gas	91-00	>5.0	26-33K
40	1010100055099004000	Gas	01-10	>5.0	3.5-4K
42	1010100055099003500	Gas	01-10	>5.0	3-3.5K
43	1010100055099002500	Gas	01-10	>5.0	2-2.5K
44	1010100055099003000	Gas	01-10	>5.0	2.5-3K
62	1020100043540019500	Diesel	91-00	3.5-4.0	16-19.5K
65	1020100043540007000	Diesel	91-00	3.5-4.0	6-7K

Table 4-13: “Leftover” Holes Filled by PERE

SourceBinID	Fuel	MY Group	Engine Size Class (L)	Loaded Weight Class (lbs)
1010100012025003500	Gas	<80	2.0-2.5	3-3.5K
1010100013035003500	Gas	<80	3.0-3.5	3-3.5K
1010100013540003500	Gas	<80	3.5-4.0	3-3.5K
1010100013540004000	Gas	<80	3.5-4.0	3.5-4K
1010100013540004500	Gas	<80	3.5-4.0	4-4.5K
1010100014050004500	Gas	<80	4.0-5.0	4-4.5K
1010100015099006000	Gas	<80	>5.0	5-6K
1010100020020002000	Gas	81-85	<2.0	<2K
1010100022025004500	Gas	81-85	2.0-2.5	4-4.5K
1010100023035004500	Gas	81-85	3.0-3.5	4-4.5K
1010100024050003000	Gas	81-85	4.0-5.0	2.5-3K
1010100024050005000	Gas	81-85	4.0-5.0	4.5-5K
1010100024050006000	Gas	81-85	4.0-5.0	5-6K
1010100025099002500	Gas	81-85	>5.0	2-2.5K
1010100030020002000	Gas	86-90	<2.0	<2K
1010100032530002500	Gas	86-90	2.5-3.0	2-2.5K
1010100033540003000	Gas	86-90	3.5-4.0	2.5-3K
1010100034050006000	Gas	86-90	4.0-5.0	5-6K
1010100035099007000	Gas	86-90	>5.0	6-7K
1010100043035002500	Gas	91-00	3.0-3.5	2-2.5K
1010100043540002500	Gas	91-00	3.5-4.0	2-2.5K
1010100043540003500	Gas	91-00	3.5-4.0	3-3.5K
1010100044050002000	Gas	91-00	4.0-5.0	<2K
1010100045099003000	Gas	91-00	>5.0	2.5-3K
1010100995099004500	Gas	91-00	>5.0	4-4.5K
1010100045099009000	Gas	91-00	>5.0	8-9K
1010100052530004000	Gas	01-10	2.5-3.0	3.5-4K
1010100053540004000	Gas	01-10	3.5-4.0	3.5-4K
1010100984050005000	Gas	01-10	4.0-5.0	4.5-5K
1010100985099004500	Gas	01-10	>5.0	4-4.5K

In addition to the bins in Table 4-12 and 4-13, PERE was used to fill all the bins for motorcycles, since we did not have any test data to bin directly. In MOVES, motorcycles are divided into three weight categories (<500 lbs, 500-700 lbs and >700 lbs) and three engine size categories (<170 cc, 170-280 cc, and >280cc). Energy rates were calculated for each combination of engine size and weight that exists in the fleet (as captured in the MOVES default database) by running PERE using midpoint values for each category. Detailed methodology and specifications for motorcycles are described in the PERE report.

4.3.3.2 Hole Filling with Interpolation and Copying

It was not feasible to use PERE to fill all holes, so a systematic approach for interpolation or copying rates for lower priority holes based on the rates of neighboring bins was developed. Since many dimensions were available to interpolate, a hierarchical set of rules were developed to perform interpolation based on weight, engine size or model year; and barring adequate data to perform these interpolations, to copy directly from neighboring bins by model year, weight or engine size. This process was executed using a SAS script. A full discussion of this interpolation decision algorithm is contained in Appendix D.

4.3.4 Advanced Technologies & Alternative Fuels

Advanced technologies, alternative fuels and future model year vehicles are an important component of MOVES, in order to generate estimates into the future, and evaluate the effect of new technologies penetration into the fleet. Unfortunately little to no data exists advanced technology and alternative fuel vehicles (not to mention future vehicles) to allow direct binning, or even interpolation; hence a process for generating these rates was required. Advanced technologies, alternative fuels and future model year rates are not included in the emission rate table in the MOVES default database; rather, a separate database table is created by the user via a pre-processing step known as the Future Emission Rate Creator, or FERC. Details on executing the FERC can be found in the MOVES2004 User Guide; details on the design can be found in the MOVES2004 Software Design Reference Manual. This section documents the methodology used by the FERC to generate these rates, focusing on how PERE was used to generate the default technology ratios employed in the FERC.

A matrix of fuel type and engine technology combinations available for inclusion in MOVES2004 via the FERC is shown in Table 4-14, shown with checks. Hydrogen-related technologies (internal combustion, fuel cell and fuel cell hybrid) are not included in the initial release of MOVES2004 but we plan to add them to the model as soon as they are available.

Table 4-14: Advanced Technologies and Alternative Fuels in MOVES2004

	Gas	Diesel	Alt Fuels*	Electricity	Liquid H2	Gaseous H2
Conventional Internal Combustion (CIC)	✓	✓	✓			
Advanced Internal Combustion (AIC)	✓	✓				✓
Moderate Hybrid – CIC	✓	✓				
Moderate Hybrid – AIC	✓	✓				
Full Hybrid – CIC	✓	✓				
Full Hybrid – AIC	✓	✓				
Electric				✓		
Fuel Cell					✓	✓
Hybrid – Fuel Cell					✓	✓

* Alternative Fuels in MOVES are CNG, LPG, E85 and M85 - these are separate in the model, and are only combined here for presentation

The rates for advanced technologies and alternative fuels for model years 2001 through 2010 are generated in the FERC by applying ratios to the rates for “base” fuel and engine technologies for which rates already existed. This is done for all source bins in the base technology, so that advanced technologies and alternative fuels could be modeled for all classes of vehicles. For most cases, the technologies used as the base rates to which ratios were applied were gasoline conventional internal combustion. The diesel-based technologies (e.g. advanced IC and hybrids) used diesel conventional internal combustion for a base. Because the FERC generates rates for all permutations of source bin, operating mode, fuel type and engine technology, the database table produced by the FERC is about four times larger than the emission rate table included in the MOVES default database (nearly 100,000 records in the default configuration).

The source of the ratios used to generate advanced technology rates depend on technology and fuel type. PERE was used to generate ratios for gas and diesel advanced internal combustion and hybrid configurations. The PERE ratios varied by operating mode bin, which allows more flexibility than benefits expressed as aggregate ratios over standard test procedures that are common in current literature. Ratios for E85, M85, CNG and LPG were derived from the relative benefits used in GREET, and did not vary by operating mode bin. For a given fuel/technology combination, one set of ratios (by operating mode bin, if applicable) are applied to all base source bins, meaning that the relative benefit of a certain technology is assumed to be the same across all weight classes and engine sizes. This is a broad assumption that could be refined with more investigation; we did investigate applying different ratios depending on vehicle weight, but lacked enough information at this point to determine how the ratios should change by weight or other vehicle characteristics.

Model years beyond 2010 are characterized by two broad model year groups: 2011 thru 2020 and 2021 thru 2050. Rates for these model year groups were generated by a direct ratio to the 2001 thru 2010 rates for the same technology and fuel type. This provides some opportunity to model evolutionary improvements in fuel economy and emissions technology into the future. For the default case, however, a ratio of 1 was applied across the board for all technologies and fuel types, meaning that the 2011 and later rates are assumed to have the same performance as 2001 through 2010 rates. While evolutionary scenarios could certainly be modeled in MOVES through changes to FERC default inputs, we chose the most conservative future scenario for the default case based on the trend of minimal evolutionary improvement in fuel economy which has occurred over the past decade.

The FERC is a MySQL script which employs input files containing the advanced technology and alternative fuel ratios by fuel and engine technology and the base technologies to which the ratios should be applied. The default 2001 through 2010 model year ratios shown in Table 4-15 are contained in the FERC input file named “ShortTermFERC_PERE” included with the MOVES2004 installation package. The 2011 and later model year ratios (all equal to 1) are contained in the FERC input file named “LongTermFERC” included with the MOVES2004 installation package. Users wishing to use alternate advanced technology assumptions would replace the default ratios in these files prior to executing the FERC. This is designed as an automated process for customized “what-if” sensitivity analysis.

4.3.4.1 Modeling Advanced Technology Vehicles Using PERE

There are actually a number of ways that PERE could be used to generate advanced technology rates in MOVES. It can be run many times for each source bin to generate rates directly, but this would amount to hundreds of runs. Alternatively, a select few of the engine size and weight categories could be run and the rest interpolated. The approach used for MOVES2004, however, was to model a single representative vehicle, bin the results by operating mode, and then ratio these rates to the corresponding “base” conventional vehicle rates across all source bins. This approach is simpler and less time consuming, but requires broad assumptions and is considered a first step. The process for generating these ratios is described in the following steps:

Step 1: Choose the Representative Source Bin

Due to its relative frequency in the fleet, the “representative” vehicle was chosen to lie in a MOVES source bin with weight ranging from 3,500 – 4,000 lbs (test weight), and engine size ranging 3.0 – 3.5 Liters (automatic transmission).

Step 2: Define the Vehicle Specifications

Based on the engine technology source bin, the vehicle parameters such as weight, engine size, vehicle shape etc, can be defined. The weight and engine size

values are simply the central value of the bin. Thus PERE effectively models an “average” vehicle in the source bin. The corresponding values for the representative source bin described under Step 1 to a PERE conventional vehicle of 3,750 lbs, 3.25L, and 155kW (peak power).

Since engine size has limited meaning for hybrids and no meaning for electric vehicles, it is necessary to define a power surrogate for engine size.¹² Though it is not perfect, we take the convention that motor power + engine peak power = total peak power. Taking this into account, the PERE specifications for the representative source bin are an engine size of 1.7L for a full hybrid with a 72kW motor, and 2.8L for a moderate hybrid with a 22kW motor (gasoline and diesel).

Because the source bins do not have a dimension for body type, the road load coefficients are estimated based on the weight. Lighter light duty weights tend to be (compact) passenger cars, then midsize cars, luxury, compact pickups, SUVs, minivans on up to medium-duty trucks. The variation in body types will certainly lead to variation (or uncertainty) in the emission rates.

Step 3: Define the Driving Cycles

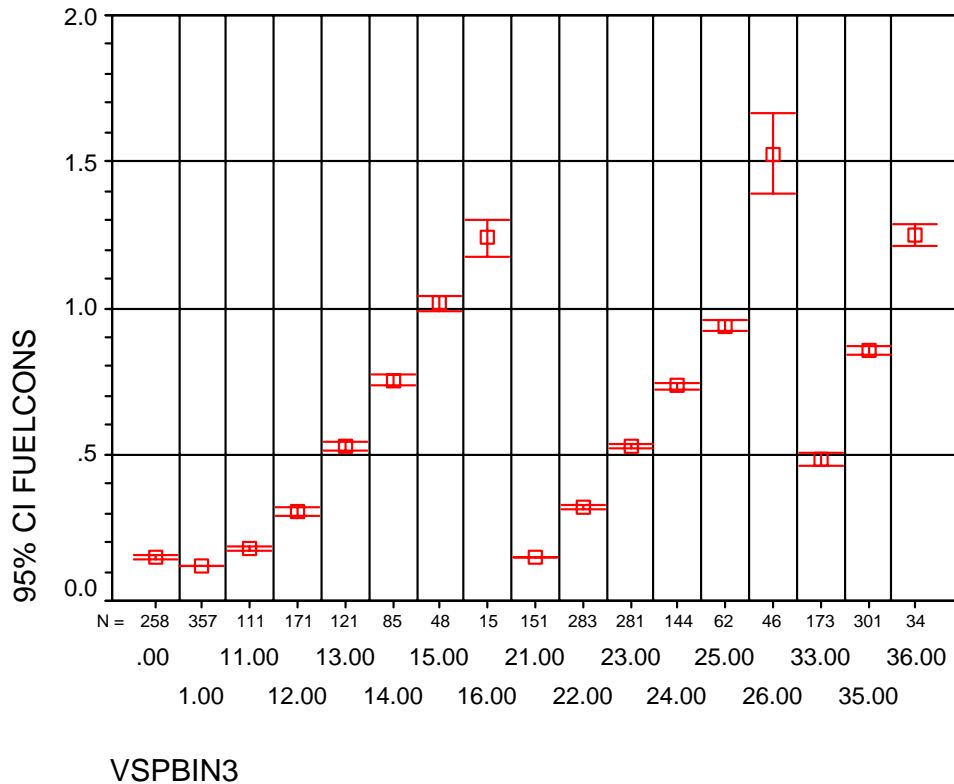
The output to PERE is second-by-second energy consumption, therefore the next step is to input the driving cycles. The driving cycles input into PERE help determine the “binned” energy consumption rates for MOVES. The binned energy consumption and emissions rates are dependent on the driving cycle input. It is important to capture a representative sampling of real-world driving. The number of cycles are reduced for advanced technologies compared to conventional vehicle cycles (described in Section 4.3.3), to speed up the model runs.

Three driving cycles were used for the advanced technology runs: FTP (urban), FTP (highway), and LA92. The latter is a self-weighting cycle and includes harder accelerations and higher speeds. For hybrids, the state-of-charge was maintained over the total of the three cycles.

Step 4: Run PERE and Bin Output

After PERE was run, the second-by-second results were binned into the 17 operating modes described in Section 4.1.3. Figure 4-3 shows a sample of predicted fuel consumption as a function of VSP for a fictional hybrid passenger car. This is an indication of how an output from PERE would be translated into an emissions rate in MOVES. The uncertainty bars are from PERE generated variations within a bin, and do not adequately reflect the true uncertainty of the emission rates, which will be added in future versions of PERE.

Figure 4-3: Hybrid Fuel Consumption by Operating Mode Bin from PERE (g/s)



The binned results for each advanced technology model by PERE were divided by the binned results for the baseline conventional technology runs to generate the ratios. As PERE produces fuel consumption rates which reflect differences in energy content, PERE results for non-gasoline fuels were corrected to gasoline equivalent so that the ratios reflect differences in energy consumption, not fuel consumption.

The ratios generated by PERE for all fuel types and engine technologies included in MOVES are shown in Table 4-15. These ratios are the same as included in the “ShortTermFERC_PERE” file, except for diesel technology ratios, which are shown relative to gasoline conventional internal combustion here but are expressed relative to diesel conventional internal combustion in this file.

Table 4-15: Advanced Technology Running Energy Consumption Ratios (relative to Gasoline Conventional IC)

	Operating Mode Bin																
Fuel/Technology	0	1	11	12	13	14	15	16	21	22	23	24	25	26	33	35	36
Gasoline Conventional IC (CIC)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Gasoline Advanced IC (AIC)	0.88	0.86	0.88	0.89	0.90	0.90	0.90	0.90	0.88	0.89	0.90	0.90	0.90	0.90	0.89	0.90	0.90
Gasoline CIC Hybrid Mild	0.00	0.00	0.00	0.33	1.02	1.08	1.13	1.16	0.00	0.36	0.98	1.05	1.13	1.18	0.82	1.05	1.14
Gasoline CIC Hybrid Full	0.00	0.00	0.00	0.18	0.87	0.97	1.06	1.22	0.00	0.11	0.85	0.95	1.10	1.27	0.63	0.93	1.09
Gasoline AIC Hybrid Mild	0.00	0.00	0.00	0.29	0.92	0.97	1.00	1.03	0.00	0.32	0.88	0.95	1.02	1.07	0.70	0.91	1.01
Gasoline AIC Hybrid Full	0.00	0.00	0.00	0.15	0.78	0.87	0.95	1.14	0.00	0.07	0.76	0.85	0.99	1.21	0.55	0.82	1.01
Diesel Fuel Conventional IC	0.52	0.47	0.54	0.66	0.72	0.75	0.77	0.79	0.51	0.64	0.71	0.73	0.77	0.81	0.63	0.70	0.74
Diesel Fuel Advanced IC	0.48	0.41	0.50	0.62	0.69	0.71	0.74	0.76	0.47	0.61	0.67	0.70	0.74	0.77	0.60	0.67	0.71
Diesel CIC Hybrid Mild	0.00	0.00	0.00	0.15	0.57	0.78	0.84	0.90	0.00	0.45	0.70	0.79	0.87	0.94	0.54	0.72	0.84
Diesel CIC Hybrid Full	0.00	0.00	0.00	0.10	0.52	0.73	0.81	0.96	0.00	0.19	0.63	0.74	0.86	1.02	0.44	0.67	0.83
Diesel AIC Hybrid Mild	0.00	0.00	0.00	0.14	0.54	0.75	0.80	0.86	0.00	0.43	0.67	0.75	0.83	0.90	0.52	0.69	0.81
Diesel AIC Hybrid Full	0.00	0.00	0.00	0.10	0.50	0.69	0.77	0.92	0.00	0.18	0.61	0.71	0.83	0.98	0.42	0.64	0.80
Electric	-0.54	0.08	-0.04	0.17	0.31	0.36	0.39	0.44	-0.08	0.20	0.32	0.40	0.47	0.50	0.23	0.36	0.47

Figures 4-4 and 4-5 show the ratios for gas hybrid and diesel hybrid technologies. Note that it is possible for the ratios to exceed 1 in some operating mode bins. This is due to the fact that the hybrids are heavier vehicles, driven with smaller engines, yet still follow the same driving trace. Thus, it is quite possible for the fuel consumption to be higher than the conventional vehicle during certain modes of driving. However, over an entire driving cycle, the fuel consumption is significantly lower. In most driving cycles, the time spent in modes where the ratio is greater than 1 is usually quite small (high acceleration events). These results support the well-known evidence that the advantage of hybrids is seen in stop-and-go driving.

Modeling hybrid performance by operating mode greatly increases the power of the model to assess real-world performance, compared to the use of a single fixed fuel economy ratio (over an FTP cycle for example) as other studies have proposed. This is due to the fact that hybrid fuel economy depends in large part on the type of driving, idle time, decelerations etc., which are washed out in a single cycle average number.

Figure 4-4. The ratio of energy (or fuel consumption) for gasoline moderate and full hybrid vehicles, relative to gasoline conventional IC

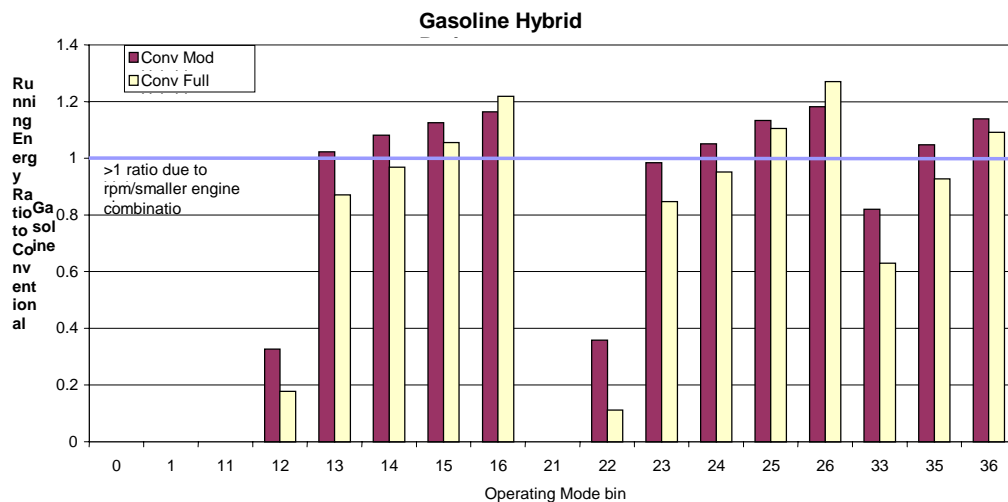
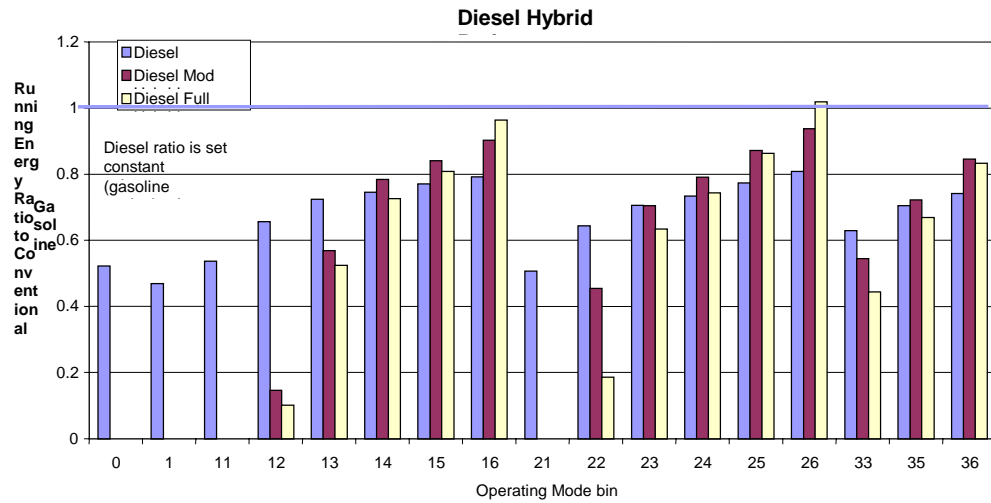


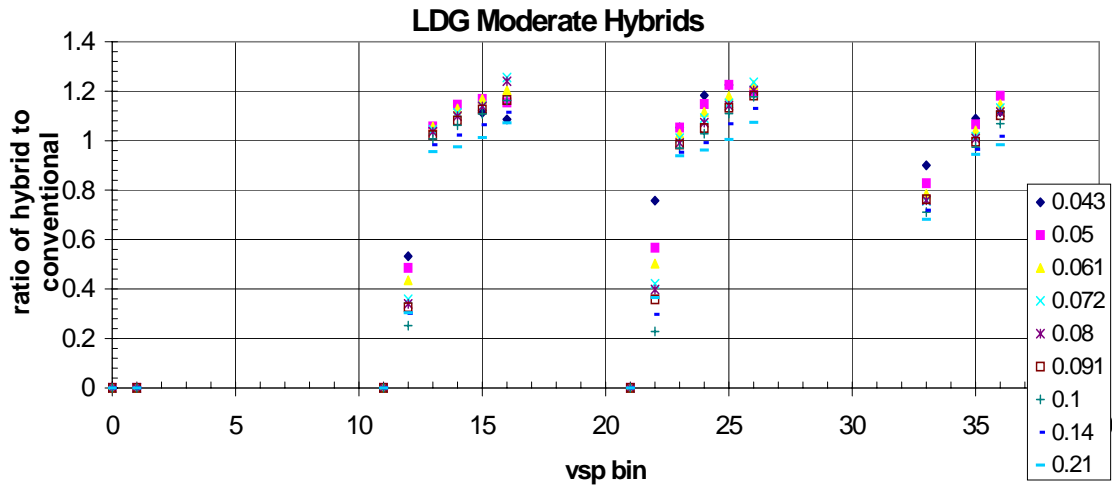
Figure 4-5. The ratio of energy (or fuel consumption) for diesel moderate and full hybrid vehicles, relative to gasoline conventional IC



4.3.4.2 Investigation of Weight-Based Ratios

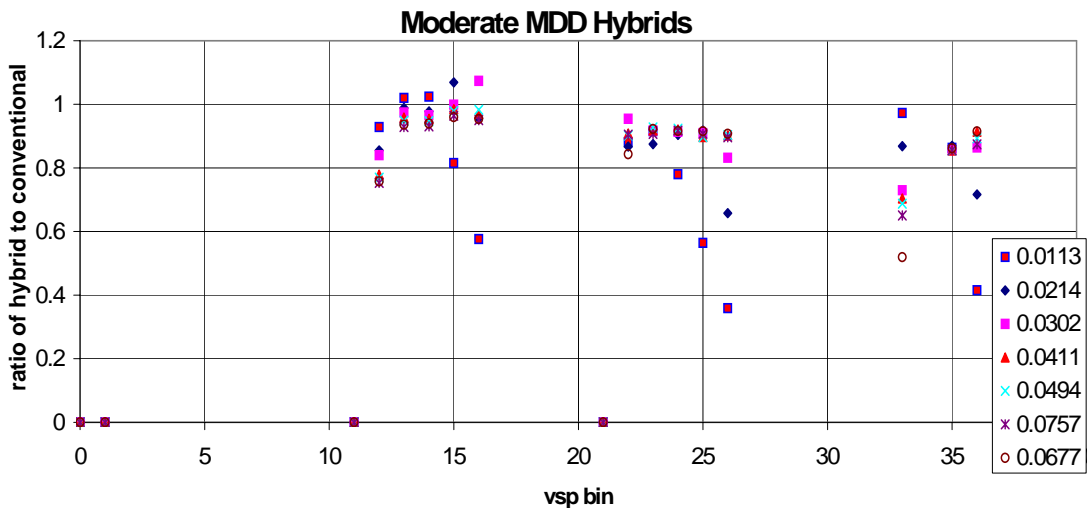
The development of ratios based on a single source bin has limitations, which are presently explored for the purpose of considering next steps for modeling advanced technology vehicles in the evolution of MOVES. The ratios are based on a single (though common) representative vehicle class. We might expect vehicles of different weight but same power-to-weight ratio to give similar rates. However, vehicles with lower power-to-weight ratios would (for example) require more assistance from the engine, both during launch and assist (power boost). This would necessarily change the shape of the curves. An example of this is shown in Figure 4-7. The varying power-to-weight ratios having different shapes. The effect is mainly pronounced in the “2” bins, i.e. bins 12, and 22. These correspond to moderate acceleration from low to medium speed (launch).

Figure 4-7. Fuel consumption ratios for a series of power-to-weight ratio (kW/kg) light-duty gasoline hybrid vehicles.



The shape difference is even more pronounced in medium-duty vehicles where the power-to-weight ratio can fall below 0.01. This is shown in Figure 4-8 for a series of medium-duty diesel hybrid vehicles. Note that the fuel consumption drops at higher loads for the underpowered vehicles. This is due to the fact that the engine cannot keep up with the cycle by itself and more of the total power is supplied by the battery/motor. It is also likely that these vehicles are unable to follow the driving cycles. These hybrids were run on the following medium-duty driving cycles: CBD (Central Business District), NYBUS, NYCCT, and UDDS-D. Although weight-based ratios were not used in MOVES2004, these results merit further consideration of how to improve the resolution of advanced technology modeling in MOVES.

Figure 4-8. Fuel consumption ratios for a series of power-to-weight ratio (kW/kg) medium duty diesel hybrid vehicles.



For future versions of MOVES, the rates may likely be separated by power-to-weight (P/Wt) ratio bins, rather than engine displacement or weight separately in order to differentiate the energy ratios for the different vehicle classes. An example of such a split is shown in Table 4-16. The bands of color represent possible P/Wt ratios to group together into bins. The boxed cells are modeled using PERE.

Table 4-16. Possible Power-to-weight ratio bins (kW/kg) for turbo diesel hybrids.

lo wt bin lbs	avg wt in bin kg	Engine displacement or avg pwr					
		2 95.2	2.5 116.3	3 137.5	3.5 158.6	4 190.3	5 291.9
2000	1021	0.0933	0.1140	0.1347	0.1554	0.1865	0.2860
2500	1247	0.0763	0.0932	0.1102	0.1271	0.1526	0.2340
3000	1474	0.0646	0.0789	0.0933	0.1076	0.1291	0.1980
3500	1701	0.0560	0.0684	0.0808	0.0932	0.1119	0.1716
4000	1928	0.0494	0.0603	0.0713	0.0823	0.0987	0.1514
4500	2155	0.0442	0.0540	0.0638	0.0736	0.0883	0.1355
5000	2495	0.0382	0.0466	0.0551	0.0636	0.0763	0.1170
6000	2948	0.0323	0.0394	0.0466	0.0538	0.0645	0.0990
7000	3402	0.0280	0.0342	0.0404	0.0466	0.0559	0.0858
8000	3856	0.0247	0.0302	0.0357	0.0411	0.0494	0.0757
9000	4309	0.0221	0.0270	0.0319	0.0368	0.0442	0.0677
10000	5443	0.0175	0.0214	0.0253	0.0291	0.0350	0.0536
14000	6804	0.0140	0.0171	0.0202	0.0233	0.0280	0.0429
16000	8051	0.0118	0.0144	0.0171	0.0197	0.0236	0.0363
19500	10319	0.0092	0.0113	0.0133	0.0154	0.0184	0.0283
26000	13381	0.0071	0.0087	0.0103	0.0119	0.0142	0.0218

4.3.4.3 Modeling Alternative Fuels Using GREET

Energy ratios for CNG, LPG, E85 and M85 were taken directly from GREET, which in turn are based on analysis of available certification data on vehicles using these fuels.¹³ GREET uses aggregate ratios of energy consumption, so no operating mode split was available. However, for these fuels, the relative benefit of the technology would be expected to be more uniform across operating mode, and the need to split by operating mode is less important than for advanced technologies, particularly hybrids. The ratios for these fuels derived from GREET are shown in Table 4-17, and are in the “ShortTermFERC_PERE” file used by the FERC to generate future rates.

**Table 4-17: Alternative Fuel Energy Consumption Ratios
(relative to Gasoline Conventional IC)**

Fuel	Ratio
CNG	1.05
LPG	1.0
E85	0.95
M85	0.95

4.4 Start Energy Rate Development

The definition of start energy consumption for MOVES follows the approach initially developed and documented with MOBILE6. With this approach, “start” energy is defined as the energy consumed at startup over and above the energy which would be consumed had the vehicle followed the same trajectory during running (warmed-up) operation. Start energy rates are therefore the incremental amount of energy consumed at start-up, and start rates in the model are in the units of KJ per start. Starts were not separated into operating mode in MOVES2004. The main ramification of this is that MOVES2004 does not differentiate between hot start and cold start energy consumption (e.g soak time). We expect to expand the approach used for start emissions to include soak time and load-based effects when the model is developed for criteria pollutants.

4.4.2 Data Sources

Data used in the development of start rates analysis came from EPA's Mobile Source Observation Database (MSOD) as of April 2003. The initial analyses were limited to the FTP tests that were performed within the temperature range of 68 degrees to 86 degrees Fahrenheit (i.e., at a nominal temperature of 75° F). This restriction produced a database of 18,676 FTPs performed on 10,422 vehicles. Only FTP tests were used, since the basis of start emission rates is the difference between Bag 1 cold start and Bag 3 hot start. Whereas running energy rates are based entirely on second-by-second data, start energy rates are based entirely on aggregate FTP bag results.

4.4.3 Methodology for Rate Development

4.4.3.1 Assessing Hot Start vs. Hot Running

The approach used to estimate the amount of fuel consumed during an engine start requires identifying vehicles that were tested over driving cycles that differed only by the presence of an engine start. What we found were 192 vehicles (in the MSOD) that were tested over a Federal Test Procedure (FTP) driving cycle followed by an immediate "Hot-Running 505" (HR-505).

The FTP driving cycle consists of three individual operating modes:

- The first mode (Bag-1) is a 505-second driving cycle that begins with a cold start (i.e., following a soak of 12 to 36 hours). This mode is referred to as a "Cold-Start 505" (CS-505).
- The second mode (Bag-2) is an 867-second driving cycle that involves no engine start.
- The third mode (Bag-3) is a 505-second driving cycle identical to the first mode but begins with a hot start (i.e., following a soak of only 10 minutes). This mode is referred to as a "Hot-Start 505" (HS-505).

Following the completion of the FTP (but without shutting off the engine), 192 vehicles then drove another 505-second driving cycle identical to both the first and third modes. Since this cycle involved NO engine start, it is referred to as a "Hot-Running 505" (HR-505).

Subtracting the fuel consumed over the HR-505 from the fuel consumed over the corresponding CS-505 should provide an estimate of the fuel consumed during an engine cold-start itself. However, this data set is relatively small considering the analyses we wish to perform. Hence, our approach will be to first develop a substantially larger data set for estimating fuel consumed during engine start. Specifically, we will consider whether we can use the fuel consumed over the HS-505 as a surrogate for the fuel consumed over the HR-505.

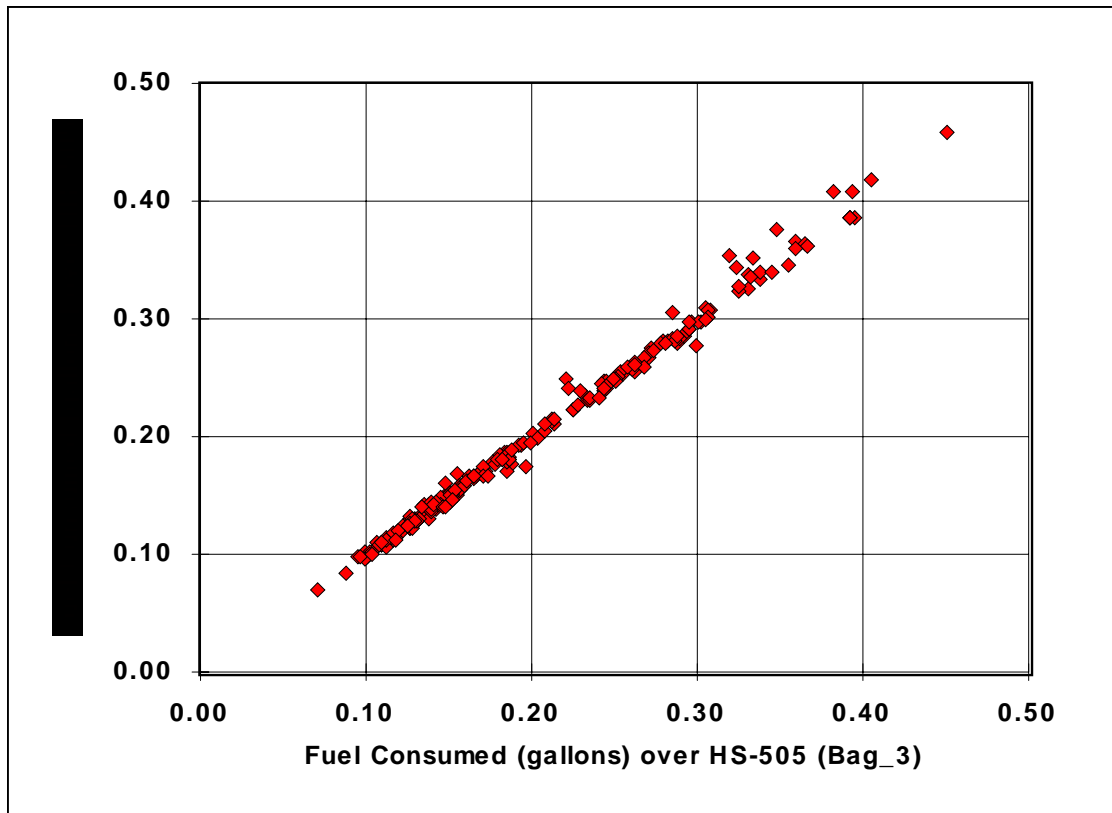
To compare the fuel consumed over the "Hot-Start 505" with the fuel consumed over the corresponding "Hot-Running 505" cycle, we used 244 test pairs (FTP and "Hot-Running 505") on those 192 vehicles. In Figure 4-9, we plotted the fuel consumed over the HR-505 versus the fuel consumed over the corresponding HS-505 cycle. Even the most cursory visual inspection of that graph reveals a very strong match between these two quantities. A more rigorous approach using a linear regression produces the equation for the fuel consumed (in gallons), with an R-squared value of 0.99:

$$\text{Eqn 4-1 } HR505 = (1.0095 * HS505) - 0.002$$

The slope in the preceding equation has a 95 percent confidence interval of 0.999 to 1.019 (with a P-value less than 0.00001), and the intercept has a 95 percent confidence interval of -0.0041 to 0.00015 (with a P-value of 0.069). This strong linear correlation and the virtual equality implied by this regression equation suggests that fuel consumed during the HS-505 is a reasonable estimate of the fuel used during the HR-505. (Additionally, in 107 of those 244 test pairs, the fuel consumed during the HR-505 exceeded the amount consumed over the HS-505. In the remaining 137 pairs the reverse was true. This nearly equal split also argues for the HS-505 being a reasonable estimate of the HR-505.)

Therefore, for the remainder of this analysis, we used the difference of the fuel consumption on Bag-1 of the FTP minus the fuel consumption on Bag-3 of the FTP as an estimate of fuel consumed during a cold engine start.

Figure 4-9: Comparison of Fuel Consumed Over HS-505 and HR-505



4.4.3.2 Determining Source Bin Variables

From the MSOD, we identified 20,156 FTPs that were performed on 10,516 vehicles. The vehicle and test result data from those tests were analyzed using a stepwise regression process in order to determine the most important variables to consider for source bins.

The stepwise regression process first uses the Pearson Product-Moment to select the independent variable that has the highest correlation with the "Fuel Consumed During Start" (estimated by subtracting the Bag-3 fuel consumption from the corresponding Bag-1 fuel consumption). The difference between the best linear estimate using that variable and the "Fuel Consumed During Start" term (i.e., the residuals) is then compared with the set of remaining variables to identify the variable having the next highest correlation. This process continues as long as the "prob" values do not exceed (an arbitrary) five percent, thus, creating a sequence of variables in descending order of statistical correlation. The rank ordering produced by this process is dependent upon the independence of the variables. In this instance, there is some collinearity among the variables, which may reduce the usefulness of this statistical tool.

For each of those 20,156 FTPs, the difference of the fuel consumed on Bag-1 minus the fuel consumed on Bag-3 was calculated as an estimate of fuel consumed during an engine start. Those values (in gallons) were used as the dependent variable along with the following 12 variables that were considered as independent variables in this stepwise analysis:

- Age (estimated as "Model Year" minus test year)
- Ambient Temperature
- Engine Displacement (in cubic inches)
- Number of Cylinders
- Model Year (four-digit year ranging from 1965 through 2000)
- Odometer
- Test Weight
- Car v. Truck
- Fuel Delivery System (Carbureted v. Fuel Injected)
- Fuel Injection System (TBI v. PFI)
- Gasoline v. Diesel
- Light-Duty (cars & trucks) v. Heavy-Duty

The last five variables on the preceding list are categorical variables (i.e., ones and zeros).

Some of these variables are potentially collinear, and that collinearity could affect the usefulness of the regression process. Thus, we were cautious about using (in the same regression) those pair of potentially collinear variables. Those pairs are:

- Engine Displacement / Number of Cylinders,
- Model Year / Fuel Delivery System, and
- Odometer / Age.

The stepwise regression process continued for eight steps, the last of which is shown in Table 4-18.

Table 4-18: Stepwise Regression of Start (Bag 1 – Bag 3) Fuel Consumption

Dependent variable is:		Start (Bag-1 minus Bag-3)		
No Selector				
20156 total cases of which 3960 are missing				
R squared = 38.5% R squared (adjusted) = 38.5%				
s = 0.0144 with 16197 - 9 = 16188 degrees of freedom				
Source	Sum of Squares	df	Mean Square	F-ratio
Regression	2.08917	8	0.261146	1267
Residual	3.3355	16188	0.000206	
Variable	Coefficient	s.e. of Coeff	t-ratio	prob
Constant	1.40283	0.0786	17.9	≤ 0.0001
Disp (CID)	0.000072	0.0000	17.7	≤ 0.0001
Avg_Temp	-0.000710	0.0000	-64.1	≤ 0.0001
Model_Year	-0.000674	0.0000	-17.0	≤ 0.0001
Carb / FI	-0.003503	0.0003	-11.1	≤ 0.0001
LD / HD	0.007635	0.0011	7.2	≤ 0.0001
Estimated_Age	0.000312	0.0000	4.45	≤ 0.0001
No._of_Cylinders	0.001212	0.0002	6.11	≤ 0.0001
Odometer (miles)	0.000000	0.0000	-3.12	0.0018

From this analysis we concluded that engine displacement and model year were important variables to bin by. The importance of average temperature also indicated the need for a separate temperature effect, discussed in Section 9. The coefficients values also suggest that the difference in actual fuel consumption during engine starts may be substantial between light-duty and heavy-duty vehicles.

4.4.4 Emission Rate Development

With no operating mode bins and only aggregate bag data to work with, the development of emission rates was straightforward. Energy rates, in terms of KJ per start, were calculated for each combination of engine displacement and model year group from Table 4-1 by averaging the start (Bag 1 minus Bag 3) results for all FTP tests from vehicles falling in those bins. Vehicle weight did not appear as an important variable in the stepwise regression analysis, so test data was not segregated by weight for this analysis. However, since total energy source bin definitions include the vehicle weight class, the start energy rates were set at the same level for all weight classes within an engine displacement / model year group bin.

The start emission rates by source bin category which resulted from this analysis are presented in Appendix E.

4.4.5 Hole Filling

Despite the large number of tests, a number of source bins lacked adequate data to fill directly – mostly for heavy-duty and/or diesel vehicles. We filled these holes mostly by copying from other bins where data existed, or applying ratios to existing rates based on trends observed in the data. This section provides details for this process. The calculations as described here were performed in fuel consumption space first, then converted to energy consumption for use in MOVES. The rates developed from this process are presented in Appendix E.

In comparing the differences between the fuel consumed during engine starts for the light-duty gasoline cars and trucks (LDGs) and for the heavy-duty gasoline-fueled vehicles (HDGVs), we noted:

- For the HDGVs, there were data only in the 1986-90 and 1991+ model year bins and in only the 4.1-5.0 and 5.1+ displacement bins.
- Within each of the two displacement bins, the effect of changing model year ranges was not statistically significant.
- After combining all model year data, the resulting (two) displacement bins were not statistically different from the corresponding displacement bins for the Pre-1981 light-duty trucks & cars. (Probably due to lack of CAFE standards for the HDGVs.)

Therefore, we used the results (i.e., means) from the Pre-1981 light-duty gasoline-fueled cars and trucks for the HDGVs.

In comparing the differences between the fuel consumed during engine starts for the heavy-duty gasoline-fueled vehicles (HDGVs) and for the heavy-duty diesel-fueled vehicles (HDDVs), we noted:

- For the HDDVs, there were data only in the 1986-90 and 1991+ model year bins and in only the 5.1+ displacement bin.
- Although the difference in average engine-start fuel consumption was statistically significant for those two model year bins, the fact that the 1986-90 bin contained data on only two vehicles makes it difficult to justify splitting by model year.
- Combining results on the 2 tests from the 1986-90 model year range with the 29 tests from the 1991 and newer range, produces a mean of 0.0277 gallons per start with a standard deviation of 0.0247.
- Comparing the mean and standard deviation with those from the HDGVs (over 5.0 liter displacement), we note the standard deviations are close and that the fuel consumption (in gallons) was reduced by 46 percent.

Therefore, we calculated the average start fuel consumption of the HDDVs as the number of gallons of gasoline used by the corresponding HDGVs reduced by 46 percent.

In comparing the differences between the fuel consumed during engine starts for the light-duty gasoline-fueled vehicles (LDGVs) and for the light-duty diesel-fueled vehicles (LDDVs), we noted:

- Except for the 2.6-3.0 liter displacement bin, the amount of diesel fuel consumed (in gallons) for each engine start is only 60 to 80 percent of the number of gallons of gasoline consumed for each corresponding engine start.
- Although there are no results for the 1986-1990 model year diesel-fueled vehicles, we note that, for the corresponding gasoline-fueled vehicles, the amount of gasoline consumed for an engine start is about 3 percent higher than for the corresponding 1991 and newer vehicles. Applying this factor to the diesel fuel consumed (by that single test) of the 1991+ vehicle permits us to obtain an estimate for the 1986-90 diesel light-duty cars and trucks.

Therefore, we propose to estimate the average engine-start fuel consumption of the LDDVs, by:

- Multiplying by 70 percent (the mid-point from the first observation) the number of gallons of gasoline consumed by the light-duty vehicles to estimate the number of gallons of diesel fuel consumed (for an engine start) for all but the 2.6-3.0 liter displacement bin.
- Using the calculated mean start fuel consumption for the LDDVs with 2.6-3.0 liter engines for all model year ranges except the 1986-1990.
- For the 1986-90 model year diesels with 2.6-3.0 liter engines, use the fuel consumption of the corresponding 1991+ vehicles increased by 2.91 percent (where 2.91 percent was the increased start fuel consumption of the 1986-1990 LDDVs compared to their 1991+ counterparts).

The final category of holes to fill were motorcycles. In MOVES, motorcycles are stratified into two groups, “uncontrolled” and “controlled”. The uncontrolled motorcycles have engines similar to but usually smaller than those used in the pre-1981 model year passenger cars and light-duty trucks. The controlled motorcycles have engines similar to but usually substantially smaller than those used in the 1991-2000 model year passenger cars and light-duty trucks. Performing linear regressions on the fuel consumption estimates for those two model year ranges produce the following equations for fuel consumed during an engine start:

$$\begin{aligned} \text{Eqn 4-2 Uncontrolled Motorcycle Energy Consumption (KJ)} \\ = 2048.1 + (717.16 * \text{Displacement}) \end{aligned}$$

$$\begin{aligned}\text{Eqn 4-3 Controlled Motorcycle Energy Consumption (KJ)} \\ = 1402.15 + (477.97 * \text{Displacement})\end{aligned}$$

Using these regression equations, we are able to estimate the fuel consumed (as functions of engine displacement) during an engine start of each type of motorcycle.

4.4.6 Advanced Technologies and Alternative Fuels

Advanced technology and alternative fuel start rates were developed in the same manner as for running energy rates; by applying ratios to base conventional technology via the Future Emission Rate Creator (FERC). There is a dearth of data on cold start factors for advanced technology vehicles, so some assumptions were required to develop these ratios. A comparison of cold start energy consumption was conducted on two hybrid vehicles (details available in the PERE report), from which it was decided that hybrid cold start factors may be consistent with those for conventional vehicles. For gasoline and diesel-fueled vehicles, we therefore based the ratios solely on the changes in engine displacement assumed in the development of the running energy ratios. Advanced internal combustion engines were therefore given no additional benefit, but hybrid vehicles are assigned benefit based on the expectation that engine size will be reduced.

Cold start factors for hybrids were approximated by developing a linear function of start fuel consumption as a function of engine displacement on the conventional vehicles, plugging in the smaller engine displacement assumed for hybrids, and calculating the ratio of conventional start fuel energy to hybrid start energy for use in the FERC. Based on an analysis of start fuel consumption rates versus engine displacement from the data detailed in Section 4.4.1, the following regression equations were used for gasoline and diesel:

$$\text{Eqn 4-4 Gasoline Consumed (gallons)} = 0.0114 + (0.0039 * \text{Displacement})$$

$$\text{Eqn 4-5 Diesel Consumed (gallons)} = 0.0072 + (0.0028 * \text{Displacement})$$

Alternative fuel ratios were derived from GREET, and are the same as the running ratios presented in Section 4.3. The ratios used for start are shown in Table 4-19. These ratios are the same as included in the “ShortTermFERC_PERE” file, except for diesel technology ratios, which are shown relative to gasoline conventional internal combustion here but are expressed relative to diesel conventional internal combustion in the FERC file.

**Table 4-19: Advanced Technology & Alternative Fuel
Start Energy Consumption Ratios (relative to Gasoline Conventional IC)**

Fuel / Technology	Start Energy Ratio
Gasoline Advanced IC	1.0
Gasoline CIC Hybrid Moderate	0.93
Gasoline CIC Hybrid Full	0.75
Gasoline AIC Hybrid Moderate	0.83
Gasoline AIC Hybrid Full	0.67
Diesel Fuel Conventional IC	0.77
Diesel Fuel Advanced IC	0.73
Diesel CIC Hybrid Moderate	0.71
Diesel CIC Hybrid Full	0.57
Diesel AIC Hybrid Moderate	0.67
Diesel AIC Hybrid Full	0.54
Compressed Natural Gas (CNG) CIC	1.05
Liquid Propane Gas (LPG) CIC	1.0
Ethanol (E85) Conventional IC	0.95
Methanol (M85) Conventional IC	0.95
Electric	1.0

4.5 Extended Idle

The extended idle process was added to MOVES primarily to account for “hoteling” of long-haul heavy-duty trucks, whose activity isn’t explicitly accounted for in VMT estimates or in the driving schedules used to characterize on-road operation. Extended idle for long-haul trucks generally occurs at truck stops, to allow drivers to heat or air condition the truck cab overnight. Energy and emission rates tend to be higher for extended idle relative to on-road idle operation, since RPM is usually set higher (e.g. 1200 RPM versus 600 RPM for on-road idle) to provide adequate power for the auxiliary systems.

Limited emission data exists for extended idle conditions, so direct binning of test data was not an option across the range of source bins in MOVES. However, EPA has conducted a test program to assess the relative change in energy and some emissions on short periods of idle vs. longer periods.¹⁴ To develop energy rates for extended idle, this work was used to generate an adjustment factor to be applied to the “idle” bin data for running total energy (Bin 1). A series of tests conducted in May 2002 on 5 trucks at paired varying idle speeds, with and without air conditioning. The most common low-high idle speed pair was 600 RPM (simulating on-road idle) and 1200 RPM (simulating extended idle), so from the original sample, we analyzed the difference in fuel consumption between all points at 600 RPM and 1200 RPM. The results are shown in Table 4-20.

Table 4-20: Fuel Consumption Rates At Varying Idle Speeds

RPM	A/C Status	Fuel Rate (gallons/hour)
600	Off	0.503
1200	Off	1.112
600	On	0.565
1200	On	1.373
Relative Increase	Off	2.22
Relative Increase	On	2.43

For the A/C off case, fuel consumption increased by a factor of 2.22 between 600 and 1200 RPM.

The default case for MOVES2004 only includes extended idle for combination long-haul trucks, so that emission rates are only needed for those source bins which are mapped to this category. However, to provide the possibility for modeling extended idle for all source types, we generated extended idle rates for all total energy source bins (this would allow extended periods of idling in fast-food drive-thrus, school zones, etc. to be accounted for if desired). To generate extended idle rates for all source bins, we applied the 2.22 multiplier from Table 4-20 to the Bin 1 on-road idle rates within a given source bin, for all loaded weight categories greater than or equal to 33,000 pounds (the de facto “heavy duty” cutpoint). Since we did not have data for the lighter vehicle weight range, we set extended idle rates equal to the Bin 1 on-road idle rates for loaded weights below 33,000 lbs.

Advanced technology and alternative fuel rates for extended idle were generated using the same adjustment process described for running and start in Sections 4.2 and 4.3. The adjustment ratios used for extended idle were estimated based on the adjustments for the idle bin under running energy consumption, although some modifications were made to account for the fact the length of idle period would preclude the battery as a power source for hybrid vehicles. The ratios used for extended idle are presented in Table 4-21, and are included in the FERC input table “ShortTermFERC_PERE”.

**Table 4-21: Advanced Technology & Alternative Fuel
Extended Idle Energy Consumption Ratios
(relative to Gasoline Conventional IC)**

Fuel/Technology	Running Idle	Extended Idle	Comments
Gasoline Conventional IC (CIC)	1	1	
Gasoline Advanced IC (AIC)	0.86	0.86	
Gasoline CIC Hybrid Moderate	0.00	1.00	assumed to run off engine
Gasoline CIC Hybrid Full	0.00	1.00	assumed to run off engine
Gasoline AIC Hybrid Moderate	0.00	1.00	assumed to run off engine
Gasoline AIC Hybrid Full	0.00	1.00	assumed to run off engine
Diesel Fuel Conventional IC	0.47	0.47	
Diesel Fuel Advanced IC	0.41	0.41	
Diesel CIC Hybrid Moderate	0.00	1.00	
Diesel CIC Hybrid Full	0.00	1.00	
Diesel AIC Hybrid Moderate	0.00	1.00	
Diesel AIC Hybrid Full	0.00	1.00	
Electric	0.08	0.08	

5. Petroleum and Fossil Energy Calculations

In MOVES2004, petroleum and fossil energy are the quantities of energy consumption derived from petroleum or fossil-based sources. Petroleum energy is a subset of fossil energy, the latter including all of fuel subtypes CNG, LPG and M85 as well as gasoline and diesel. Well-to-pump petroleum and fossil energy rates are generated directly by GREET, and passed into the GreetWellToPump table. Estimates of petroleum and fossil-based energy are calculated for the pump-to-wheel processes (running , start, extended idle) in MOVES by multiplying total energy consumption results by the fraction of energy which is either petroleum or fossil-based. With this approach, direct rates of petroleum and energy consumption are not required in the emission rate database.

The petroleum and fossil energy fractions for the pump-to-wheel processes were derived from GREET values for on-road energy consumption. Specifically, Table 2 of the “Results” sheet within GREET (spreadsheet version) contains gram/mile total, petroleum and fossil energy results for the range of fuels and vehicles in GREET. To generate the petroleum and fossil fractions for MOVES, we divided the petroleum and fossil energy gram/miles results attributed to “vehicle operation” by the corresponding total energy value. The results for each MOVES fuel subtype are shown in Table 5-1, including the corresponding GREET fuel / vehicle category used to derive the results.

Table 5-1: Pump-To-Wheel Petroleum and Fossil Energy Fractions by Fuel Subtype

Fuel Subtype	GREET Fuel / Vehicle Category	Petroleum Fraction	Fossil Fraction
Conventional Gasoline	Baseline Gasoline	0.95	1.0
Reformulated Gasoline	Baseline Gasoline	0.95	1.0
E10	Low-level EtOH Blend	0.94	0.94
Conventional Diesel	Conventional Diesel	1.0	1.0
Biodiesel (BD20)	BD20	0.81	0.81
Fisher-Tropsch (FT100)	FT100	0	1.0
CNG	Dedicated CNGV	0	1.0
LPG	LPGV: Dedicated	0.4	1.0
Ethanol (E85)	EtOH: FFV E81 Corn	0.26	0.26
Methanol (M85)	MeOH FFV: M85 nNA NG	0.26	1
Gaseous Hydrogen	FCV: GH2	0.01	0.94
Liquid Hydrogen	FCV: LH2	0.01	0.94
Electricity	Electric	0.02	0.87

6. Carbon Dioxide (CO₂) Calculations

MOVES2004 does not currently estimate CO₂ emissions, but is planned for future releases using the methodology presented here. CO₂ will be calculated from total energy consumption results, rather than through the direct use of CO₂ emission rates. CO₂ will be calculated in this way according to equation 6-1:

$$\text{Eqn 6-1 } \text{CO}_2 = \text{Total Energy Consumed} * \text{Carbon Content} * \text{Oxidation Fraction} * (44/12)$$

This equation methodology is consistent with methods to calculate CO₂ inventories used in the U.S. Inventory of U.S. Greenhouse Gases Emissions and Sinks (hereafter referred to as Emissions & Sinks), in line with International Panel on Climate Change (IPCC) guidelines.¹⁵ In this equation, carbon content is in terms of energy, expressed in grams per KJ, and the oxidation fraction is the percent of carbon which winds up as CO₂ in the atmosphere – the portion which remains unoxidized is generally black carbon (particulate matter) emissions.

The FuelSubType table in the MOVES default database contains the carbon content and oxidation fraction values used in equation 6-1, by fuel subtype. Carbon content values were derived from GREET estimates of carbon weight fraction and Lower Heating Values (LHVs) from Heywood.¹⁶ The Heywood LHVs were used for consistency, since they were used in the binner program to convert fuel rates to energy rates. The oxidation fractions were taken directly from the Emissions & Sinks report. Carbon content, the inputs to carbon content and oxidation fraction are shown by fuel subtype in Table 6-1:

Table 6-1: Carbon Content and Oxidation Fraction by Fuel Subtype

Fuel Subtype	Lower Heating Value (KJ/g)	Carbon Weight %	Carbon Content (g/KJ)	Oxidation Fraction
Conventional Gasoline	44.0		0.0196	0.99
Reformulated Gasoline	42.9		0.0196	0.99
E10	-	-	0.0196*	0.99
Conventional Diesel	43.2		0.0200	0.99
Biodiesel (BD20)	-	-	0.0199*	0.99
Fisher-Tropsch (FT100)	41.6	85.3%	0.0205	0.99
CNG	45.0	72.4%	0.0161	0.995
LPG	46.4	82.0%	0.0177	0.995
Ethanol (E85)	-	-	0.0194*	0.99
Methanol (M85)	-	-	0.0189*	0.99
Gaseous Hydrogen	120	0%	0.0	0
Liquid Hydrogen	120	0%	0.0	0
Electricity	-	-	0.0	0
M100**	20.0	37.5%	0.0188	-
E100**	26.9	52.2%	0.0194	-
BD100**	40.2	77.6%	0.0193	-

* weighted average of blended fuels

** reference fuels only

7. Methane (CH₄) and Nitrous Oxide (N₂O) Rates

7.1 Source Bin Definitions

Unique source bin categories were defined for CH₄ and N₂O since the factors important for their formation differ from energy consumption, and because the data used to generate rates were much more limited than for energy. Important vehicle characteristics for these pollutants are likely in line with those important for HC and NO_x, e.g. emission standards, vehicle age, emitter category, and vehicle class. The limited test data on CH₄ and N₂O restricts the ability to split source bins by all of these categories.

Emission rates for CH₄ and N₂O are reported in the Emissions & Sinks report according to fuel type, broad vehicle class (e.g. light-duty and heavy-duty) and model year. We decided to stay consistent with this classification, but added the full range of fuel types and advanced technologies to allow modeling of these bins alongside energy consumption. The source bin classifications for CH₄ and N₂O are shown in Table 7-1. The changes from energy consumption source bin categories are: finer definitions of model year group to account for changes in standards, elimination of loaded weight and engine size, and addition of a new category, “regulatory class”, meant to capture differences due to vehicle class-based emission standards.

Table 7-1: Source Bin Categories for CH₄ & N₂O

Fuel Type	Engine Technology	Model Year Group	Regulatory Class
Gas	Conventional IC (CIC)	1972 and earlier	Motorcycle
Diesel	Advanced IC (AIC)	1973	Light Duty Vehicle
CNG	Moderate Hybrid - CIC	1974	Light Duty Truck
LPG	Full Hybrid - CIC	.	Heavy Duty Vehicle
Ethanol (E85)	Moderate Hybrid - AIC	.	
Methanol (E85)	Full Hybrid - AIC	.	
Gas H ₂	Fuel Cell	1999	
Liquid H ₂	Hybrid - Fuel Cell	2000	
Electric		2001-2010	
		2011-2020	
		2021-2050	
	(See Table 4-14 for combinations of fuel type and engine type used in MOVES2004)		

7.2 Operating Mode Definitions

Virtually no second-by-second data were available for CH₄ and N₂O, making it unfeasible to split emissions by the 17 bins devised for total energy. Much of the test data was conducted over the standard Federal Test Procedure, which limited to ability to develop emission rates at varying operating conditions. We therefore decided to define a single operating model for running and a single operating mode for start, with FTP bag data used as the basis to develop rates for these modes.

7.3 Data Sources

EPA has published estimated on-road rates for both CH₄ and N₂O in the report entitled "Direct and Indirect Emissions from Mobile Combustion Sources".¹⁷ Those values were based on estimates from Annex E of the Emissions & Sinks report. In both documents, the rates represented emissions produced (in grams per mile) over the FTP test.

Under contract, EPA had ICF Consulting revise those estimates of CH₄ and N₂O emissions: 1) making use of all available test results, 2) splitting the emissions associated with engine start from the running emissions, and 3) converting the running emissions into units of grams per hour (rather than grams per mile). The ICF analyses are documented in the report entitled "Update of Methane and Nitrous Oxide Emission Factors for On-Highway Vehicles".¹⁸

EPA provided ICF with test results containing methane measurements over 13,277 FTP tests on 6,950 vehicles and 14,636 non-FTP tests on 2,963 vehicles; and with test results containing nitrous oxide measurements on 95 FTP tests on 64 vehicles and 232 non-FTP tests on 74 vehicles. The non-FTP tests included a hot running 505 as well as several other driving cycles not utilized in this report. Methane tests were performed in various U.S. locations during the period between April 1982 and June 2000. Nitrous oxide tests were performed in various U.S. locations during the period between June 1998 and January 2000. The analyses performed by ICF were limited to the FTP tests that were performed within the temperature range of 68 degrees to 86 degrees Fahrenheit (i.e., at a nominal temperature of 75° F).

7.4 Running and Start Rate Development

Since the goal of the ICF analysis was to develop separate emission rates for both the running operation and engine starts, the analyses focused on the FTP tests which contained both of those two types of vehicle operation. The first step was to determine how to split the FTP emissions into start and running emission rates.

7.4.1 N₂O Emission Rates

The FTP test results were reported separately for each of three "bags" (or modes). Bags 1 and 3 each begin with an engine start, and they have identical driving cycles (3.59 miles requiring 505 seconds of vehicle operation). The Bag 2 mode has no engine start and requires 867 seconds of vehicle operation to travel 3.86 miles. A relatively small number of those vehicles had a fourth mode performed following the FTP. This additional mode was identical to the first bag/mode (hence, identical to the third bag/mode) with the exception that it contained no engine start. (The fourth mode is referred to as a HR-505 for "hot running 505.") A comparison among those three 505-second modes could, therefore, lead to estimates of start emissions. However, the small

number of vehicles receiving this additional testing limited the usefulness of such a direct comparison.

For the FTPs for which N₂O was measured, ICF identified a total of 21 for which that additional 505-second additional mode was performed. (Those 21 vehicles consisted of 9 gasoline-fueled passenger cars and 12 gasoline-fueled light-duty trucks, all 21 were determined to be Tier-1 vehicles.) Within each of these two vehicle classes, ICF found that the difference in N₂O emission rates (measured in units of grams per mile) between the Bag-2 mode and this fourth mode were not statistically significant (at the 95 percent confidence level). A similar comparison was performed using emission rates in units of grams per hour. In this comparison, the differences between the emissions of the two modes were found to be statistically significant. We therefore used the Bag-2 N₂O emission rates (in grams per mile) as estimates of the N₂O "running" emissions.

The FTP test is actually a weighted average of those three modes. Thus, the FTP simulates a driving cycle, nominally 7.44 miles in length, requiring 1,372 seconds of vehicle operation, and with a single "generic" engine start. The "generic" start is a weighted average of 57 percent of a hot-start and 43 percent of a cold-start.

Subtracting the Bag-2 N₂O emission rate from the weighted FTP emission rate (both in units of grams per mile) and then multiplying that difference by the actual weighted FTP distance (varies with each test, but about 7.44 miles) yields an estimate of the N₂O emissions (in grams per start) associated with each generic start. A similar analyses could have been performed separately for Bag-3 and for Bag-1 to estimate the individual cold-start and hot-start emission, respectively. However, the MOVES2004 model will not take advantage of that level of precision.

In a few instances, this approach to estimating engine start emissions led to estimates of the N₂O start emissions being negative. That negative value is inconsistent with the mechanism forming N₂O emissions. Therefore, in those few cases, the estimates of negative N₂O start emissions were rounded up to zero.

Multiplying the FTP N₂O emission rate (in grams per mile) by the actual weighted FTP distance traveled and then subtracting the calculated grams per start emissions produces the estimated total N₂O emissions (in grams) from the "running" operations. Dividing that value by the duration of the driving cycle, 0.38 hours, yields an estimate of the N₂O emissions (in grams per hour) associated with "running" operations of the vehicle.

These calculations were performed for each of the vehicle type/control technology groups used in the development of EPA's greenhouse gas inventory reports,¹⁹ shown in Table 7-2. This grouping required conversion to the model year-based source bins proposed for MOVES2004. Therefore, it was necessary to weight the ICF estimates together for each model year bin in MOVES according to fractions of control technology and emission standard by model year used in the Emissions & Sinks report. The resulting

model year-based rates, which were used directly in the Emission Rate table of the MOVES default database, are shown in Appendix F.

Table 7-2: Proposed Start and Running N₂O Rates By Control Technology

Vehicle Type	Control Technology	N2O Running (g/hr)	N2O Start (g/start)
Gasoline Passenger Cars (LDGV)			
	LEVs	0.00841	0.09015
	Tier 1	0.28501	0.11280
	Tier 0	0.81124	0.09183
	Oxid Catalyst	0.63214	0.07156
	Non-Catalyst	0.25054	0.02836
	Uncontrolled	0.25054	0.02836
Gasoline Light-Duty Trucks (LDGT)			
	LEVs	0.02346	0.05891
	Tier 1	0.79287	0.20046
	Tier 0	1.35376	0.15324
	Oxid Catalyst	0.82198	0.09305
	Non-Catalyst	0.28336	0.03208
	Uncontrolled	0.28686	0.03247
Gasoline Heavy-Duty Vehicles (HDGV)			
	LEVs	0.04791	0.12032
	Tier 1	1.61934	0.40942
	Tier 0	2.76667	0.31318
	Oxid Catalyst	1.71067	0.19364
	Non-Catalyst	0.62240	0.07045
	Uncontrolled	0.65375	0.07400
Diesel Passenger Cars (LDDV)			
	Advanced	0.02303	0
	Moderate	0.02354	0
	Uncontrolled	0.02790	0
Diesel Light Duty Trucks (LDDT)			
	Advanced	0.03209	0
	Moderate	0.03144	0
	Uncontrolled	0.03633	0
Diesel Heavy Duty Vehicles (HDDV)			
	Advanced	0.09598	0
	Moderate	0.09598	0
	Uncontrolled	0.09598	0
Motorcycles (MC)			
	Controlled	0.10315	0.01168
	Uncontrolled	0.13173	0.01491

7.4.2 CH₄ Emission Rates

The approach used to split the running and the engine start CH₄ emissions paralleled the one used for N₂O emissions. The contractor (ICF) first identified test pairs (FTP and HR-505s) for which methane was measured. A total of 345 such pairs were identified. Those 345 vehicles consisted of 89 gasoline-fueled passenger cars, 84 gasoline-fueled light-duty trucks, and 172 gasoline-fueled heavy-duty trucks.

As with the N₂O analyses, ICF compared the methane emissions on the Bag-2 mode with the corresponding methane emissions on the Hot-Running 505 (HR-505) mode. In the N₂O analyses, ICF found that the Bag-2 emissions (in units of grams per mile) were a better estimate of the HR-505 than using units of grams per hour; for the methane emissions, the comparisons were more mixed. However, the Bag-2 methane emissions in grams per mile, again, appeared to be the better surrogate for the HR-505 emissions.

For estimating methane emissions associated with the "generic" engine start (in grams per start), we simply subtracted the Bag-2 rate from the weighted FTP rate (both in grams per mile), multiplied that difference by the actual distance traveled, and rounded any negative estimates up to zero.

We then estimated the "running" methane emissions by first multiplying the FTP emissions (in grams per mile) by the actual distance traveled (to estimate total grams of methane). We then subtracted the calculated methane emissions associated with the engine start, and then divided that value by the elapse time of the driving cycle (i.e., 0.38 hours).

As with the N₂O estimates, these calculations were performed for each of the vehicle type/control technology groups used in EPA's Climate Leaders Greenhouse Gas Inventory Protocol, with results shown in Table 7-3. This grouping again required transformation to the model-year based MOVES bins according to the weighting of control technology and emission standard by model year. The model year-based rates are shown in Appendix F.

Table 7-3: Proposed Start and Running CH₄ Emissions By Control Technology

Vehicle Type	Control Technology	CH₄ Running (g/hr)	CH₄ Start (g/start)
Gasoline Passenger Cars (LDGV)			
	LEVs	0.17026	0.03189
	Tier 1	0.23927	0.05521
	Tier 0	1.20220	0.03425
	Oxid Catalyst	2.57799	0.00882
	Non-Catalyst	3.02017	0.05906
	Uncontrolled	3.17229	0.06203
Gasoline Light-Duty Trucks (LDGT)			
	LEVs	0.20721	0.04599
	Tier 1	0.45605	0.08235
	Tier 0	1.20728	0.07245
	Oxid Catalyst	2.54222	0.09948
	Non-Catalyst	3.41579	0.06680
	Uncontrolled	3.63215	0.07103
Gasoline Heavy-Duty Vehicles (HDGV)			
	LEVs	0.42320	0.09393
	Tier 1	0.47775	0.16309
	Tier 0	3.78625	0.18304
	Oxid Catalyst	3.50183	0.21549
	Non-Catalyst	7.50273	0.14672
	Uncontrolled	8.27757	0.16187
Diesel Passenger Cars (LDDV)			
	Advanced	0.01946	0
	Moderate	0.01989	0
	Uncontrolled	0.02358	0
Diesel Light Duty Trucks (LDDT)			
	Advanced	0.02713	0
	Moderate	0.02658	0
	Uncontrolled	0.03070	0
Diesel Heavy Duty Vehicles (HDDV)			
	Advanced	0.08112	0
	Moderate	0.08112	0
	Uncontrolled	0.08112	0
Motorcycles (MC)			
	Controlled	1.24339	0.02431
	Uncontrolled	1.66798	0.03262

7.5 Advanced Technologies & Alternative Fuels

The Emissions & Sinks report contains estimates of CH₄ and N₂O emission rates for five alternative fuels (i.e., methanol, ethanol, CNG, LNG, and LPG) for three vehicle classes (light-duty cars and trucks, heavy-duty trucks, and buses). Those estimated rates, reproduced from Table E-14 of Emissions & Sinks, are shown in Table 7-4:

Table 7-4: Emission Factors for CH₄ and N₂O for Alternative Fueled Vehicles

Vehicle Type	Fuel	N ₂ O (g/mi)	CH ₄ (g/mi)
Light-Duty Vehicles	Methanol	0.063	0.014
	CNG	0.113	0.914
	LPG	0.152	0.609
	Ethanol	0.076	0.043
Heavy-Duty Vehicles	Methanol	0.217	0.646
	CNG	0.297	9.629
	LNG	0.440	6.857
	LPG	0.150	0.108
	Ethanol	0.307	1.975
Buses	Methanol	0.217	0.646
	CNG	0.162	12.416
	Ethanol	0.364	2.079

To work with MOVES, these rates required separation into start and running emissions, and conversion to grams per hour (running) and grams per start (start). To split the start and running emissions, we referred to the comparable gasoline-fueled vehicles. We split the emissions of the alternative fueled vehicles so that the ratio of the start to running would be the same as for the corresponding gasoline-fueled vehicles. The final rates are shown in Appendix F.

Advanced technology rates for CH₄ and N₂O were developed by applying ratios to the conventional technology rates – the same method used for energy consumption. PERE doesn't model CH₄ and N₂O, and we weren't aware of any data on which to base ratios. We therefore developed ratios based on those used for energy consumption, under the assumption that the CH₄ : energy and N₂O : energy ratios would stay the same between conventional and advanced technologies. Start energy ratios were used directly for CH₄ and N₂O. Running ratios were more involved. Since the running energy ratios were broken down by operating mode, a composite ratio was needed to apply to the single running mode used for CH₄ and N₂O. These were calculated by weighting the running energy ratios by operating mode with the national default operating mode distribution generated by MOVES. The advanced technology ratios for CH₄ and N₂O

are shown in Table 7-5. These ratios are in the “ShortTermFERC” file used by the FERC, although the ratios for diesel advanced technologies are relative to gasoline conventional IC in this table and relative to diesel conventional IC in the FERC file.

**Table 7-5: Advanced Technology CH₄ and N₂O Ratios
(relative to Gasoline Conventional IC)**

Fuel / Technology	Running Ratio	Start Ratio
Gasoline Advanced IC	0.89	1.0
Gasoline CIC Hybrid Moderate	0.60	0.93
Gasoline CIC Hybrid Full	0.53	0.75
Gasoline AIC Hybrid Moderate	0.53	0.83
Gasoline AIC Hybrid Full	0.48	0.67
Diesel Fuel Conventional IC	0.64	0.77
Diesel Fuel Advanced IC	0.60	0.67
Diesel CIC Hybrid Moderate	0.38	0.65
Diesel CIC Hybrid Full	0.35	0.51
Diesel AIC Hybrid Moderate	0.36	0.61
Diesel AIC Hybrid Full	0.33	0.48
Electric	0	0

8. CO₂ Equivalent Calculation

MOVES2004 does not currently estimate CO₂ Equivalent emissions, but is planned for future releases using the methodology presented here. CO₂ equivalent is a combined measure of greenhouse gas emissions weighted according to the global warming potential of each gas, relative to CO₂. Although the mass emissions of CH₄ and N₂O are much smaller than CO₂, the global warming potential is higher, which increases the contribution of these gases to overall greenhouse effect. CO₂ equivalent is calculated from CO₂, N₂O and CH₄ mass emissions according to equation 8-1, meaning that there will be no direct CO₂ equivalent rates in MOVES.

$$\text{Eqn 8-1 } \text{CO}_2 \text{ Equivalent} = \text{CO}_2 * \text{GWP}_{\text{CO}_2} + \text{CH}_4 * \text{GWP}_{\text{CH}_4} + \text{N}_2\text{O} * \text{GWP}_{\text{N}_2\text{O}}$$

GWP is Global Warming Potential. The values used for this, shown in Table 8-1, are contained in the Pollutant table of the MOVES Default Database, and are taken from the Emission & Sinks report. The values used are for a 100-year time horizon, according to the guidelines of the United Nations Framework Convention on Climate Change (UNFCCC).²⁰

Table 8-1: 100-Year Global Warming Potentials

Pollutant	Global Warming Potential
CO ₂	1
CH ₄	21
N ₂ O	320

9. Adjustments

The modal binning approach eliminates the need for many of the correction factors used in MOBILE – for example speed correction factors and off-cycle correction factors. However, sufficient data doesn't exist to bin other important effects such as temperature, air conditioning and fuel effects. As in MOBILE, correction factors are necessary to account for these effects. This section discusses the development of correction factors in MOVES2004.

9.1 Temperature

Temperature effects were only generated for the start process for total energy only, to capture the effect of cold starts on energy consumption.

9.1.1 Adjustments for Gasoline Vehicles

9.1.1.1 Data Sources

The vast majority of the FTPs in the MSOD (i.e., 18,676 of 20,156) were performed within a narrow temperature range (68° to 86° Fahrenheit). This concentration of test results can result in the analyses being skewed. Therefore, we limited the analyses of temperature effects to the 580 vehicles (from the MSOD) that had been tested at both ambient temperatures outside the FTP temperature range as well as within the FTP temperature range. There were a total of 2,818 FTPs performed on these 580 vehicles.

9.1.1.2 Start Analysis

We used a regression analysis to fit the start fuel consumption (Bag 1 minus Bag 3) using a quadratic curve. We then repeated the analysis using a logarithmic fit. Both approaches fit the accepted physical model in which, as the ambient temperature increases: 1) the fuel consumption improves (i.e. decreases) and 2) the rate of improvement also decreases (i.e., the changes at colder temperatures are more dramatic than at higher temperatures).

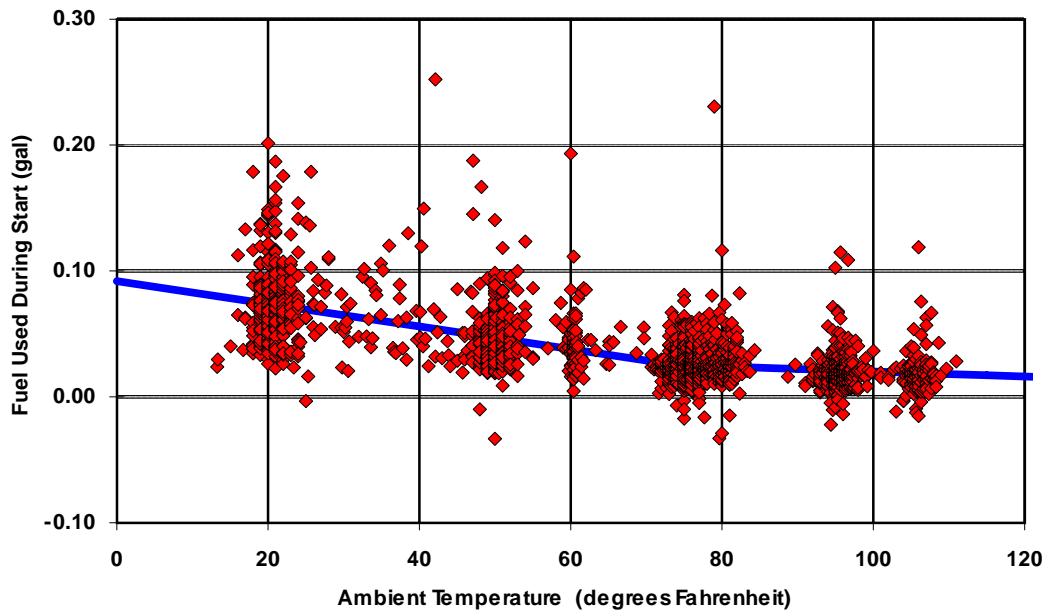
Mathematically, this suggests that we want an equation that is decreasing but "concave up" (at least between zero and 110° F). A relatively simple approach involves two linear equations with a common point (which turned out to be 73.9° F). The linear equations are:

$$\begin{aligned}\text{Eqn 9-1 Fuel Consumed During Engine Start (gallons)} &= \\ &= 0.0919 - (0.0009 * \text{Temperature}), \text{ where Temp} < 73.9^\circ \text{ F} \\ &= 0.0402 - (0.0002 * \text{Temperature}), \text{ where Temp} > 73.9^\circ \text{ F}\end{aligned}$$

This piecewise linear fit, which is illustrated in Figure 9-1, indicates improving fuel consumption (with increasing temperature) of 0.0009 gallons per each degree Fahrenheit

up to 73.9° F. Then, the improvement continues but at a reduced rate of 0.0002 gallons per degree above that temperature.

Figure 9-1: Comparison of Fuel Consumed During Engine Start versus Temperature for MSOD sample



In a series of recent testing programs, EPA's Office of Research and Development measured the emissions of vehicles with simulated malfunctions operating on a variety of test fuels. In these studies, nine gasoline-fueled, fuel injected, passenger cars were tested over the FTP at nominal temperatures of 75°, 20°, 0°, and -20° F. Additionally, five of those nine cars were also tested at an intermediate temperature of 40° F.

The fuel consumed during engine starts was calculated for all of the tests. Restricting the analyses to the tests using standard fuel and with no induced malfunctions, and then regressing the fuel consumed versus the ambient temperature, we obtained the following linear regression equations:

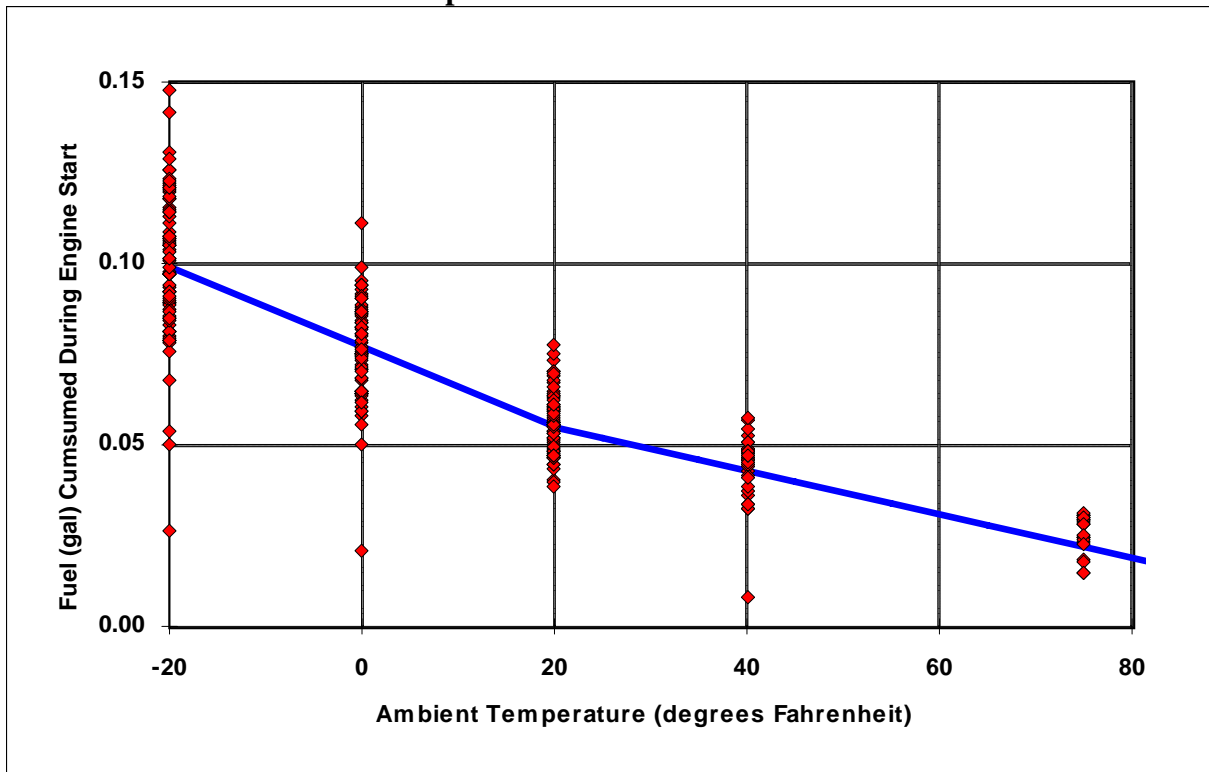
Eqn 9-2 Fuel Consumed During Engine Start (in gallons) =

$$= 0.0771 - (0.0011 * \text{Temperature}), \text{ where } \text{Temp} \leq 20^\circ \text{ F}$$

$$= 0.0669 - (0.0006 * \text{Temperature}), \text{ where } \text{Temp} \geq 20^\circ \text{ F}$$

These two line segments meet at a temperature of 20.4° F. This piecewise linear fit is illustrated in Figure 9-2.

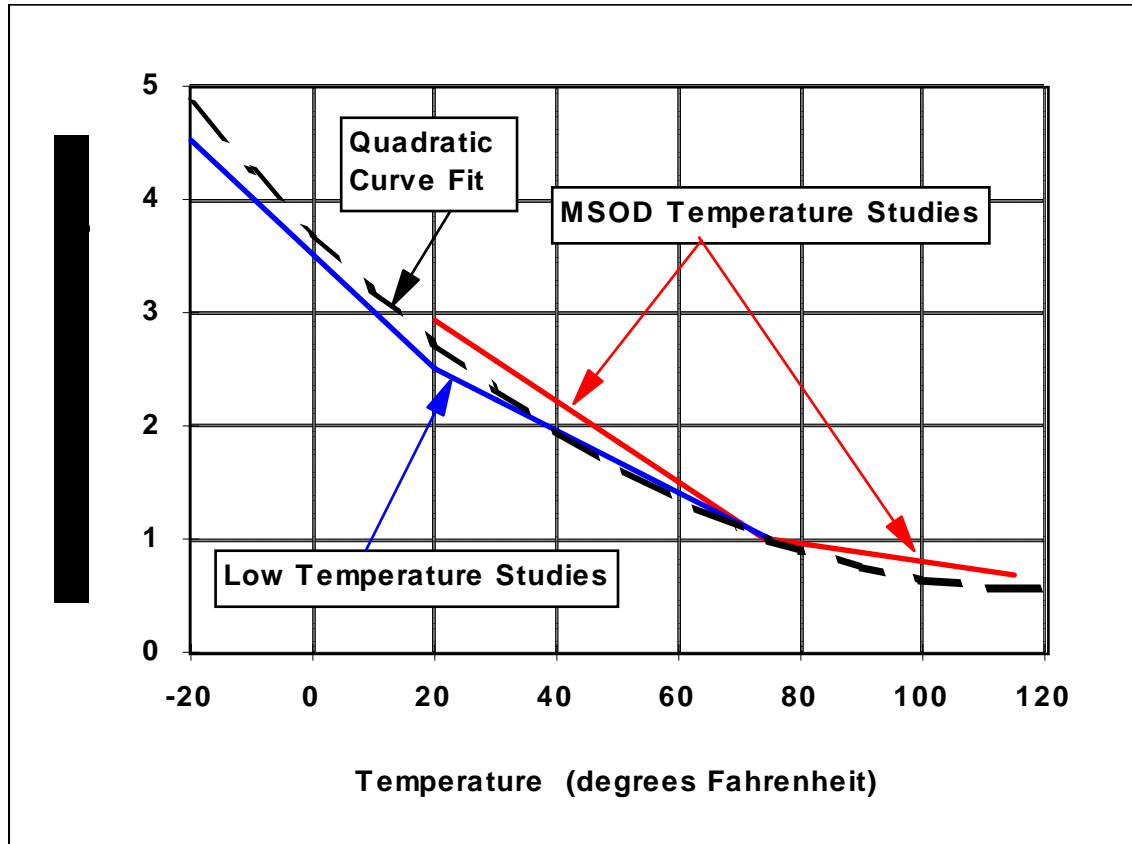
Figure 9-2: Comparison of Fuel Consumed During Engine Start versus Temperature for EPA/ORD Dataset



This piecewise linear fit indicates improving fuel consumption (with increasing temperature) of 0.0011 gallons per each degree Fahrenheit up to 20.4° F. Then, the improvement drops to 0.0006 gallons per degree from that temperature up to 75° F.

Using Equation 9-1 and 9-2, we can predict the cold-start fuel consumption (for each vehicle sample) over the temperature range from -20° through 120° Fahrenheit. Normalizing those predictions (so that the fuel consumption at 75° F is 1.0) for each of the two analyses produces multiplicative adjustment factors that are illustrated in Figure 9-3.

**Figure 9-3 Normalized Fuel Consumed During Engine Start versus Temperature
With Quadratic Curve Fit**



We then developed a single quadratic function that would approximate those piecewise linear segments. Rather than simply performing a regression analysis on the normalized test results, we first established the following two criteria that the quadratic function (which would become our temperature adjustment) must satisfy:

1. The quadratic function must have a value of 1.0 at the temperature that is the nominal value for the baseline FTPs (i.e., 75° F).
2. The engine-start fuel consumption must improve (i.e., decrease) with increasing temperatures, for all temperatures within the range of the test data (i.e., from -20° F through 115° F). And, as we approach the maximum temperature, the improvement in fuel consumption becomes very small.

The first criterion was met by using an equation of the form:

$$\text{Eqn 9-3 Factor} = A * (\text{Temp} - 75)^2 - B * (\text{Temp} - 75) + 1.0$$

The second criterion was met by selecting the vertex of the curve (parabola) to occur at the temperature of 120° F. This yielded the following relationship between those two coefficients:

$$\text{Eqn 9-4 } B = 90 * A$$

The resulting quadratic curve that approximates the piecewise linear segments (pictured as a dotted curve in the preceding figure) has as its equation:

$$\text{Eqn 9-5 Factor} = 0.000219 * (\text{Temp} - 75)^2 - 0.01971 * (\text{Temp} - 75) + 1.0$$

Equation 9-5 was used in MOVES2004 to estimate the effects of engine-start energy consumption of gasoline-fueled vehicles. The coefficients shown above are stored directly in the table TemperatureAdjust, as TempAdjustTermA (-0.01971) and TempAdjustTermB (0.000219). A placeholder exists in the database for TermAdjustTermC, but is not currently used.

9.1.1.3 Running Analysis

Subtracting the fuel consumed during engine starts from the FTP results leaves the fuel consumed over a hot-running LA-4 driving cycle. Analyzing the calculated fuel consumed over that hot-running LA-4 for the 2,818 FTPs performed on the 580 vehicles (from the MSOD) that had been tested at both ambient temperatures outside the FTP temperature range as well as within the FTP temperature range, we found virtually no change in fuel consumption. Therefore, we are using a temperature adjustment factor of 1.0 for fuel consumed during "running" operations.

9.1.2 Adjustments for Diesel Vehicles

The difference in the process of vaporizing the fuel between gasoline-fueled vehicles and diesel-fueled vehicles suggests that the diesel-fueled vehicles would be less sensitive to the ambient temperature than would the gasoline-fueled vehicles (relative to the amount of fuel consumed during engine start). To test this hypothesis, we examined the database of 20,156 FTPs and identified 93 FTPs performed on 66 diesel-fueled cars and trucks. Repeating the regression analysis from Section 2.2, we obtained the results in Table 9-1 (on this much smaller subset of the test data).

**Table 9-1: Regression of Start (Bag-1 minus Bag-3) Fuel Consumption
(Diesel-Fueled Cars and Trucks)**

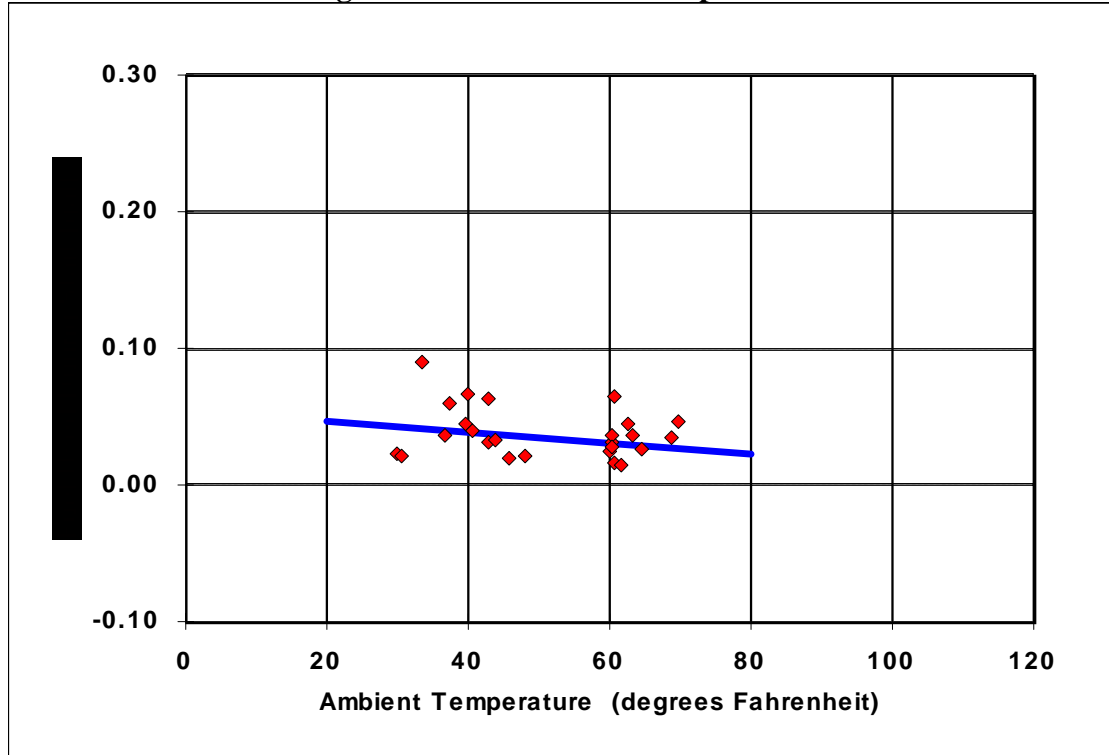
Dependent variable is:			Start (Bag-1 minus Bag-3)	
No Selector				
R squared = 13.8% R squared (adjusted) = 11.8%				
s = 0.0180 with 93 - 3 = 90 degrees of freedom				
Source	Sum of Squares	df	Mean Square	F-ratio
Regression	0.004667	2	0.002334	7.17
Residual	0.029274	90	0.000325	
Variable	Coefficient	s.e. of Coeff	t-ratio	prob
Constant	0.055671	0.0103	5.38	≤ 0.0001
Avg_Temp	-0.000536	0.0002	-3.43	0.0009
Disp (CID)	0.000043	0.0000	2.64	0.0099

Comparing these two regressions to the previous analyses suggests that for the diesel-fueled vehicles, the fuel consumption during engine start is substantially less sensitive to changes in ambient temperature than are the corresponding gasoline-fueled vehicles (at least within the range of the test data, between 29° and 70° F). To determine the magnitude of the temperature effect on diesel fuel consumption (during engine starts), we paralleled the approach used with the larger (gasoline-fueled) sample. That is, we first identified (within this 66-vehicle sample), 12 diesel-fueled vehicles that were tested (25 FTPs) at multiple temperatures (29° to 70° F). Then performing a regression analysis on these tests produces Equation 9-6 (note: the coefficients are different than Table 9-1 because this is a different sample than that used to generate Table 9-1):

Eqn 9-6: Fuel Consumed During Engine Start (gallons) = 0.0564 - (0.0004 * Temperature)

This is illustrated by Figure 9-4.

Figure 9-4: Comparison of Fuel Consumed During Engine Start At Various Temperatures



This regression equation indicates improved fuel consumption (with increasing temperature) of 0.0004 gallons per degree within the temperature range of the test data (29° to 70° F). While the regression equation in Section 2.3 indicates improved fuel consumption (with increasing temperature) of 0.0009 gallons per degree for that same temperature range. This suggests that the change in fuel consumption during engine starts resulting from change in the ambient temperature for diesel-fueled vehicles is about 0.44 (i.e., 0.0004 / 0.0009) times the corresponding change in the fuel consumption for gasoline-fueled vehicles.

If the magnitude of the effect of ambient temperature on the engine-start fuel consumption of diesels is 44 percent of the corresponding effect on gasoline-fueled engines, then the temperature adjustment factor for diesel vehicles is derived by multiplying the gasoline temperature adjustment coefficients from Equation 9-5 by 0.44, resulting in Equation 9-7:

$$\text{Eqn 9-7 Factor} = 0.000096 * (\text{Temp} - 75)^2 - 0.00867 * (\text{Temp} - 75) + 1.0$$

MOVES2004 employs these coefficients (in the TemperatureAdjust table) to adjust, for varying ambient temperatures, the engine-start fuel consumption of diesel-fueled vehicles. As with the gasoline-fueled vehicles, we propose a temperature adjustment factor of 1.0 for fuel consumed during "running" operations for the diesel-fueled vehicles.

9.2 Air Conditioning

MOVES2004 accounts for increased energy consumption due to air conditioning usage through an air conditioning adjustment factor. The air conditioning adjustment is a combined measure of air conditioning activity (the fraction of time the air conditioning compressor is engaged) and the increase in energy consumption when the compressor is engaged. This approach was first established in MOBILE6.^{21,22} In MOVES2004, air conditioning adjustments are applied only to running total energy. Analysis performed for MOBILE6 indicates that start energy and emissions rates are not highly impacted by A/C usage, and there is not sufficient data of the effects of A/C usage on CH₄ and N₂O emissions to merit an adjustment for these pollutants.

A/C adjustment factors stored in the MOVES Default database energy usage reflect “full-usage”, e.g. when the compressor is engaged. Within MOVES2004 these are scaled back according to the fraction of time the compressor is engaged, the penetration of A/C in the fleet, and an estimate of malfunctioning systems. The latter three components are documented in the Fleet & Activity report.

Full-usage energy adjustment factors are stored in the FullACAdjustment table in MOVES Default. The are broken down by the 17 operating mode bins discussed in Section 4.2 (Table 4-3). The adjustments were broken down by operating mode because test data shows that the relative load increase caused by the A/C compressor varies significantly depending on engine load – at low loads the relative increase is high, at high loads the relative increase is low.

9.2.1 Data Sources

A subset of data from the MSOD contains a sample of second-by-second data on a variety of test schedules tested with A/C on and off. The majority of these data are from a program EPA conducted when gathering data from MOBILE6, which ran 38 LDVs and LDTs over several drive cycles at high A/C load conditions. Additional details of this testing can be found in MOBILE6 documentation.

9.2.2 Analysis Methodology

Using the subset of vehicles with paired A/C on and off tests, the binner program generated energy consumption rates with and without the A/C on, grouped by the source bins for which the test vehicles fell into. The ratio of A/C-on to A/C-off was calculated for each operating mode as the FullACAdjustment, by source bin. Since the source bin coverage of the 38-vehicle dataset wasn't broad enough to merit keeping the A/C adjustment split by source bin, we developed a composite adjustment as a weighted average across source bins, using the source bin fraction (i.e. relative occurrence of each source bin in the fleet) as a weighting factor. The resulting adjustments are shown in Table 9-2.

Table 9-2 Full A/C adjustments by operating mode bin

Braking: 1.34			
Idle: 1.36			
VSP \ Speed	0-25mph	25-50	>50
< 0 kw/tonne	1.31	1.29	
0 to 3	1.25	1.22	
3 to 6	1.19	1.19	
6 to 9	1.17	1.17	
9 to 12	1.15	1.16	
12 and greater	1.13	1.13	1.20
6 to 12			1.16
< 6			1.14

Since paired A/C on and off data is very limited, we decided to calculate a single set of adjustment and apply them to all source use types in the model. This means light-duty A/C effects would be applied to all vehicle classes, including heavy-duty.

10. Well-To-Pump Energy & Emission Rates

An updated version of the GREET model, developed by Argonne National Laboratory, has been integrated into MOVES2004 to provide energy and emission rates for the well-to-pump process. A database table named GREETWellToPump in the MOVES Default database contains default well-to-pump rates generated by the integrated version of GREET; the user only needs to run GREET if a change from default well-to-pump assumptions is desired. These rates are expressed in terms of well-to-pump energy use (or emissions for CH₄ and N₂O) per unit pump-to-wheel energy consumption – i.e. well-to-pump KJ per pump-to-wheel KJ, or well-to-pump grams per pump-to-wheel KJ.

A main feature of the updated GREET model is the ability to generate these rates by calendar year, taking into account changes in fuel production efficiencies or pathway mixes over time. The rates used in MOVES therefore vary by fuel subtype and calendar year. Rates by fuel subtype for two calendar years, 1999 and 2020, are shown in Tables 10-1 and 10-2. The full set of calendar year rates can be accessed by querying the GREETWellToPump table in the MOVES Default database. Because the energy rates in Table 10-1 are expressed as KJ upstream per KJ downstream, the numbers are simply well-to-pump energy consumption expressed as the fraction of pump-to-wheel energy consumption. Results for the hydrogen pathways are under development and will be added to the model when they are ready.

Table 10-1:
Default Well-To-Pump Energy Rates for Calendar Year 1999 & 2020
(Well-To-Pump KJ / Pump-To-Wheel KJ)

Fuel Subtype	Total Energy		Petroleum Energy		Fossil Energy	
	1999	2020	1999	2020	1999	2020
Conventional Gasoline	0.229	0.236	0.107	0.110	0.225	0.232
Reformulated Gasoline	0.266	0.270	0.105	0.108	0.262	0.266
E10	0.264	0.268	0.105	0.109	0.259	0.265
Conventional Diesel	0.179	0.213	0.082	0.099	0.176	0.210
Biodiesel (BD20)	0.275	0.291	0.101	0.111	0.269	0.286
Fisher-Tropsch (FT100)	0.725	0.643	0.020	0.020	0.724	0.643
CNG	0.151	0.146	0.008	0.006	0.140	0.137
LPG	0.114	0.111	0.030	0.025	0.113	0.110
Ethanol (E85)	0.631	0.613	0.088	0.088	0.620	0.604
Methanol (M85)	0.520	0.452	0.055	0.055	0.519	0.451
Electricity	1.654	1.499	0.095	0.067	1.624	1.476

gas and diesel subtypes account for phase-in of low sulfur fuel in required years

Table 10-2:
Default Well-To-Pump Emission Rates for Calendar Year 1999 & 2020
(Well-To-Pump Grams / Pump-To-Wheel KJ)

Fuel Subtype	CO ₂		CH ₄		N ₂ O	
	1999	2020	1999	2020	1999	2020
Conventional Gasoline	0.017	0.017	1.00E-4	1.01E-4	2.77E-7	2.91E-7
Reformulated Gasoline	0.018	0.018	1.16E-4	1.16E-4	2.95E-7	3.08E-7
E10	0.015	0.015	1.01E-4	1.02E-4	3.68E-6	3.48E-6
Conventional Diesel	0.014	0.016	9.71E-5	9.99E-5	2.24E-7	2.66E-7
Biodiesel (BD20)	0.004	0.006	9.51E-5	9.61E-5	1.73E-6	1.60E-6
Fisher-Tropsch (FT100)	0.026	0.022	1.16E-4	1.16E-4	9.51E-8	9.56E-8
CNG	0.011	0.010	2.35E-4	2.35E-4	1.59E-7	1.72E-7
LPG	0.008	0.008	1.08E-4	1.11E-4	1.44E-7	1.41E-7
Ethanol (E85)	-0.006	-0.009	1.08E-4	1.07E-4	3.91E-5	3.67E-5
Methanol (M85)	0.022	0.018	1.37E-4	1.35E-4	3.76E-7	3.68E-7
Electricity	0.203	0.189	2.81E-4	2.55E-4	2.69E-6	3.08E-6

These rates are a product of the default GREET inputs for each fuel pathway, including variables such as gasoline oxygenate content and natural gas export share – there are in total several hundred inputs that are too numerous to list here, but can be found in separate documentation. A report to be included with MOVES2004 technical documentation entitled “GREET Users Manual and Technical Issues for MOVES Integration”, prepared by Argonne Labs, documents many of the “top-level” default inputs. More extensive documentation of underlying GREET methodologies can be found on the GREET website (available online at <http://www.transportation.anl.gov/software/GREET/index.html>)

11. Data Source Identification

From the preceding sections it should be clear that emission rates were generated from several different approaches: direct binning of test data, a variety of hole-filling methods, and estimation of rates based on ratios of advanced technology to conventional technology performance. To provide a means of keeping track of how a specific emission rate in the default database was generated, we added the field “data source ID” to each record of the emission rate table. DataSourceID is a meta-data field which tells the user which method was used to generate the rate found in that record, distinguishing among the many methods discussed in this report. The full description of data source IDs are contained in the DataSourceID table in the MOVES default database, and are also shown in Tables 11-1 and 11-2.

Table 11-1: Data Source ID descriptions

ID	Description
1001	Data: Binned from second-by-second data
1002	Data: Calculated from bag data
2001	Hole Filling: Manual
3001	Hole Filling: Physical Emission Rate Estimator
4011-4242	Hole Filling: Interpolator / Copier Program (see Table 11-2)
5001	Hole filling: Pre-2001 Alt Fuel & Adv Tech script
5002	Hole filling: 2001-2010 Diesel Conv IC script
5003	Hole filling: Extended Idle Calculation script
6000	Future Emission Rate Creator

Table 11-2: Data Source ID descriptions for Interpolator / Copier Program

ID	Description
4011	all opmodes except idle interpolate by weight class in first pass
4012	all opmodes except idle interpolate by weight class in second pass
4021	all opmodes except idle interpolate by displacement class in first pass
4022	all opmodes except idle interpolate by displacement class in second pass
4031	all opmodes except idle interpolate by model-year group in first pass>
4032	all opmodes except idle interpolate by model-year group in second pass>
4041	all opmodes except idle adopt meanbaserate from neighboring sourcebin by model-year group in first pass
4042	all opmodes except idle adopt meanbaserate from neighboring sourcebin by model-year group in second pass
4051	all opmodes except idle adopt meanbaserate from neighboring sourcebin by weight class in first pass
4052	all opmodes except idle adopt meanbaserate from neighboring sourcebin by weight class in second pass
4061	all opmodes except idle adopt meanbaserate from neighboring sourcebin by displacement class in first pass
4062	all opmodes except idle adopt meanbaserate from neighboring sourcebin by displacement class in second pass
4111	idle opmode interpolate by displacement class in first pass
4112	idle opmode interpolate by displacement class in second pass
4121	idle opmode interpolate by weight class in first pass
4122	idle opmode interpolate by weight class in second pass
4131	idle opmode interpolate by model-year group in first pass>
4132	idle opmode interpolate by model-year group in second pass>
4141	idle opmode adopt meanbaserate from neighboring sourcebin by model-year group in first pass
4142	idle opmode adopt meanbaserate from neighboring sourcebin by displacement class in first pass
4151	idle opmode, adopt meanbaserate from neighboring sourcebin, by displacement class, in first pass
4152	idle opmode adopt meanbaserate from neighboring sourcebin by displacement class in second pass
4161	idle opmode adopt meanbaserate from neighboring sourcebin by weight class in first pass
4162	idle opmode adopt meanbaserate from neighboring sourcebin by weight class in second pass
4241	idle opmode adopt meanbaserate from neighboring sourcebin by model-year group in second pass
4242	idle opmode adopt meanbaserate from neighboring sourcebin by displacement class in second pass

12. Acknowledgments

The authors gratefully acknowledge the following individuals for their ideas and support towards the work presented in this report: Chad Bailey, Connie Hart and David Brzezinski of EPA's Office of Transportation and Air Quality; Michael Wang, Yu We, Dan Santini, Feng An, Anant Vias and Amgad Elgowainy of Argonne National Laboratory; Matt Barth, Ted Younglove and George Scora of UC Riverside; Chris Frey of North Carolina State University; and Nigel Clark and Ralph Nines of West Virginia University.

Appendix A: Binning Methodology Proof-Of-Concept

A.1 Background

The section lays out the methodology for determining the 17 operating mode bins used for running total energy consumption. This analysis builds on previous studies conducted in the preliminary stages of MOVES, including 1) the on-board emission “shootout” involving multiple participants,²³ 2) the proof-of-concept modal binning analysis conducted by North Carolina State University (NCSU),²⁴ and 3) follow-on investigation of the binning approach by EPA.²⁵ The reader should consult the referenced reports for a full explanation of these findings. The conclusions drawn from these studies which form the basis of the work presented in this section are:

- The process of binning – i.e. calculating average mass per time emission rates from second-by-second data within a pre-defined range of vehicle operation – is an effective means of characterizing emission changes from vehicle operational changes.
- Vehicle Specific Power (VSP) is an effective parameter for binning, but does result in some bias at lower speeds if not supplemented with additional variables.
- Average cycle speed was initially proposed as a supplemental binning variable, with some success. However, a drawback of this approach is not having the ability to define bins based on “real-time” vehicle operation.

A.2 Investigation of Alternatives: Engine Friction and RPM

Additional analysis focused on the role of engine friction in generating the bias which was appearing with the VSP-only approach. The effect of engine friction on fuel consumption is relatively large at low loads. To account for this we investigated modifying the VSP equation to explicitly include engine friction. While showing some promise, this approach was ultimately scrapped because it wasn’t feasible to determine an appropriate value (or algorithm) for the additional engine friction “gamma” term which could apply across vehicles.²⁶

Engine friction is directly proportional to RPM, so the idea of binning by VSP and RPM was considered. It was rejected because RPM is not directly available in many datasets, and would need to be derived from vehicle speed and estimates of transmission logic and gear ratio. To do this on a vehicle-to-vehicle basis would not be feasible.

A.3 VSP/Speed Binning Proof-Of-Concept

Instantaneous vehicle speed was chosen as the second binning parameter. Speed serves as a reasonable surrogate for engine friction and is readily available across the range of datasets which would be considered for MOVES. NCSU’s analysis of modal binning approaches showed that speed was an important binning variable, and that model performance would be improved by using it in conjunction with VSP.

The primary objective of this analysis was to develop bin definitions using VSP and instantaneous speed, which could be applied across all source use types and pollutants, and which improved on VSP-only approaches. The VSP-only approach used as the basis for comparison was a 14 VSP bin approach developed by NCSU and documented in their modal binning analysis report. Another objective of the analysis, in response to comment received on the proposed methodology (Appendix G), was to show the feasibility of the binning approach on a vehicle sample which included heavy-duty vehicles and high-emitting light-duty vehicles, as the preliminary binning analyses focused on later model light-duty vehicles.

A.3.1 Data Sources

The datasets used for the light-duty analysis was compiled from 26 LDVs and LDTs tested by ARB over their Unified Cycle Correction (UCC) cycles, and 11 LDVs tested on-road by EPA as part of the “shootout” dataset. For the latter, only the segments of testing run under warmed-up conditions with the A/C off were used. The model year range for the ARB dataset was from 1983 – 1998, and included high emitting vehicles in the sample. High emitters were purposely included to address the question of whether the modal binning approach would work on high as well as low emitting vehicles. The EPA dataset consisted of 1996 and newer vehicle. The combined dataset consisted of 253 unique “trips”.

The dataset used for the heavy-duty analysis was compiled from 11 Class 8 trucks tested by CE-CERT on-road using their heavy-duty trailer configuration, over a variety of cycles, and EPA shootout data on 15 Transit Buses. These data were separated into 64 “trips”.

A.3.2 Bin Options Evaluated

Five binning approaches were developed for the analysis, termed Bin Option 1 through 5 (BO1 through BO5). The approaches are shown in Table A-1. Each of the options defined idle and “braking” as distinct modes; this approach was initially proposed as part of NCSU’s proof-of-concept work. We did change “decel” to “braking” in order to segregate actual braking operation, which is useful for modeling hybrid vehicles with regenerative braking. Whereas the NCSU work defined one mode (bin) each for cruise and accel, we expanded cruise and accel operation into subdivisions of VSP and speed.

Table A-1: Bin Approaches Analyzed

Braking (common)	Accel < -2 mph/s, or < -1 mph/s for 3 consecutive seconds			
Idle (common)	Speed = 0 (used between -1 and 1 mph to account for signal noise)			
Cruise/Accel	VSP Bin Ranges (kw/ton)		Instantaneous Speed Ranges (mph)	
Bin Option 1	≤ 2 2-6 6-10 10-14 ≥ 14		0-15 15-45 45-60 ≥ 60	
Bin Option 2	≤ 1 1-4 4-7 7-10 10-14 14-20 ≥ 20		0-40 40-55 ≥ 55	
Bin Option 3	≤ 1 1-4 4-7 7-10 10-13 13-16 16-19 ≥ 19		0-40 ≥ 40	
Bin Option 4	≤ 1 1-4 4-7 7-10 10-13 ≥ 13		0-30 ≥ 30	
Bin Option 5	< 50 mph ≤ 0 0-3 3-6 6-9 9-12 ≥ 12	≥ 50 mph ≤ 6 6-12 ≥ 12	0-25 25-50	≥ 50

The determination of bin options began with performing Hierarchical Tree-Based Regression (HTBR) on the analysis dataset, using the open-source statistical software package R. Separate HTBR runs were performed on the light-duty and heavy-duty sets, and by pollutant. Since the goal of the analysis was to come up with a common set of bin definitions across vehicle class and pollutant, the bin approaches were composites of the HTBR runs. In general speed cutpoints were above 40 mph. VSP cutpoints were not as cleanly defined, so the VSP points used were picked arbitrarily. BO1 was an initial attempt at a composite of the HTBR results. BO2 refined BO1 to more evenly distribute time/emissions for both light-duty and heavy-duty, add higher VSP bin to improve CO prediction. BO3, 4 and 5 were based on bins which could be filled by IM240 alone: BO3 was the initial cut, BO4 a refinement to improve distribution of time (i.e. making sure bins would be adequately filled by the IM240), and BO5 adds higher speed bins to

improve fuel prediction and reshuffles VSP and speed cutpoints to maintain distribution of time in each bin.

The latter three bin approaches focused on the IM240 because (as detailed in Section 4.3.1) much of the second-by-second data for light-duty vehicles used in MOVES2004 was from the New York Instrumentation Protocol Assessment (NYIPA) program, which included over 10,000 lab-grade IM240 tests. Since this program has the most vehicle coverage by far of any of the MOVES2004 data sources, allowing the bins to be completely filled with IM240 would improve coverage across the source bins. This approach also leaves the door open for using IM240 program data directly, which could make emitter distributions for criteria pollutants more straightforward.

A.3.3 VSP Calculation

VSP was calculated according to Equation A-1. Road load coefficients for light-duty were derived from the 50 mph road load horsepower values reported for each vehicle. The coefficients for heavy-duty were derived from Petrushov.²⁷ The weight of the CE-CERT trailer test rig, as estimated by CE-CERT staff (44,000 lbs), was added to the reported cab weight.

Equation A-1:
$$\text{VSP} = (A * \text{Speed} + B * \text{Speed}^2 + C * \text{Speed}^3 + \text{Mass} * \text{Speed} * \text{Accel}) / \text{Mass}$$

Where:

VSP is in KW/Metric Ton
Speed is in meters/second (mps)
Accel is in meters/second²
A is rolling resistance term in KW / mps
B is friction term in KW / mps²
C is aerodynamic drag term in KW / mps³
Mass is in metric tons (1000 kg)

A.3.4 Source Bin Assignment

The analysis followed the same steps MOVES would in modeling an aggregate estimate of fuel consumption and emissions. The first step was to assign the vehicles in the analysis set to source bins. For fuel consumption, the MOVES2004 source bin categories for energy consumption (Table 4-1) were used to assign each vehicle to a bin. For HC, CO, and NO_x source bins were defined based on categories of pollutant/CO₂ ratio by trip. The criteria for defining a bin were that the range of emission points within any bin should not be more than 25 percent of the range of the entire sample. We originally looked at defining categories based on vehicle emissions (e.g. emitter categories). Large trip-to-trip variability reduced the effectiveness of this approach.

A.3.5 Analysis Methodology and Results

A random sample of trips selected out of from light-duty and heavy-duty analysis sets (validation sample). For light-duty, 20 percent of the trips were chosen (64 trips); for heavy-duty, 30 percent of trips (23 trips). Binned fuel consumption and emission rates were produced by averaging all one-second values in each source bin and operating mode bin for the remaining trips (prediction sample). Fuel consumption and emission rates predicted for each second in the validation sample using the fuel consumption and emission rates from predicted sample, for the source bin and operating mode bin in that second. Fuel consumption and emissions aggregated over each trip in the validation sample to generate predicted and observed trip-average gram per second fuel consumption and emission rates.

Since MOVES will predict aggregate emissions produced by several vehicles over a period of time (e.g. one hour), the primary evaluation metric for this analysis is the magnitude of difference between the average of the observed and predicted fuel/emissions over all trips in the validation sample (sample average). Table A-2 shows the results of this comparison for each of the alternatives evaluated, in terms of percent difference from observed for the average per-trip emissions. These comparisons are shown for a) all trips in the sample, b) all trips with an average speed < 30 mph, and c) all trips with an average speed > 30 mph. The purpose of the latter two comparisons is to identify biases for lower speed or higher speed cycles; in some cases very good agreement in the “all trip” case can mask large differences in each speed range. The VSP row is for the 14 VSP bin approach developed by NCSU.

Table A-2: Sample Average Results (percent difference observed vs. predicted)

All Trips								
Light-Duty					Heavy-Duty			
	Fuel	HC	CO	NOx	Fuel	HC	CO	NOx
VSP	9%	1%	6%	5%	1%	15%	13%	-3%
BO1	4%	2%	4%	-2%	-3%	9%	15%	-5%
BO2	4%	2%	5%	-3%	-2%	10%	16%	-5%
BO3	5%	1%	6%	-1%				
BO4	6%	0%	3%	1%				
BO5	4%	1%	3%	-3%	-1%	10%	14%	-4%
Trips w/ Average Speed < 30								
Light-Duty					Heavy-Duty			
	Fuel	HC	CO	NOx	Fuel	HC	CO	NOx
VSP	22%	14%	8%	14%	10%	36%	25%	19%
BO1	8%	6%	5%	-8%	1%	25%	20%	5%
BO2	7%	6%	5%	-8%	1%	25%	18%	5%
BO3	10%	9%	5%	-5%				
BO4	14%	9%	4%	0%				
BO5	8%	6%	4%	-7%	0%	23%	21%	7%
Trips w/ Average Speed > 30								
Light-Duty					Heavy-Duty			
	Fuel	HC	CO	NOx	Fuel	HC	CO	NOx
VSP	-1%	-6%	6%	1%	-6%	-15%	-4%	-16%
BO1	2%	-1%	4%	0%	-6%	-12%	7%	-8%
BO2	2%	0%	6%	0%	-4%	-11%	13%	-11%
BO3	2%	-3%	6%	0%				
BO4	0%	-6%	3%	1%				
BO5	1%	-2%	3%	-1%	-2%	-9%	3%	-11%

In general adding speed bins improves prediction relative to the VSP-only approach, based on the results by speed range. This is particular the case for fuel consumption. For light-duty, BO1, BO2 and BO5 predict average trip fuel consumption within 10 percent over all trips and for both speed ranges. For heavy-duty, BO1, BO2 and BO5 generally predicts average trip fuel consumption and emission within 15 percent over all trips and for high-speed trips; HC and CO predict high on lower speed trips. Overall, BO2 and BO5 performed the best.

We also looked at the distribution of time and emissions in each bin. Tables A-3 and 4 show the distribution of time and emissions in the analysis datasets for two bin options, BO2 and BO5 (plus the distribution of time on the IM240 cycle for BO5). This comparison shows that the distribution of time and emissions is more balanced BO2. Basing bin definitions on the IM240 results in unbalanced bin distributions for heavy-duty, where the majority of operation occurs above 50 mph. However, BO5 prediction results were not hurt by this in this analysis.

Table A-3: Distribution of Time, Fuel Consumption and Emissions by Bin for BO2

Bin (Spd/VSP)	Light-Duty					Heavy-Duty				
	Time	Fuel	HC	CO	NOx	Time	Fuel	HC	CO	NOx
Brake	12.3%	6.4%	8.3%	5.8%	6.6%	5.3%	0.5%	1.1%	1.4%	0.9%
Idle	11.5%	4.6%	5.8%	2.6%	2.5%	18.8%	2.9%	3.4%	5.9%	3.5%
<40/<1	12.0%	6.4%	7.7%	5.7%	5.3%	17.3%	4.1%	6.8%	9.7%	4.4%
40-55/<1	3.4%	2.5%	2.8%	2.1%	2.5%	2.4%	1.7%	3.7%	2.2%	1.9%
>55/<1	2.4%	3.0%	2.9%	3.5%	4.1%	5.8%	3.7%	6.6%	3.5%	3.8%
<40/1-4	7.8%	5.7%	5.9%	4.7%	4.3%	6.7%	5.4%	3.1%	10.5%	4.9%
40-55/1-4	2.7%	2.4%	2.1%	1.6%	2.2%	4.0%	5.8%	9.6%	4.3%	6.3%
>55/1-4	2.6%	3.5%	2.8%	3.0%	4.6%	4.8%	6.7%	9.1%	4.2%	7.3%
<40/4-7	6.1%	5.7%	5.5%	4.7%	4.0%	4.6%	6.0%	2.0%	10.9%	4.8%
40-55/4-7	2.9%	2.9%	2.5%	2.0%	2.6%	4.6%	8.6%	12.3%	6.3%	9.4%
>55/4-7	4.2%	6.2%	4.6%	4.7%	8.1%	6.5%	11.1%	12.4%	6.4%	12.4%
<40/7-10	4.6%	5.2%	5.5%	4.9%	3.6%	3.3%	5.8%	1.4%	10.4%	4.1%
40-55/7-10	2.2%	2.6%	2.4%	2.3%	2.3%	2.6%	6.0%	7.2%	4.6%	6.1%
>55/7-10	5.0%	8.2%	6.2%	6.9%	11.6%	4.2%	8.6%	10.0%	4.9%	8.9%
<40/10-14	3.9%	5.2%	5.6%	5.5%	3.5%	1.9%	3.7%	0.8%	4.4%	2.6%
40-55/10-14	1.9%	2.5%	2.5%	2.7%	2.1%	1.3%	3.3%	2.8%	2.4%	3.3%
>55/10-14	5.2%	9.4%	7.7%	10.1%	13.4%	2.6%	6.1%	4.3%	3.1%	6.0%
<40/14-20	2.4%	3.9%	4.9%	5.4%	2.6%	0.5%	1.2%	0.2%	0.9%	0.8%
40-55/14-20	1.0%	1.6%	2.0%	2.4%	1.4%	0.7%	2.3%	0.8%	1.0%	2.3%
>55/14-20	3.3%	6.3%	5.8%	8.5%	7.3%	1.4%	4.1%	1.6%	1.8%	3.9%
<40/>20	0.7%	1.3%	2.2%	2.8%	0.8%	0.0%	0.1%	0.0%	0.1%	0.1%
40-55/>20	0.4%	0.9%	1.2%	2.4%	0.9%	0.2%	0.8%	0.3%	0.4%	0.8%
>55/>20	1.7%	3.6%	3.1%	5.9%	3.8%	0.5%	1.6%	0.6%	0.7%	1.5%

Table A-4: Distribution of Time, Fuel Consumption and Emissions by Bin for BO5

Bin (Spd/VSP)	Light-Duty						Heavy-Duty				
	Time	Time (IM240)	Fuel	HC	CO	NOX	Time	Fuel	HC	CO	NOX
Brake	12.3%	13.3%	6.4%	8.3%	5.8%	6.6%	5.3%	0.5%	1.1%	1.4%	0.9%
Idle	11.5%	4.6%	4.6%	5.8%	2.6%	2.5%	18.8%	2.9%	3.4%	5.9%	3.5%
<25/<0	3.9%	6.7%	1.8%	2.1%	1.5%	1.2%	6.6%	1.0%	2.8%	2.9%	1.2%
25-50/<0	6.2%	3.8%	3.6%	4.7%	3.4%	3.8%	4.3%	1.0%	1.5%	2.3%	1.3%
<25/0-3	5.5%	9.2%	3.5%	3.7%	2.7%	2.2%	9.7%	4.6%	4.4%	10.0%	4.1%
25-50/0-3	5.7%	4.2%	4.0%	4.2%	3.3%	3.7%	3.8%	2.9%	2.5%	3.3%	3.2%
<25/3-6	2.8%	5.4%	2.5%	2.3%	1.9%	1.4%	2.6%	3.5%	1.3%	8.1%	2.4%
25-50/3-6	5.8%	13.8%	5.1%	5.0%	4.0%	4.4%	4.0%	4.5%	3.3%	4.2%	4.6%
<25/6-9	2.3%	3.3%	2.5%	2.4%	1.9%	1.3%	2.0%	3.7%	0.8%	8.8%	2.3%
25-50/6-9	4.5%	2.1%	4.7%	4.9%	4.4%	4.1%	2.8%	4.5%	2.4%	4.0%	4.1%
<25/9-12	1.7%	4.6%	2.1%	2.1%	1.9%	1.1%	1.0%	2.1%	0.4%	4.2%	1.2%
25-50/9-12	3.1%	1.7%	3.8%	4.1%	4.0%	3.1%	1.6%	3.2%	1.4%	2.7%	2.8%
<25/>12	1.8%	2.5%	2.8%	3.1%	3.2%	1.4%	0.4%	0.9%	0.1%	1.1%	0.5%
25-50/>12	4.3%	6.3%	7.2%	9.4%	11.4%	5.8%	1.5%	3.7%	1.0%	2.2%	3.2%
>50/<6	7.9%	5.8%	15.3%	13.9%	21.3%	18.0%	3.6%	10.3%	4.7%	5.0%	10.1%
>50/6-12	11.0%	7.5%	17.8%	13.7%	15.8%	24.0%	11.0%	23.7%	27.1%	14.8%	24.8%
>50/>12	9.9%	5.4%	12.5%	10.4%	10.9%	15.7%	20.9%	27.0%	41.6%	19.1%	29.8%

Although BO2 has a more even distribution of time and emissions for the analysis dataset, we chose to use BO5 because it met the objective of the analysis (good performance across vehicle class and pollutant) and represents the best balance between performance and source/operating mode bin coverage with available data sources, particularly for light-duty (BO2 has bins which the IM240 couldn't fill). Using this approach will enable the New York IPA dataset to populate bins across the spectrum of the light-duty fleet without requiring PERE to fill more extreme operating mode bins. Possible downsides of BO5 are the unbalanced bin distributions for heavy-duty, and the heavier reliance on the relatively lax IM240 cycle to populate even the aggressive driving bins. We will revisit bin definitions for future implementations, as more data (particularly PEMS data) becomes available. The MOVES design is flexible in that bin definitions can be changed with only minor programming, database, and binner program changes.

To understand patterns across the spread of fuel consumption and emissions, predicted vs. observed per-trip results are shown in the following charts (Figures A-1 through A-8) for Bin Option 5, with 1:1 line superimposed. Per-trip predictions for BO5 show good performance across the range of fuel consumption and emissions, indicating the robustness of the approach across emitter class and driving behavior.

Fig A-1

Light Duty Fuel Consumption

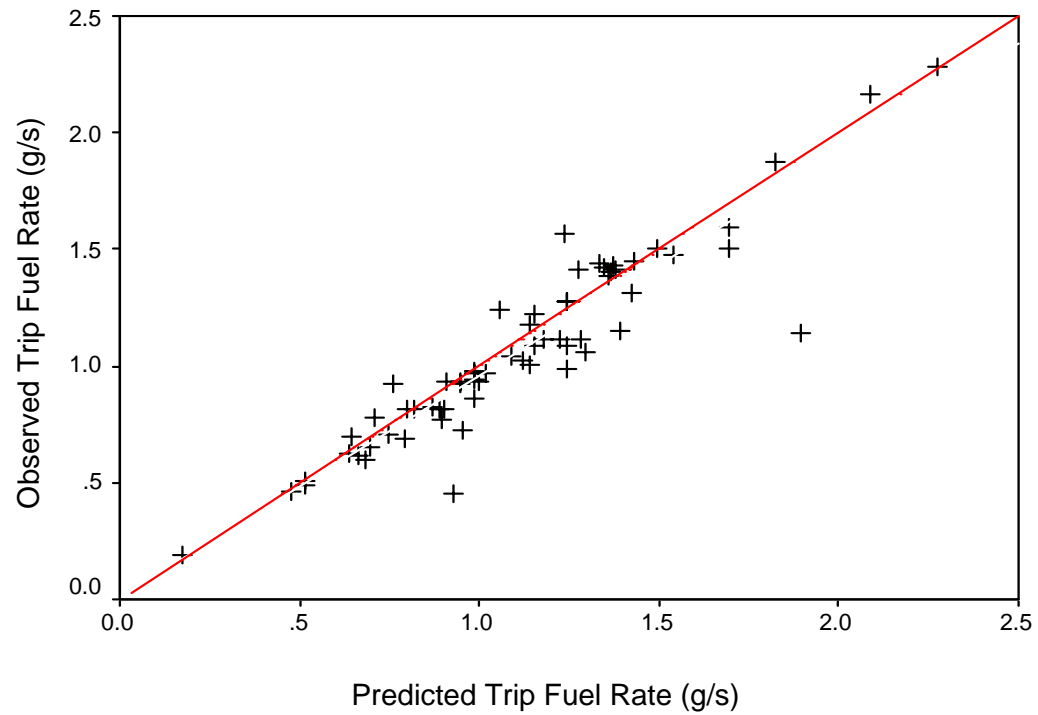
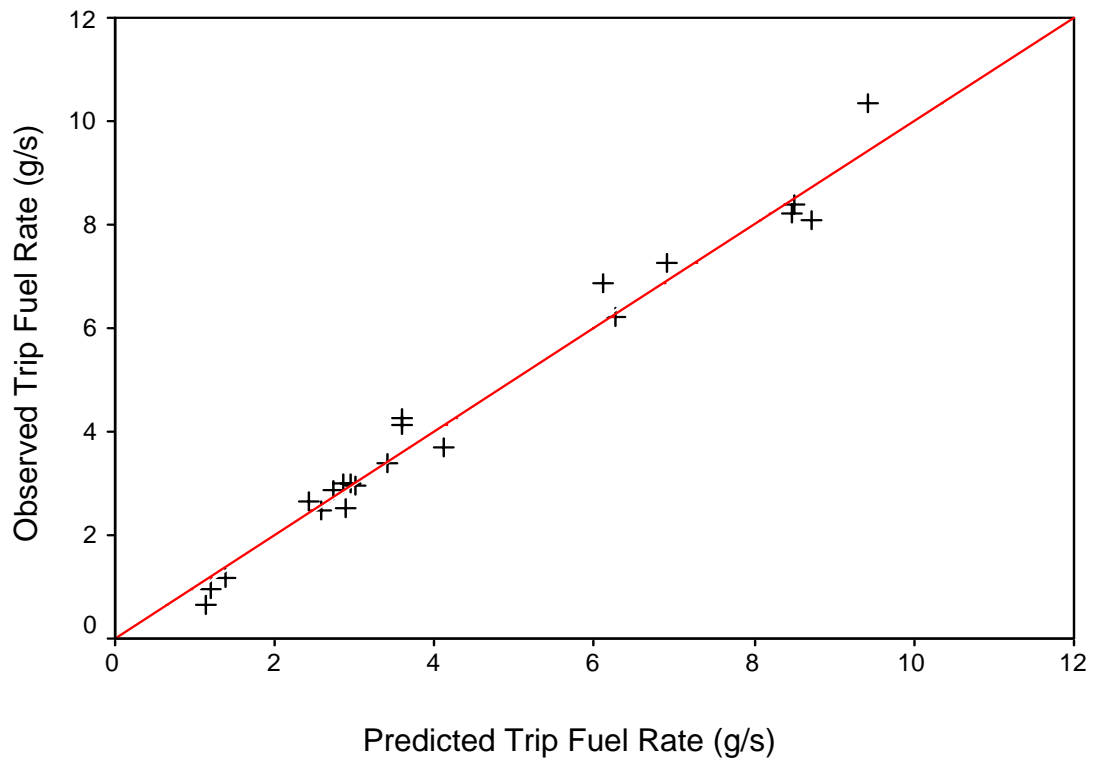


Fig A-2

Heavy Duty Fuel Consumption



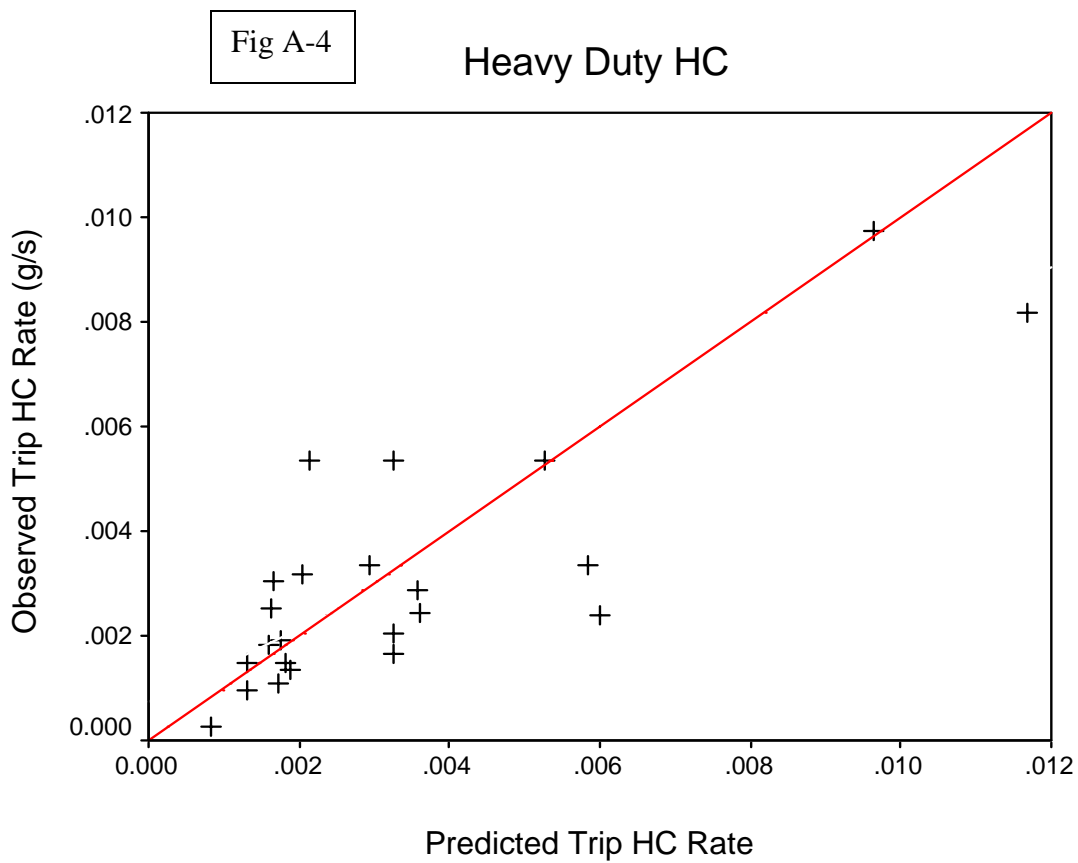
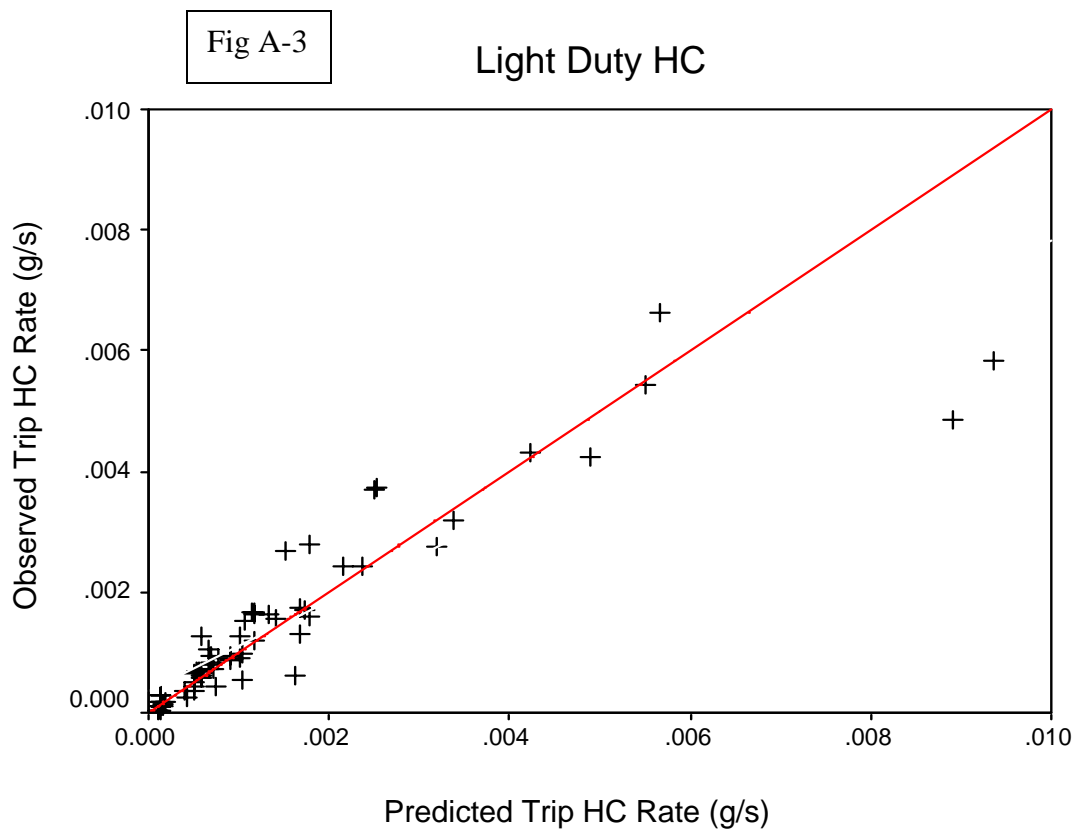


Fig A-5

Light Duty CO

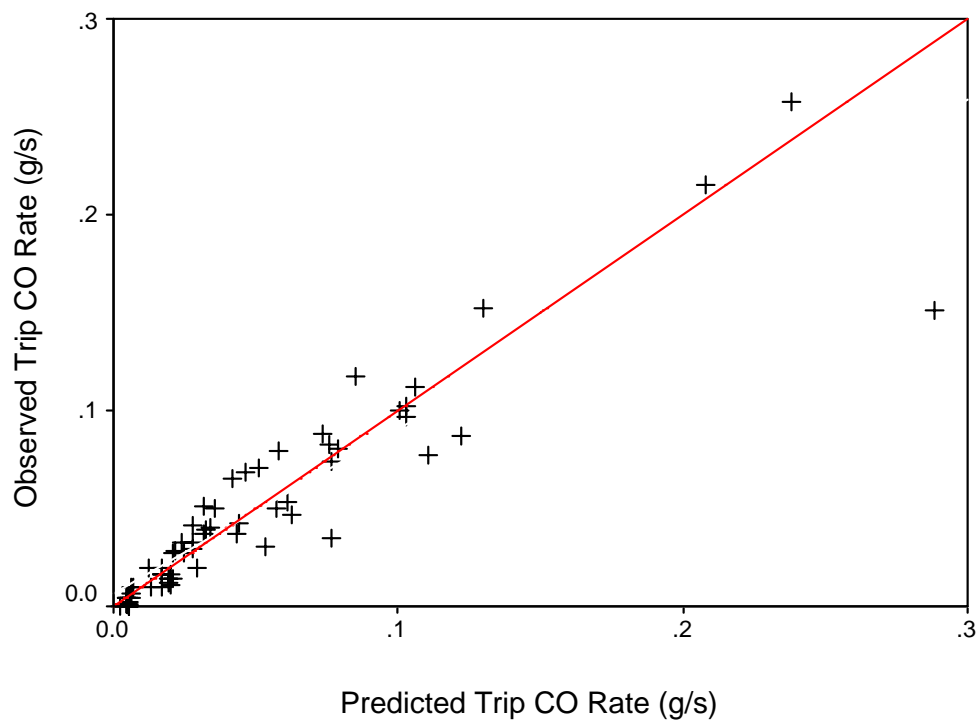
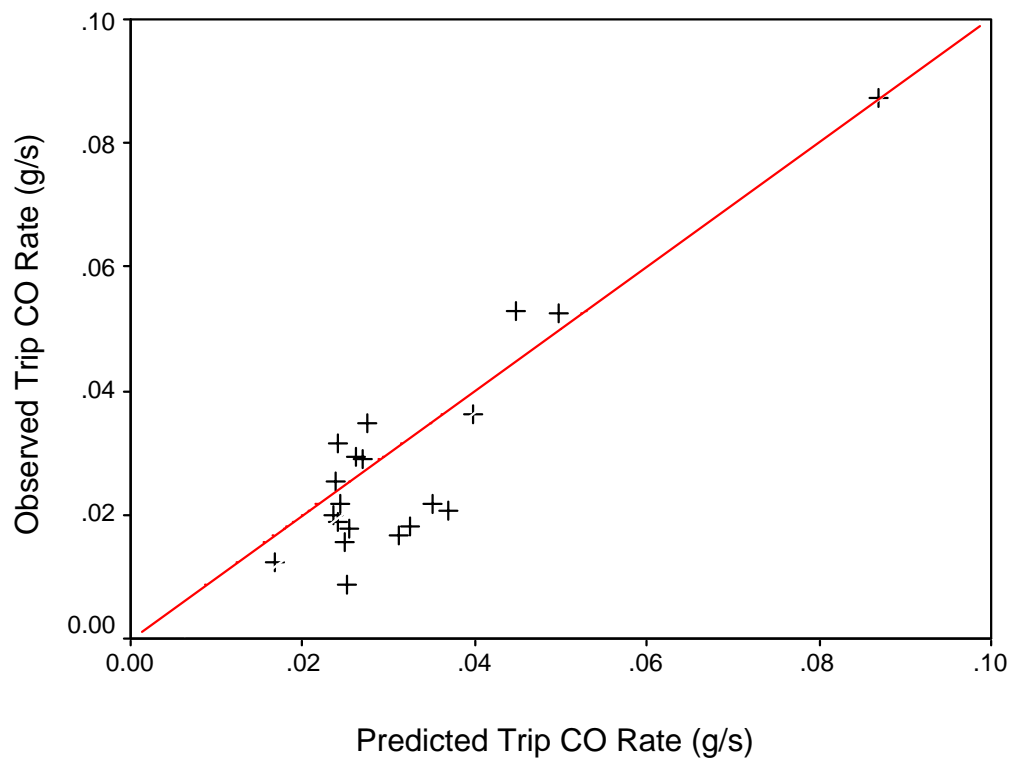
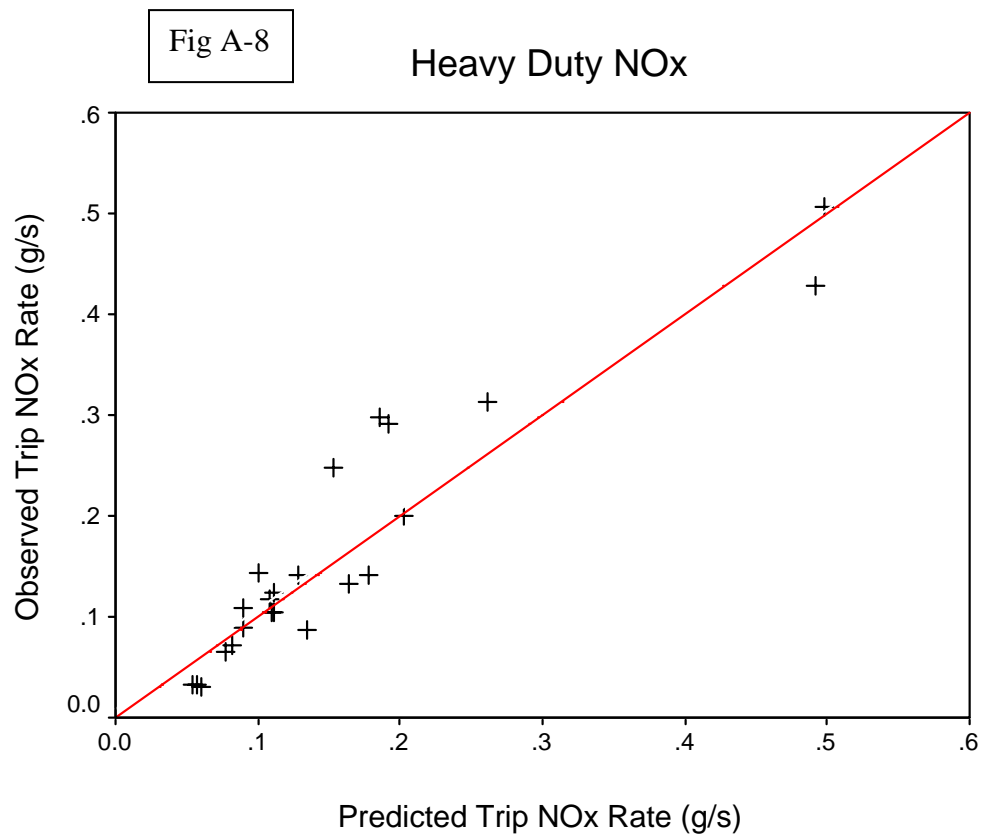
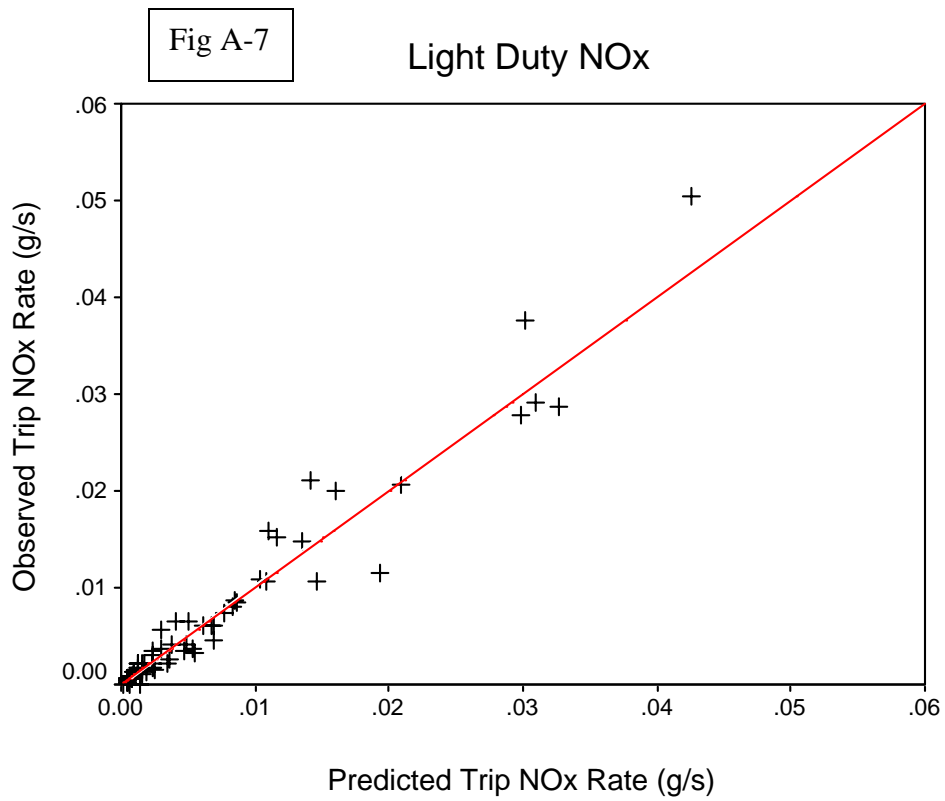


Fig A-6

Heavy Duty CO





Appendix B: Calculation of Running Energy Consumption Rates Using “The Binner Program”

B.1 The EmissionRate Table

This section describes the development of emission rates for runningEnergy consumption rates from motor vehicles in the MOVES emissionRate table. The emissionRate Table includes five fields, as shown in Table B-1. Consistent with the MOVES modal approach, the table contains mean base emission rates (meanBaseRate) and associated estimates of uncertainty in these means for motor vehicles classified as “emissions sources” (sourceBinID), and by “operating mode” (opModeID). The uncertainty estimates are expressed as coefficients of variation for the mean (meanBaseRateCV). In this section, we will describe the processes of data classification by source bin and operating mode, calculation of energy consumption, and statistical evaluation of the results.

Table B-1. Fields in the MOVES Emission Rate Table

Label	Symbol	Description
SourceBinID	---	Source Bin
OpModeID	---	Operating Mode (defined specifically for a given pollutant process)
MeanBaseRate	\bar{y}_{cell}	Cell mean (for Energy Consumption)
MeanBaseRateCV	$CV_{\bar{y}}$	Coefficient of variation of the cell mean
PolProcessID	---	Pollutant Process = 9101 (running energy)

B.2 Data

The data used to populate the emissionRate Table were compiled from selected test programs and studies. Following quality-assurance, the data was loaded into the Mobile Source Observation Database (MSOD).²⁸ The identity and scope of specific programs and studies is described in Section 3.

B.2.1 Fields Extracted from the Mobile-Source Observation Database

Consistent with the approach adopted to estimate modal emission rates, this analysis relied on data collected on a continuous “second-by-second” basis, i.e., on a measurement frequency of approximately 1.0 hertz. No data representing integrated “bag” samples was included. In addition, all data used was collected under laboratory conditions using chassis dynamometers following a variety of prescribed test procedures; no data collected “in-use” using portable instrumentation was included. Prior to data

classification and analysis, we built two temporary data tables by extracting specific data fields from MSOD. One table, VEHICLE, contains descriptive information for vehicles and tests, and the second table, DYNO, contains emissions measurements. The fields used and their sources in MSOD are listed in Table B-2. The data aggregation, classification and analysis was performed by a program written in structured query language. During this process, the test identifier field RESULTID served as a primary key linking corresponding information in the two tables.

Table B-2 Data Fields Extracted from the Mobile Source Observation Database (MSOD) for Use in Development of the MOVES Emission Rate Table

Program Table	Description	MSOD Table	MSOD Field
VEHICLE	Data source: study or program	RESULT	WA_ID
VEHICLE	Unique vehicle identifier	VEHICLE	MS_ID
VEHICLE	Unique test identifier	RESULT	RESULTID
VEHICLE	Test cycle	RESULT	TEST_PROC
VEHICLE	Test schedule	RESULT	SCHED_ID
VEHICLE	Nominal temperature for cycle (°F)	RESULT	NOM_TEMP
VEHICLE	Initial temperature for test (°F)	DYNOTEST	INIT_TEMP
VEHICLE	Vehicle class	VEHICLE	VEHCLASS
VEHICLE	Preconditioning for test cycle	DYNOTEST	PRECOND
VEHICLE	Fuel type	M_SOURCE	FUELTYPE
VEHICLE	Vehicle make	VEHICLE	MAKE
VEHICLE	Vehicle model	VEHICLE	MODEL_NAME
VEHICLE	Vehicle model year	VEHICLE	MODEL_YR
VEHICLE	Test weight	DYNOTEST	TESTWGHT
VEHICLE	Engine displacement (cubic in.)	M_SOURCE	DISP_CID
VEHICLE	Engine displacement (L)	M_SOURCE	DISP_LITER
VEHICLE	System Disablements	RESULT	DISABLE
VEHICLE	Engine Technology	M_SOURCE	FI_TYPE
DYNO	Cumulative time counter within cycle (sec)	DYNOTIME	DYNOSECS
DYNO	Vehicle speed (continuous at 1.0 Hz)	DYNOTIME	SPEED
DYNO	THC emission rate (continuous at 1.0 Hz)	DYNOTIME	R_THC
DYNO	CO emission rate (continuous at 1.0 Hz)	DYNOTIME	R_CO
DYNO	NOx emission rate (continuous at 1.0 Hz)	DYNOTIME	R_NOX
DYNO	CO2 emission rate (continuous at 1.0 Hz)	DYNOTIME	R_CO2

B.2.2 Scope of Analysis

The goal of the analysis was to estimate base energy-consumption rates for running operation. To subset for running operation, two conditions applied. First, if the test cycle were preceded by a warm-up cycle or another test cycle, all data were retained. Second, for test cycles not preceded by a warm-up cycle, the first 100 seconds of data from each test were removed. Cycles assumed to include a cold start included the EPA74, FLA4, FTP, FTPSS, ST01, and LA92. Finally, if the DISABLE field indicated that air conditioning was on during the test, corresponding measurements were removed.

B.3 Assignment of Data to MOVES SourceBins

B.3.1 SourceBin Definition for RunningEnergy

Within the MOVES modal framework, motor vehicles will be characterized as emissions sources on the basis of fuel, engine and vehicle characteristics. To estimate energy consumption, emissions data is classified on the basis by six parameters: (1) fuel type, (2) engine technology, (3) regulatory class (4) model-year, (5) engine displacement and (6) vehicle weight. A specific six-way combination of these characteristics is defined as a “Source Bin.” Within the MOVES database, each source bin is identified by a specific 20-character numeric code. For purposes of the current analysis, i.e., estimation of energy consumption, the regulatory-group attribute was not applied. Thus, the source bins as defined for this analysis effectively represent a five-way classification by the remaining attributes.

Each of the sourcebin characteristics listed above is defined as a MOVES attribute. The attributes corresponding to each characteristic are (1) fuelTypeID, (2) engTechID, (4) modelYearGroupID (shortMdYrGrpID), (5) engSizeID and (6) weightClassID. Each of the attributes is further defined below.

B.3.1.1 Fuel Type (fuelTypeID)

As the name implies, this characteristic is defined simply in terms of the type of fuel used in the engine. For example, most common fuel types currently in the database are “gasoline” or “diesel.” The attribute will also account for the introduction of future fuel types, such as “hydrogen” or “electric” vehicles. Assignment of fuelType based on values of the MSOD FUELTYPE field are presented in Table B-3.

B.3.1.2 Engine Technology (engTechID)

With respect to currently available emissions measurements, this attribute is defined in terms of fuel delivery technology. All currently available fuel delivery technologies are designated as “conventional.” The attribute is also defined to contain values for anticipated future technologies, as shown in Table B-4. Assignment of values for current technologies based on the MSOD field “FI_TYPE” is also shown.

B.3.1.3 Model Year Group (modelYearGroupID, shortModYrGroupID)

This attribute assigns emissions data to classes based on the model years of vehicles represented in the emissions database. We assigned model-year group designations based on predefined model-year ranges, as shown in Table B-5. Note that the “short” model-year group labels are used in the sourceBin attribute labels, rather than the “long” values, in order to save space.

B.3.1.4 Engine Size (engSizeID)

We define engine size in terms of displacement, expressed in liters (L). On this basis, we classified the data into predefined displacement classes, as shown in Table B-6.

B.3.1.5 Vehicle Loaded Weight (weightClassID)

This attribute represents the weight of the vehicle bearing a load during operation. It is represented by the test weight of the vehicle as recorded for each vehicle test. We classified the data into a predefined set of weight classes, as shown in Table B-7.

Table B-3. Equivalence between MSOD and MOVES databases for the SourceBin Attribute “Fuel Type” (fuelTypeID)

MSOD Field “FUELTYPE”		MOVES Attribute	
Value	Definition	fuelTypeID	Definition
GAS	GASOLINE POWERED	1	Gasoline
DIES	DIESEL POWERED	2	Diesel
CNG	COMPRESSED NATURAL GAS	3	Compressed Natural Gas
LPG	LIQUID PETROLEUM GAS	4	Liquid Petroleum Gas
E85	85% ETHANOL, 15% GASOLINE	5	Ethanol (E85 or E95)
E00	100% ETHANOL	5	Ethanol (E85 or E95)
M85	85% METHANOL, 15% GASOLINE	6	Methanol (M85 or M95)
M00		6	Methanol (M85 or M95)
<no value exists>		7	Gaseous Hydrogen
<no value exists>		8	Liquid Hydrogen
<no value exists>		9	Electricity
Any other non-null value		NULL	
NULL		NULL	

Table B-4 Equivalence between MSOD and MOVES databases for the SourceBin Attribute “Engine Technology” (engTechID)

MSOD Field “FI_TYPE”		MOVES Attribute	
Value	Definition	EngTechID	Definition
PFI	PORT FUEL INJECTION	1	Conventional
TBI	THROTTLE-BODY FUEL INJECTION	1	Conventional
NOTFI	NOT FUEL INJECTED	1	Conventional
FICARB	CARBURETED <i>(IS THIS IN MSOD?)</i>	1	Conventional
DIRECT	INTO CYLINDER INJECTION	1	Conventional
Any other non-null value		1	Conventional
NULL		NULL	
<no value exists>		10	Hybrid Electric IC-A
<no value exists>		11	Hybrid Electric IC-B
<no value exists>		12	Hybrid Electric IC-C
<no value exists>		15	Hybrid Electric DI-A
<no value exists>		16	Hybrid Electric DI-B
<no value exists>		17	Hybrid Electric DI-C
<no value exists>		20	Fuel Cell
<no value exists>		25	Hybrid Electric – Fuel Cell A
<no value exists>		26	Hybrid Electric – Fuel Cell B
<no value exists>		27	Hybrid Electric – Fuel Cell C

Table B-5 Definition of the MOVES Model-Year-Group Attribute (shortModYrGrpID) for Light-duty and Heavy-duty Motor Vehicles

Model Year	MOVES Attribute ¹	
	modelYearGroupID	shortModYrGrpID
1980 and earlier	19601980	01
1981 – 1985	19811985	02
1986 – 1990	19861990	03
1991 – 2000	19912000	04
2001 – 2010	20012010	05
¹ The short attribute values were used in actual assignment of model-year groups. The long attribute values (modelYearGroupID) are included here for completeness.		

Table B-6 Definition of the MOVES Engine-Size Attribute (engSizeID) in terms of Engine Displacement

Engine Displacement (L)	engSizeID
displacement < 2.0	0020
2.0 ≤ displacement < 2.5	2025
2.5 ≤ displacement < 3.0	2530
3.0 ≤ displacement < 3.5	3035
3.5 ≤ displacement < 4.0	3540
4.0 ≤ displacement < 5.0	4050
5.0 ≤ displacement	5099
NULL	NULL

Table B-7 Definition of the MOVES Vehicle Weight-Class Attribute (weightClassID) for Light-duty and Heavy-duty Motor Vehicles

Vehicle Weight (lb) ¹	weightClassID
weight < 2,000	20
2,000 ≤ weight < 2,500	25
2,500 ≤ weight < 3,000	30
3,000 ≤ weight < 3,500	35
3,500 ≤ weight < 4,000	40
4,000 ≤ weight < 4,500	45
4,500 ≤ weight < 5,000	50
5,000 ≤ weight < 6,000	60
6,000 ≤ weight < 7,000	70
7,000 ≤ weight < 8,000	80
8,000 ≤ weight < 9,000	90
9,000 ≤ weight < 10,000	100
10,000 ≤ weight < 14,000	140
14,000 ≤ weight < 16,000	160
16,000 ≤ weight < 19,500	195
19,500 ≤ weight < 26,000	260
26,000 ≤ weight < 33,000	330
33,000 ≤ weight < 40,000	400
40,000 ≤ weight < 50,000	500
50,000 ≤ weight < 60,000	600
60,000 ≤ weight < 80,000	800
80,000 ≤ weight < 100,000	1000
100,000 ≤ weight < 130,000	1300
130,000 ≤ weight	9999
NULL	NULL
¹ Defined as the equivalent test weight for a given vehicle test, as obtained from the MSOD field DYNOTEST.TESTWGHT.	

B.3.1.6 Source Bin Identifier (sourceBinID).

Assignment of the attributes just described allowed assignment of the source-bin identifier. The identifier is a 19-digit numeric label, of the form *1fftteyysssswww00*, where each component is defined as follows:

1 is the literal value “1,” which serves as a leading value to set the magnitude of the entire label,
ff represents the fuelTypeID,
tt represents the engTechID,
ee represents the regClassID (This attribute is not defined for running energy consumption in MOVES2004. The value 0 is inserted as a place holder),
yy represents the shortModYrGrpID,
ssss represents the engSizeID,
www represents the weightClassID, and
00 is the literal value “00,” which serves to provide two trailing zeroes at the end of the label.

The individual attributes are assembled in the proper sequence by constructing the sourceBinID as a pattern variable, where

$$\begin{aligned}\text{sourceBinID} = & 1 \times 10^{18} + \\ & \text{fuelTypeID} \times 10^{16} + \\ & \text{engTechID} \times 10^{14} + \\ & \text{emisTechID} \times 10^{12} + \\ & \text{shortModYrGrpID} \times 10^{10} + \\ & \text{engSizeID} \times 10^6 + \\ & \text{weightClassID} \times 10^2\end{aligned}\tag{B-1}$$

As an example, Table B-8 shows the construction of the sourceBin label (1010100040020002500) for light-duty gasoline vehicles, with displacement less than 2.0 L, weighing between 2,000 and 2,500 lb, and manufactured in 1991 and later.

Table B-8 Definitions of Source Bin Attributes for sourceBinID = 1010100032025003000

Description	Attribute	Form	Value	Definition
Leading integer	---	---	1	--
Fuel	fuelTypeID	<i>ff</i>	01	Gasoline
Engine Technology	engTechID	<i>tt</i>	01	Conventional
Regulatory Class	regClassID	<i>ee</i>	00	---
Model year Group	shortMdYrGrpID	<i>yy</i>	03	1986-1990
Displacement Class	engSizeID	<i>ssss</i>	2025	2.0 – 2.5 L
Weight Class	weightClassID	<i>www</i>	0030	$2,500 \leq \text{weight} < 3,000$
Trailing zeroes	---	---	00	---

B.4 Operating Modes

Within source bins, we further sub-classified data on the basis of “operating mode,” designated as the MOVES attribute “opModeID.” For motor vehicles, operating mode is defined in terms of seventeen classes defined in terms of vehicle-specific power (VSP), vehicle speed and vehicle acceleration. The derivation of the operating mode classes is discussed in Appendix A.

B.4.1 Calculation of Vehicle-Specific Power (VSP)

The first step in assigning operating mode is to calculate vehicle-specific power (VSP) for each emissions measurement. At a given time t , the instantaneous VSP _{t} (kW/tonne, at a frequency of 1.0 Hz) represents the vehicle’s tractive power $P_{\text{trac},t}$ normalized to its weight m_{tonne} . The VSP parameter is expressed as a third-order polynomial in speed, with additional terms describing acceleration and road-grade effects. The coefficients for this expression are defined in terms of rolling-resistance and aerodynamic drag coefficients, as

$$\begin{aligned} VSP_t &= \frac{P_{\text{trac},t}}{m_{\text{tonne}}} \\ &= (\mu_0^{\text{tireroll}} g)v_t + (\mu_1^{\text{tireroll}} g)v_t^2 + \left(\frac{(C_d \rho_{\text{air}} R)}{m_{\text{tonne}}} + \mu_2^{\text{tireroll}} g \right) v_t^3 + a_t v_t + g v_t \sin \theta_t \end{aligned} \quad (\text{B-2})$$

where

- μ_0^{tireroll} = zero-order tire rolling-resistance coefficient (unitless),
- μ_1^{tireroll} = first-order tire rolling-resistance coefficient (sec/m),
- μ_2^{tireroll} = second-order tire rolling-resistance coefficient (sec²/m²),
- C_d = aerodynamic drag coefficient of the vehicle (unitless),
- R = cross-sectional frontal area of the vehicle (m²),
- ρ_{air} = density of air (1.202 kg/m³),
- v_t = vehicle speed at time t (m/sec),
- a_t = vehicle acceleration at time t (m/sec²),
- m_{tonne} = vehicle weight (metric tonne),
- g = acceleration due to gravity (9.8 m/sec²),
- θ_t = road grade (radians)

In practice, a simplified form of the equation is substituted for the theoretical form presented above. In the simplified expression, the polynomial coefficients for vehicle speed are expressed as track road-load coefficients (A , B , C). In addition, because vehicle speed and weight are provided to the calculation in english units (mph and lb), appropriate conversion factors are inserted in each term. Finally, because we relied on

laboratory data to calculate emission rates, road grade = 0 and the corresponding term drops out. The revised expression is

$$\text{VSP}_t = \left(\frac{c_1}{c_2} \right) \frac{A v_t}{m} + \left(\frac{c_1^2}{c_2} \right) \frac{B v_t^2}{m} + \left(\frac{c_1^3}{c_2} \right) \frac{C v_t^3}{m} + c_1^2 v_t a_t \quad (\text{B-3})$$

where

- A = the rolling resistance coefficient (rollingTermA, kW·sec/m),
- B = the rotational resistance coefficient (rotatingTermB, kW·sec²/m²),
- C = the aerodynamic drag coefficient (dragTermC, kW·sec³/m³),
- m = vehicle weight (lb)
- v_t = instantaneous vehicle velocity at time t (1.0 Hz, mi/hr),
- a_t = instantaneous vehicle acceleration (1.0 Hz, mi/hr·sec),
- c_1 = a conversion factor for speed (0.44704 m·hr/mi·sec),
- c_2 = conversion factor for vehicle weight (0.4536 kg/lb)(0.001 tonne/kg),

B.4.2 Track Road-Load Coefficients

We estimated values for the track road-load coefficients separately for light-duty and heavy-duty vehicles, as described below. As a first step, we assigned data to “light-duty” and “heavy-duty” duty classes, on the basis of vehicle-class and vehicle weight. We defined vehicle weight as the vehicle’s test weight, as recorded for a given vehicle or set of measurements (See Table B-9).

Table B-9 Definition of Vehicle Duty Class in terms of Vehicle Class and Weight

MSOD Field “VEHCLASS”		Weight (lb) ¹	Duty Class
Value	Definition		
CAR	Car	All	LIGHT
LDV	Light-duty Vehicle	All	LIGHT
LDT	Light-duty Truck	All	LIGHT
LDT1	Class 1 Light-duty Truck	All	LIGHT
LDT2	Class 2 Light-duty Truck	All	LIGHT
TRUCK	Truck	≤ 8,500	LIGHT
TRUCK	Truck	> 8,500	HEAVY
BUS	Bus	All	HEAVY
HDDV7	Class 7 Heavy-duty Vehicle	All	HEAVY
HDDV8	Class 8 Heavy-duty Vehicle	All	HEAVY
¹ Defined as test weight for a given vehicle test.			

B.4.2.1 Track Road-Load Coefficients: Light-Duty Vehicles

For light-duty vehicles, we calculated the track load coefficients from an additional parameter, the “track road load horsepower at 50 mph” (TRLHP), based on equations B-4a – c).

$$\begin{aligned} A &= PF_A \cdot \left(\frac{TRLHP \cdot c_1}{v_{50} \cdot c_2} \right) \\ B &= PF_B \cdot \left(\frac{TRLHP \cdot c_1}{(v_{50} \cdot c_2)^2} \right) \\ C &= PF_C \cdot \left(\frac{TRLHP \cdot c_1}{(v_{50} \cdot c_2)^3} \right) \end{aligned} \quad (B-4a, b, c)$$

Where:

- PF_A = default power fraction for coefficient A at 50 mi/hr (0.35),
- PF_B = default power fraction for coefficient B at 50 mi/hr (0.10),
- PF_C = default power fraction for coefficient C at 50 mi/hr (0.55),
- c_1 = a constant, converting TRLHP from hp to kW (0.74570 kW/hp),
- v_{50} = a constant vehicle velocity (50 mi/hr),
- c_2 = a constant, converting mi/hr to m/sec (0.447 m·hr/mi·sec)).

In the process of performing these calculations, we converted from english to metric units, in order to obtain values of the track road-load coefficients in SI units, as listed above.

Values of TRLHP for most domestic and imported vehicles manufactured between 1973 and 2000 can be obtained from the “I/M Lookup Table,” a dataset developed for USEPA for use in Inspection and Maintenance programs.²⁹ Parameters contained in the file include TRLHP, test-weight, displacement, and body style. To estimate values of TRLHP for specific vehicles represented in the emissions database, we used a simple regression model to relate TRLHP to testweight and vehicle class. Graphical presentation of the data shows the expected strong relationship between TRLHP and test weight (See Figure B-1). The plot also suggests the value of vehicle class as a categorical predictor. However, The I/M Lookup table does not contain vehicle class as classified by EPA, however, the “body style” can be readily translated to vehicle class, as shown in Table B-10, along with frequencies of measurements in the dataset, cross-tabulated by body style and vehicle class. In the Lookup Table, each record represents a specific combination of division, model and model year (e.g., Toyota , Corolla, 1973).

Figure B-1. Track road-load horsepower (TRLHP) for light-duty vehicle models over the period 1973-1999, versus test weight. Source: I/M Lookup Table, version 1.8.5 (Sierra Research 2000).

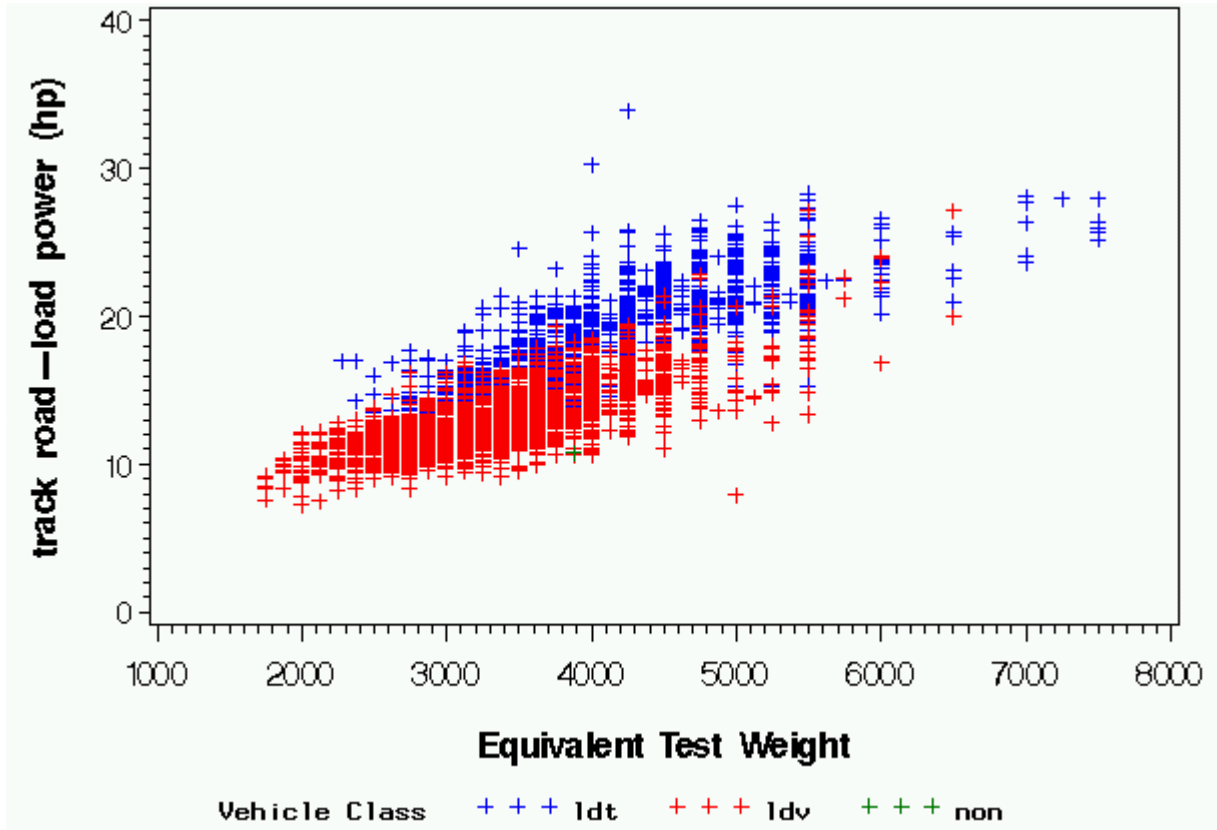


Table B-10. Frequencies of Records in the I/M Lookup Table, by Body Style and Vehicle Class

Body Style	Light-duty Vehicles (LDV)	Light-duty Trucks (LDT)
Sedan	8,209	
Wagon	1,699	
Pickup		3,204
Sport-utility Vehicle		1,954
Mini-Van		463
Full-size Van		2,087
Total	9,908	7,708

To estimate TRLHP in terms of test weight and vehicle class, we developed a general linear model designed to fit different intercepts and slopes for LDVs and LDTs. The model is in the form

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon \quad (\text{B-5})$$

where

- y = TRLHP for a given combination of division, vehicle model and model year,
- x_1 = vehicle test weight (lb),
- x_2 = an indicator variable for vehicle class:
 - = 1 for light-duty vehicles (LDV),
 - = 0 for light-duty trucks (LDT),
- β_0 = a y-intercept term,
- β_1 = a coefficient for the vehicle-class indicator, representing the difference in the intercept between LDV and LDT,
- β_2 = a slope coefficient for vehicle test weight,
- β_3 = slope coefficient for an interaction term between test weight and vehicle class, representing the difference between slopes between LDV and LDT, and
- ε = residual error between the mean estimated value of TRLHP, and the specific value for a given division/model combination, for a given value of test weight.

Use of the indicator variable for vehicle class and the interaction term allowed us to fit separate intercepts and slopes for LDV and LDT, within a single model. Table B-11 presents coefficients and goodness-of-fit statistics.

Table B-11 Coefficients and Goodness-of-Fit Statistics for the Regression Model Relating Track Road-Load Horsepower (TRLHP) for Light-duty Vehicles to Test Weight and Vehicle Class ($R^2 = 0.8064$)

Coefficient	Mean	Standard Error	t-statistic	p-value (Pr > t)
Intercept (β_0)	5.978016174	0.11881559	50.31	< 0.0001
Test weight (β_1)	0.003165941	0.00002766	114.47	< 0.0001
Vehicle Class (β_2)	-1.617898959	0.14343255	-11.28	< 0.0001
Test weight \times Vehicle Class (β_3)	-0.000390014	0.00003574	-10.91	< 0.0001

B.4.2.2 Track Road-Load Coefficients: Heavy-Duty Vehicles

No public data source is available that provides track road-load parameters for specific truck models over a series of model years. However, a source for these coefficients has been published which gives estimates of the road load coefficients for general vehicle classes, e.g., cars, short-haul delivery trucks, tractor-trailer vehicles.³⁰ Drawing on this work, we have elected to estimate general values for these coefficients in terms of the vehicles' rolling resistance and aerodynamic drag coefficients. From the physical expression of VSP (Equation B-2), the relationships defining the road-load coefficients A , B and C for heavy-duty vehicles in terms of vehicle weight (m), and the rolling resistance and aerodynamic terms are

$$A = \frac{(\mu_0^{\text{tires}} g)m}{c_1}$$

$$B = 0.0$$

$$C = \frac{C_d R \rho_{\text{air}}}{2} + \frac{(\mu_2^{\text{tires}} g)m}{c_1} \quad (\text{B-6a, b, c})$$

where

c_1 = conversion factor for vehicle weight, from lb to metric tonnes (2,204.6 lb/tonne)

Because the vehicle weight (mass) is input to the calculation in pounds, an additional conversion factor c_1 from pounds to metric tonnes is needed. The final form of the above relations between the A , B , and C road load coefficients are as displayed in Table B-12 below.

Table B-12. Road-Load Coefficients for Heavy-Duty Trucks and Buses

Coefficient ¹	Heavy-duty Vehicle Weight Class, Trucks ²			Buses
	8,500-14,000 lb (3.855-6.350 tonne)	14,000-33,000 lb (6.350-14.968 tonne)	> 33,000 lb (> 14.968 tonne)	
A	$0.0996mc_1$	$0.0875mc_1$	$0.0661mc_1$	$0.0643mc_1$
B	0.0	0.0	0.0	0.0
C	$mc_1 \left(\frac{1.47}{1,000mc_1} + 5.22 \times 10^{-5} \right)$	$mc_1 \left(\frac{1.93}{1,000mc_1} + 5.9 \times 10^{-5} \right)$	$mc_1 \left(\frac{2.89}{1,000mc_1} + 4.2 \times 10^{-5} \right)$	$mc_1 \left(\frac{3.22}{1,000mc_1} + 5.06 \times 10^{-5} \right)$

¹ As in Eq. B-3 above, units for coefficients A , B and C are kW·sec/m, kW·sec²/m² and kW·sec³/m³, respectively.

² As in Eq. B-3 above, m is vehicle weight (lb) and c_1 is a conversion factor between lb and metric tons (0.0004536 tonne/lb). For C , $1,000c_1$ is a conversion factor between lb and kg (0.4536 kg/lb).

The heavy-duty truck coefficients derived using this approach were compared to an analysis of speed-time traces of coast downs for eight heavy-duty tractor-trailer vehicles.³¹ In this analysis, vehicle speed during the coast down (external forces/vehicle mass) was parameterized as a function of time and then fit to corresponding measured values of vehicle speed determined from the vehicle speed-time traces. More explicitly, the vehicle speed was assumed to have a second order dependence on time

$$v = b_0 + b_1 t + b_2 t^2 \quad (\text{B-7})$$

Relating speed to time also allows forces on the vehicle during coast-down to be parametrized in time, as

$$F = m \frac{dv}{dt} = m(b_1 + 2b_2 t) \quad (\text{B-8})$$

This equation can then be equated to the road-load equation, by normalizing force to vehicle weight

$$\frac{F}{m} = \frac{\text{VSP}}{v} \quad (\text{B-9})$$

which allows development of relationships between the fit parameters, b_0 , b_1 , and b_2 , and the road load coefficients, A , B , and C . Finally, the physical road load parameters can be determined using the relationships in Equations B-6a-c.

To address the fact that these general relationships are under-constrained, we applied additional physical constraints. For example, we ensured that acceleration values were less than zero, loss of vehicle kinetic energy was within 5% of the actual vehicle kinetic energy loss, and that vehicle aerodynamic drag and rolling resistance coefficients were within reasonable ranges.

These preliminary analyses gave confidence that the values derived from Petrushov (1997) are reasonable initial estimates of road-load coefficients for a large variety of heavy-duty truck shapes. However, at least two considerations may merit further consideration. First, vehicle sizes may differ between the Russian and U.S. fleets. European vehicles may tend to have smaller frontal areas than those in the U.S., resulting in smaller aerodynamic drag terms. Secondly, Petrushov's tire rolling resistance factor does not include a second-order term (in speed) which leads to the road load coefficient B being equated to zero. Although some studies express the rolling resistance as the sum of a constant and a quadratic term in speed, most literature sources report non-zero values for B .³² In addition, the method used to derive track road-load coefficients for light-duty vehicles in this analysis generates non-zero values for B .

B.4.3 Calculation of Vehicle Acceleration

In addition to vehicle velocity at the 1.0 Hz measurement frequency, the VSP calculation requires estimation of acceleration at the same frequency. Thus, for each measurement at time t , we calculated acceleration a_t as the first differential of vehicle velocity, or

$$a_t = v_t - v_{t-1} \quad (\text{B-10})$$

where v_t and v_{t-1} are the vehicle velocity for the current and previous measurements, respectively. In addition, we calculated the acceleration for the two seconds prior to the current measurement (a_{t-1} and a_{t-2}) as

$$\begin{aligned} a_{t-1} &= v_{t-1} - v_{t-2} \text{ and} \\ a_{t-2} &= v_{t-2} - v_{t-3} \end{aligned} \tag{B-11a, b}$$

These latter two acceleration values are not used in calculation of VSP, but rather in defining the “deceleration” operating mode, as described below.

B.4.4 Assignment of Operating Mode (opModeID)

After calculating VSP_t for each measurement, and associated values of velocity and acceleration, we assigned operating mode on the basis of these parameters. As the name implies, opModeID is a modal variable, with 17 specific modes defined as shown in Table B-13. Only the “deceleration” mode is defined in terms of acceleration, with all others defined in terms of VSP and velocity ranges. In quantitative terms, “deceleration” occurs when acceleration was less than or equal to -2.0 mph/sec during the current second, or less than -1.0 mph/sec for three consecutive seconds, including the current second and the two previous seconds.

Table B-13 Definition of the MOVES Operating Mode Attribute for Motor Vehicles (opModeID)

OpModeID	Operating Mode Description	Vehicle-Specific Power (VSP _t , kW/tonne)	Vehicle Speed (v, mi/hr)	Vehicle Acceleration (a, mi/hr-sec)
0	Deceleration/Braking			$a_t \leq -2.0$ OR ($a_t < -1.0$ AND $a_{t-1} < -1.0$ AND $a_{t-2} < -1.0$)
1	Idle		$-1.0 \leq v_t < 1.0$	
11	Cruise/Acceleration	$VSP_t < 0$	$0 \leq v_t < 25$	
12	Cruise/Acceleration	$0 \leq VSP_t < 3$	$0 \leq v_t < 25$	
13	Cruise/Acceleration	$3 \leq VSP_t < 6$	$0 \leq v_t < 25$	
14	Cruise/Acceleration	$6 \leq VSP_t < 9$	$0 \leq v_t < 25$	
15	Cruise/Acceleration	$9 \leq VSP_t < 12$	$0 \leq v_t < 25$	
16	Cruise/Acceleration	$12 \leq VSP_t$	$0 \leq v_t < 25$	
21	Cruise/Acceleration	$VSP_t < 0$	$25 \leq v_t < 50$	
22	Cruise/Acceleration	$0 \leq VSP_t < 3$	$25 \leq v_t < 50$	
23	Cruise/Acceleration	$3 \leq VSP_t < 6$	$25 \leq v_t < 50$	
24	Cruise/Acceleration	$6 \leq VSP_t < 9$	$25 \leq v_t < 50$	
25	Cruise/Acceleration	$9 \leq VSP_t < 12$	$25 \leq v_t < 50$	
26	Cruise/Acceleration	$12 \leq VSP_t$	$25 \leq v_t < 50$	
33	Cruise/Acceleration	$VSP_t < 6$	$50 \leq v_t$	
35	Cruise/Acceleration	$6 \leq VSP_t < 12$	$50 \leq v_t$	
36	Cruise/Acceleration	$12 \leq VSP_t$	$50 \leq v_t$	

B.5 Calculation of Fuel/Energy Consumption

In MOVES2004, the mean base rate field in the Emission Rate Table represents energy consumption. For each second-by-second measurement, we calculated continuous energy consumption E_t (kJ/sec) as the product of continuous fuel consumption F_t (g/sec) and an appropriate heating value HV (kJ/g). As an initial step, we calculated fuel consumption based on emissions measurements for carbon dioxide (CO₂), carbon monoxide (CO) and total hydrocarbons (HC)

$$F_t = \left(\frac{R_{CO_2,t}}{M_{CO_2}} + \frac{R_{CO,t}}{M_{CO}} \right) (M_C + yM_H) + R_{HC,t} \quad (B-12)$$

Where inputs to the calculation are defined as follows:

$R_{CO_2,t}$ = CO₂ emission rate at time t (g/sec),

$R_{CO,t}$ = CO emission rate at time t (g/sec),

$R_{HC,t}$ = HC emission rate at time t (g/sec),

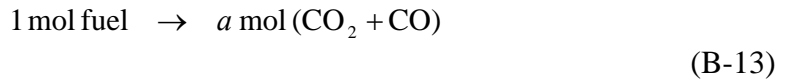
M_{CO_2} = molecular weight of CO_2 (44.0098 g/mol),
 M_{CO} = molecular weight of CO (28.0104 g/mol),
 M_{C} = molecular weight of carbon (12.0110 g/mol),
 M_{H} = molecular weight of hydrogen (1.008 g/mol),
 y = H:C ratio for fuel,
HV = heating value for fuel (kJ/g)

Table B-14 lists assumed values of H:C ratio and heating value used for each fuel.

Table B-14. Hydrogen:Carbon Ratios and Heating Values for Selected Fuels

Fuel Type	H:C Ratio (y)	Heating Value (kJ/g)
Gasoline	1.85	44.00
Diesel	1.80	43.20
Natural Gas ¹	3.80	45.02
Liquid Petroleum Gas (LPG)	3.80	46.40
Ethanol Blends ²	1.851	26.90
Methanol Blends ³	1.851	20.00
¹ Includes compressed and liquid natural gas. ² Includes ethanol blends "E85" and "E00." ³ Includes methanol blends "M85" and "M00."		

Note that the term $M_{\text{C}} + yM_{\text{H}}$ represents the mass of fuel equivalent to 1.0 mol of emissions ($\text{CO}_2 + \text{CO}$), i.e., units for this expression are g fuel/mol $\text{CO}_2 + \text{CO}$. If the fuel is assumed to contain only carbon and hydrogen, and if its nonstructural formula is represented as C_aH_b , then the fuel's molecular weight M_{fuel} can be expressed as $aM_{\text{C}} + bM_{\text{H}}$. Because the fuel's H:C ratio is defined as b/a , the number of hydrogens b is equal to ay , and M_{fuel} can also be written as $aM_{\text{C}} + ayM_{\text{H}}$, or $a(M_{\text{C}} + yM_{\text{H}})$. Further, assuming complete combustion, approximately a mols of $\text{CO}_2 + \text{CO}$ are produced by one mol of fuel, represented as



Dividing both sides by a , we see that

$$\frac{1}{a} \text{ mol fuel} \rightarrow 1 \text{ mol}(\text{CO}_2 + \text{CO}) \quad (\text{B-14})$$

If we convert the fuel to a mass basis, we see that

$$\left(\frac{1}{a} \text{ mol fuel} \right) \left(a(M_{\text{C}} + yM_{\text{H}}) \frac{\text{g fuel}}{\text{mol fuel}} \right) \rightarrow 1 \text{ mol}(\text{CO}_2 + \text{CO}) \quad (\text{B-15})$$

and finally, that

$$M_C + yM_H \text{ g fuel} \rightarrow 1 \text{ mol} (\text{CO}_2 + \text{CO}) \quad (\text{B-16})$$

Thus, we find that the molecular weight of the fuel normalized to moles of CO₂ serves as a conversion factor between moles of emissions and mass of fuel. Additionally, the molecular weight of HC emissions is assumed to equal that for the fuel, which allows the HC emission rate to be simply added to the term describing the fuel equivalent of CO₂ and CO.

B.6 Table Generation and Summary Statistics

After the steps described above, we were ready to calculate numbers to populate the EMR Table. After classifying the emissions data by source bin, we further subclassified it by operating Mode.

We calculated means and other summary statistics for each combination of sourceBinID opModeID. For simplicity, we will refer to a specific combination of sourceBinID and opModeID as a “cell.” In estimating variances for cell means, we treated the data within cells as effective cluster samples, rather than simple random samples. This approach reflects the structure of the data, which is composed of sets of multiple measurements collected on individual vehicles. Thus, measurements on a specific vehicle are less independent of other measurements on the same vehicle than of measurements on other vehicles. Accordingly, means and variances for individual vehicle tests were calculated to allow derivation of between-test and within-test variance components. These components were used in turn to calculate the variance of the mean for each cell, using the appropriate degrees of freedom to reflect between-vehicle variability.³³ To enable estimation of variances under this approach, we calculated a set of summary statistics, as listed below:

Cell mean (\bar{y}_{cell} , meanBaseRate): the arithmetic mean of all measurements in the cell.

Test mean (\bar{y}_i): the arithmetic mean of all measurements in a given test on a specific vehicle.

Cell sample size (n_{cell} , nCell), the number of individual measurements in a cell, where each count represents a measurement collected at a frequency of 1.0 Hz, (i.e., “second-by-second”).

Test sample size (n_{test} , nTest), the number of individual vehicle tests represented in a cell.

Measurement sample size (n_{meas} , nMeas): the number of measurements in a cell representing an individual test on an individual vehicle.

Test variance ($s_{\text{test},i}^2$, varTest): the variance of measurements for each vehicle test represented in a cell, calculated as the average squared deviation of measurements for a test about the mean for that test. Thus, we calculated a separate value of varTest for each test in each cell.

Between-Test variance component (s_{betw}^2 , varBetw): the component of total variance due to variability among tests, or stated differently, the variance of the test means about the overall cell mean, to be calculated as

$$s_{\text{betw}}^2 = \frac{\sum_{i=1}^{n_{\text{test}}} (\bar{y}_i - \bar{y}_{\text{cell}})^2}{n_{\text{test}} - 1} \quad (\text{B-17})$$

Within-Test Variance Component (s_{with}^2 , varWith): the variance component due to variability within tests, or the variance of measurements within individual tests about their respective test means, which can be expressed in terms of the sums of squares for each test:

$$s_{\text{with}}^2 = \frac{\sum_{i=1}^{n_{\text{test}}} \left(\sum_{j=1}^{n_{\text{meas},i}} (y_{ij} - \bar{y}_i)^2 \right)}{n_{\text{cell}} - n_{\text{test}}} \quad (\text{B-18})$$

where y_{ij} is an individual measurement i in test j . In practical terms, however, it is simpler to calculate varWith in terms of the test variances, summed over all tests in the cell:

$$s_{\text{with}}^2 = \frac{\sum_{i=1}^{n_{\text{test}}} n_{\text{meas},i} s_{\text{test},i}^2}{n_{\text{cell}} - n_{\text{test}}} \quad (\text{B-19})$$

Note that the sum of squares for test i is calculated from the test variance by multiplying by the sample size $n_{\text{meas},i}$, rather than the degrees of freedom $n_{\text{meas},i}-1$. This form is unusual with sample data, but applies in this case because the STDEV function in MySQL calculates a population standard deviation, rather than a sample standard deviation.

Variance of the cell mean ($s_{\bar{y}}^2$, varMeanCell): this parameter represents the uncertainty in the cell mean, and is calculated as the sum of the between-vehicle and within-test variance components, with each divided by the appropriate degrees of freedom.

$$s_{\bar{y}}^2 = \frac{s_{\text{betw}}^2}{n_{\text{test}}} + \frac{s_{\text{with}}^2}{n_{\text{cell}}} \quad (\text{B-20})$$

B.6.1 Cell Evaluation Criteria

Based on the summary statistics, we evaluated two criteria for each sourceBinID \times opModeID cell. These criteria determined whether the data in the cell would be considered adequate to populate the cell.

Minimum Number of Tests per Cell: To be considered as “populated,” we required that a minimum of three individual vehicle tests be represented in the cell ($n_{\text{test}} \geq 3$). This number is the bare minimum to calculate variance components for a cell and the statistics derived from them.

Coefficient of Variation of the Mean ($CV_{\bar{y}}$, CVmean): this parameter gives a relative measure of the uncertainty in the cell mean, allowing comparisons among cells. It is calculated as the ratio of the cell standard error to the associated cell mean

$$CV_{\bar{y}} = \frac{\sqrt{s_{\bar{y}}^2}}{\bar{y}_{\text{cell}}} \quad (\text{B-21})$$

For a cell to be considered “populated” we require that its standard error be no larger than one-half of its mean, or $CV_{\bar{y}} \leq 0.50$. Note that the term “CV of the mean” is synonymous with the commonly used term “relative standard error” (RSE).

B.6.2 Results: Data-Populated Cells in the EmissionRate Table

The product of the data classification and calculations described is the emissionRate Table in the MOVES database. The emissionRate Table contains five fields, as described in Table B-1. Each record in the table represents a sourceBin×opModeID cell, as previously described. For the present data compilation, the table includes results describing 2,848 cells in 154 to 171 source bins, depending on operating mode.

Appendix C: Proof-Of-Concept Assessment of Hole Filling Methods

In order to develop high-quality and dependable emissions inventory models, the MOVES model is designed to be primarily data-driven. It is assumed that as more emissions data from vehicles are collected, the emissions rates in the model can be updated and thus the fidelity should correspondingly improve. However, due to the limited existing data, there will always be data “holes” in MOVES. This appendix presents three proof-of-concept methodologies for filling these data holes for energy rate estimations in MOVES.

Preliminary analysis suggested that only about 50 percent of source bin and operating mode combination (termed “cells”) actually are filled with data (this is based on the raw number cells, not weighted for prevalence in the fleet as presented in Table 4-6). It should be noted that while all of the cells were selected because they had reflected some portion of the existing US fleet, many of these cells make up only a very small minority of the fleet mix (e.g. 1990’s compact car with diesel engine). Many of the emission rates in these empty cells may be filled by interpolating or extrapolating from nearby cells, where a physical relationship can be drawn. Others cells may be filled through disaggregating emissions bag data into their respective operating mode bins. Often bag data is present when second-by-second data is lacking. A third method of filling the operating mode bins (for a particular source bin) is through the use of a physical model such as PERE (Section 4.3.3.1). This appendix describes a proof-of-concept study that compares these three separate methods of filling empty bins. The study was performed only for light-duty cars (not trucks) since this was the extent that the data binner was prepared at the time of the study. As of the writing of this report, data quality checks on the MSOD were still being performed, the results of which will affect some of the emissions rates obtained. As a result this study does not present actual energy consumption rates as MOVES would use. The results are all at the proof-of-concept level.

Each of the methods listed above has its advantages and disadvantages. In some cases, all of the methods may be employed. The first section of this paper compares the methods for the cases when they are all applicable. In general PERE can be employed in almost all cases, though some additional developmental work may be required. However, only a fraction of the empty cells have neighboring cells populated, from which its own emission rate can be interpolated or extrapolated. Therefore, in many cases, the methodology will be guided by whether there is data available in nearby cells. This document will first identify the cells, which require filling. Then it will provide recommendations on which process is optimal for integration into MOVES by comparing the three methodologies.

C.1 Identifying Empty Bins

The following series of figures show an example of empty bins pertaining to light-duty vehicles, based on preliminary runs of the binner program described in Appendix B. Group 1 (model year pre-1980) is shown in Figure C-1. Vehicle weight is across the top, and can be determined by essentially multiplying the bin number by 100 lbs (e.g. bin 40 is 3500-4000 lbs). Engine displacement is down the left, where the bin ID defines the range of engine sizes. It is read in the following way: e.g. 4050 = 4.0 to 5.0 Liter. Green cells are cells with second-by-second data, orange cells only have bag data, and red cells have no data. Figures C-2 through C-5 shows the same chart as Figure C-1 for the other model years.

model year group 1; op mode id = 0	0	20	25	30	35	40	45	50	60	70	80	90	100	vehicle weight id
0														
20														
2025														
2530														
3035														
3540														
4050														
5099														
engine displ id														

Figure C-1. Model year group 1 (pre 1980). The (x,y) = (35, 3540) bin is the interpolated bin.

model year group 2; op mode id = 0	0	20	25	30	35	40	45	50	60	70	80	90	100	vehicle weight id
0														
20													no data	
2025													sbs data	
2530													bag data	
3035														
3540														
4050														
5099														
engine displ id														

Figure C-2. Model year group 2 (1981-1985). The (x,y) = (40, 20) bin is the extrapolated bin.

model year group 3; op mode id = 0	0	20	25	30	35	40	45	50	60	70	80	90	100	vehicle weight id
0														
20												no data		
2025												sbs data		
2530												bag data		
3035														
3540														
4050														
5099														
engine displ id														

Figure C-3. Model year group 3 (1986-1990)

model year group 4; op mode id = 0	0	20	25	30	35	40	45	50	60	70	80	90	100	vehicle weight id
0														
20													no data	
2025													sbs data	
2530													bag data	
3035														
3540														
4050														
5099														
engine displ id														

Figure C-4. Model year group 4 (1991-2000)

C.2 Hole Filling Method 1: Interpolation/Extrapolation

This is the simplest of the three methods in principle and will be described first. The process of interpolation and extrapolation involves statistically estimating the value of a particular point based on its relationship to other known values of the same function. If the unknown value has neighboring known values that are both larger and smaller, one can ‘interpolate’ between the known points to estimate the unknown. If the unknown value extends outside the range of known values (either smaller or larger) of the function, then one can ‘extrapolate’ the known points to estimate the unknown. Statistically, the tools required are identical for these estimations. However, interpolated values have much more confidence than those that are extrapolated. This is due to various unknown physical phenomena, which may occur outside the known range of values. For example, some vehicles on the “edges” may have a higher proportion of turbo driven engines, which could change the basic behavior of energy consumption in relation to the vehicles in the neighboring bins.

All of the interpolations and extrapolations conducted in this section will be within specific vehicle source bin “families”. These families are identical in all respects except for 1 “trait”. This trait defines a physical characteristic of the family. Two examples of traits are engine size and vehicle weight. It is logical to assume that there is a relationship between fuel consumption (energy) and engine size, or vehicle weight. As a rule-of-thumb, the bigger the engine, the more fuel is burned in order to follow the same driving trace. Also the heavier the car, the more work is required to move it. It is upon these simple physical principles that our justification for this methodology lies.

The previous section described all of the empty and filled cells for cars (it does not include LDT). We select the following source bin families in this study: Model year group 1 (pre-1980) for the interpolation exercise and group 2 (1981-1985) for the

extrapolation case. The interpolated bin in question is the 3540 (3.5-4.0L) cross-referenced with the 35 (3000-3500 lb) weight bin (Figure C-1). The extrapolated bin is the 20 (<2.0L) cross-referenced with the 40 (3500-4000 lb) weight bin (Figure C-2).

C.2.2 Interpolation

The interpolated emissions rates for each operating mode bin are determined using a linear regression using MS Excel. Each regression is determined from 4 known values of engine displacement bin. Figure C-5 shows the results. The vertical red line shows all the interpolated values determined from the regressions. It is interesting that there is a dip in the emissions rates measured in the 3.0-3.5L bin for all of the operating mode bins. This unphysical result, is likely due to a sampling bias. There was only a single vehicle in this bin.

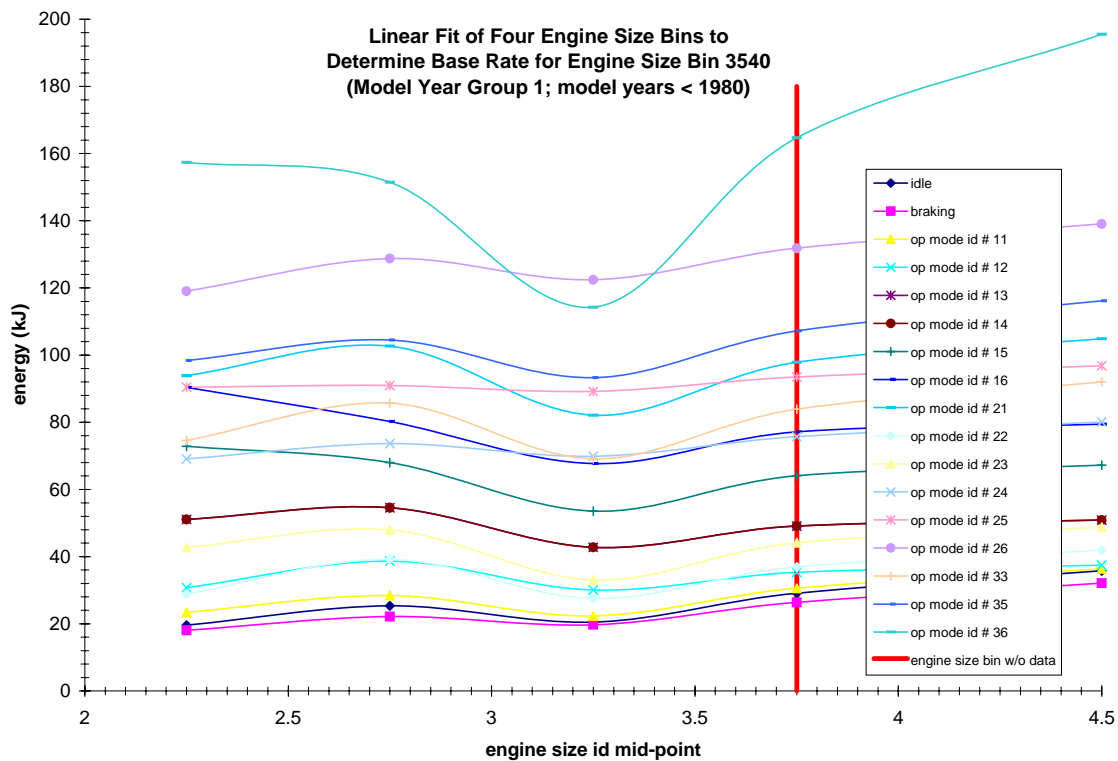


Figure C-5. Energy rates as a function of engine size for each operating mode bin. The red vertical line is the interpolated bin.

Figure C-6 shows the same results by operating mode bin. Note that there is little discrepancy between engine size bins. This will be discussed more later. The next section will address the extrapolation case.

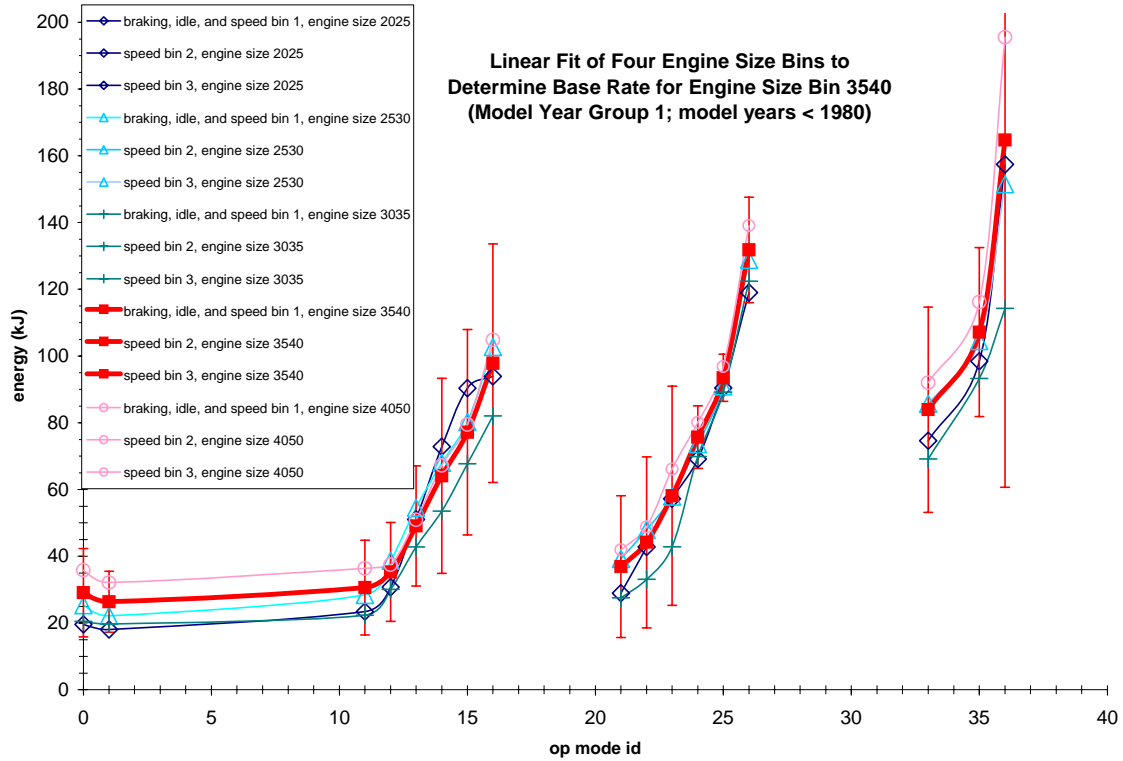


Figure C-6. Energy rates as a function of operating mode bin. The red lines are the interpolated bins.

C.2.3 Extrapolation

The cell to be extrapolated is engine size bin 20 (<2.0L) and weight bin 40 (3500-4000 lbs). The cell has 4 filled (known) weight bins that are smaller, and 6 filled engine size bins that are larger. It is interesting to compare the extrapolations from each of these two cases. Table C-1 shows the extrapolation results for each of the cases. The error bars are determined from the uncertainty in the linear regressions. The differences between the two cases are mainly pronounced in the higher VSP valued bins (16 and 26). All others are within 10%. In all cases however, they are consistent within the estimated uncertainty bounds.

Table C-1. Extrapolated energy rates in each operating mode bin from engine displacement and vehicle weight.

Op mode bin	from eng. displ	error	from weight	error	% difference
0	13.63	2.97	14.79	7.38	7.86
1	12.16	2.39	13.83	4.36	12.11
11	21.43	0.77	22.63	7.82	5.29
12	29.83	1.72	30.84	14.19	3.26
13	42.04	3.13	41.03	18.34	-2.47
14	52.00	4.66	51.58	21.41	-0.81
15	66.19	4.30	63.90	18.53	-3.58
16	84.72	5.66	71.61	25.61	-18.31
21	30.86	3.35	31.24	18.15	1.22
22	33.95	1.24	35.22	19.89	3.58
23	45.19	1.46	43.76	25.23	-3.28
24	60.54	3.50	59.62	23.77	-1.55
25	78.72	2.72	81.02	25.58	2.84
26	117.21	9.51	102.08	29.76	-14.83
33	59.47	4.44	65.62	40.57	9.38
35	73.25	5.28	78.14	36.83	6.26
36	93.85	28.61	87.28	73.07	-7.53

Figure C-7 shows the extrapolation for engine size (as in the previous case). Figure C-8 shows the extrapolation for vehicle weight. In each case, the red line represent the extrapolated values. Note in comparing these two figures that the emission rates from the various engine displacement bins are very similar, with the possible exception of the high speed bins (33, 35, and 36). This is not the case in Figure C-8, when viewed by vehicle weight. This may lead us to the conclusion that the separation of the source bins by engine size may not be necessary, or that fewer engine size bins can capture the variability. “Collapsing” the engine size bins has the advantage that there will necessarily be fewer bins, thus allowing the model to run more efficiently. As a consequence, there will also be fewer empty bins to fill. This also implies that extrapolation by vehicle weight is more reliable than that by engine displacement.

Figure C-9, demonstrates that the base rates for engine size and engine weight follow a similar behavior, as we would expect.

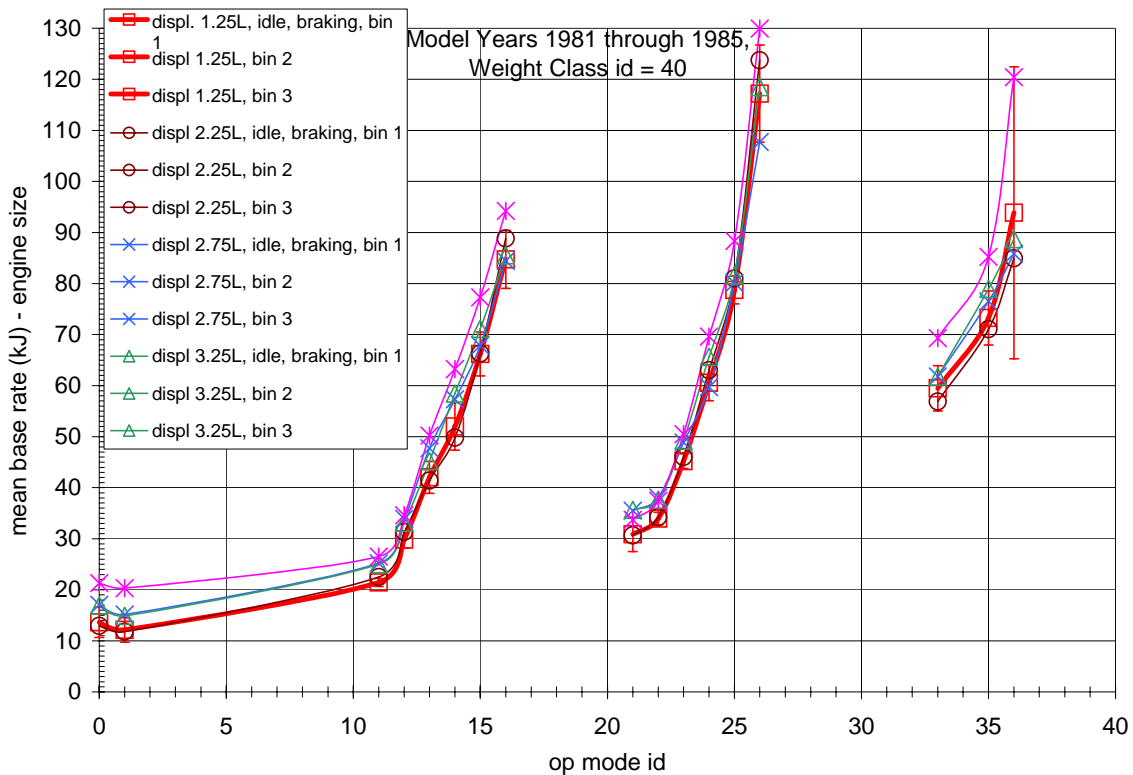


Figure C-7. Extrapolated values from engine displacement.

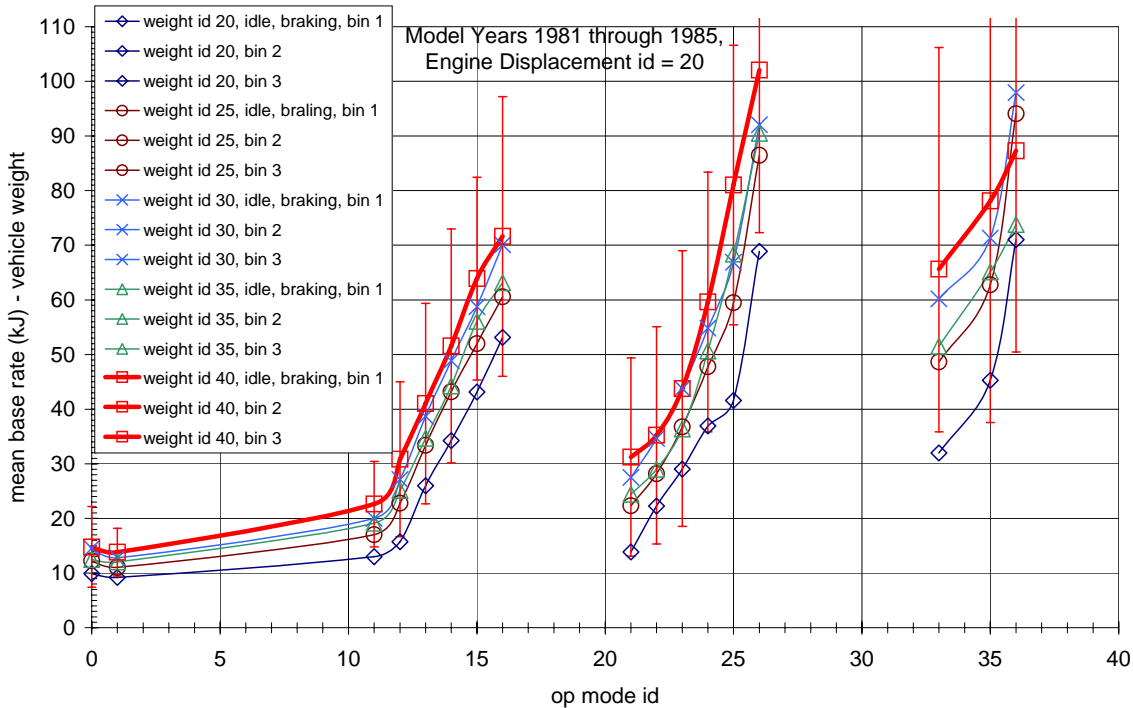


Figure C-8. Extrapolated values from vehicle weight.

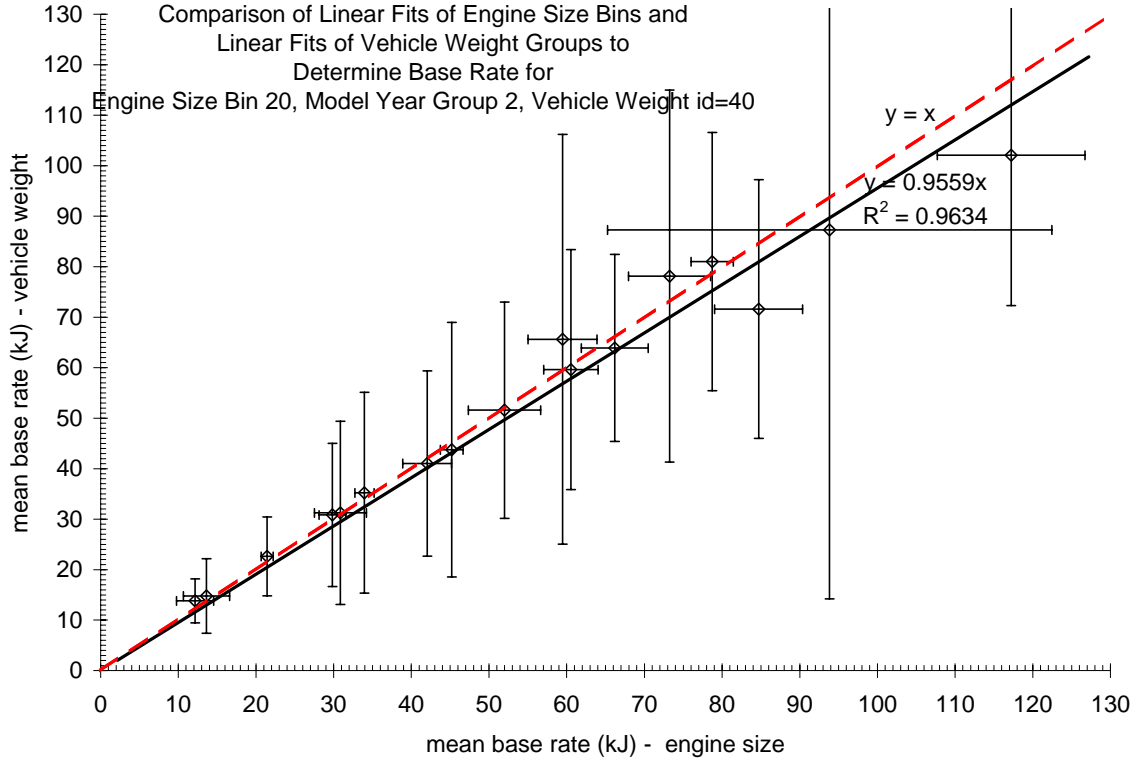


Figure C-9. A comparison of base rates for engine size and vehicle weight.

C.3 Hole Filling Method 2: Deriving Modal Rates from Bag Data

In this section, our goal is to estimate the energy consumption rate (henceforth called energy rate) within a particular operating mode bin given only the bag measurements and the VSP distribution within the driving cycle. Due to the simplifications required, we assume linear relationships only. This analysis draws on previous work performed by NC State University as part of the proof-of-concept modal binning analysis referenced in Appendix A.

In symbolic terms, we want to estimate the average energy rate in a bin: E_i , where i is the operating mode bin number ($i = 0, 1, 21, 22, \dots 36$). We know the total (or bag) energy consumption or E_{bag} , which is determined from CO₂ bag emissions measurement. For each driving cycle and vehicle, the VSP distributions are also known. Dimensionally, the product of VSP and mass is energy:

$$E_{bin_i} = \sum_t m \cdot VSP_t$$

where t is summed over the N individual second by second VSP values in a particular bin i . This is a total road-load energy, and not energy at the engine. Dimensionally, the total road-load energy over the driving cycle would then be

$$E_{cycle} = \sum_i E_{bin_i}$$

Assume that the energy at the engine is proportional to the road-load energy. This energy would be the energy consumed in the bag measurement, thus

$$E_{bag} = k \bullet E_{cycle}$$

where k is a scaling constant. The total bag energy can then be proportioned amongst the VSP determined representative energy bins:

$$E_i = k \bullet \langle E_{bin_i} \rangle$$

where

$$\langle E_{bin_i} \rangle = \frac{\sum_t m \bullet VSP_t}{N}$$

Therefore, the average value of representative energy in a bin is

$$\begin{aligned} E_i &= \frac{E_{bag}}{E_{cycle}} \bullet \langle E_{bin_i} \rangle \\ &= \frac{E_{bag}}{E_{cycle}} \bullet \frac{1}{N} \left[\sum_t m \bullet VSP_t \right]_i \end{aligned}$$

In this analysis, the bin being filled contains 5 FTP and 14 IM240 tests, for which only bag data is available. For the FTP tests, only bags 2 and 3 were used in order to minimize any potential cold start effects. In all, this source bin includes 17 distinct vehicles. The mass and TRLHP are determined for each vehicle from I/M lookup tables.³⁴ And then combined with the driving trace to determine VSP or energy distributions, i.e., the

$$m \bullet VSP_t's.$$

These energy distributions for both the IM240 and the FTP cycles were then binned according to the operating mode numbering scheme from which total cycle energies,

$$E_{cycle} = \sum_i E_{bin_i}$$

and average bin energies

$$\langle Ebin_i \rangle = \frac{\sum_t m \bullet VSP_t}{N}$$

for the two individual cycles were determined. (Note: Because there are energy losses such as engine friction and braking that are not included in the VSP equation, many VSP values are negative. All negative values of energy are set equal to zero in the averages. This will be discussed below.)

Bag data for these cycles included total NOx, total hydrocarbon, CO, and CO₂ measurements in grams/mile and total mileage. These were then used to determine the total fuel consumed and the total energy used via carbon balance and an assumed lower heating value (44MJ/kg). Finally using the bag energy, E_{bag}, and the cycle energy, E_{cycle}, a scaling factor could be determined which was used to scale the average bin values to average bag values for a particular bin.

Figure C-10 shows the energy rates determined from FTP and IM bag results. The values are considerably different from the rates determined from vehicle weight extrapolation (previous section). However, this may be an unfair comparison because the vehicles in the extrapolated bin are typical of vehicles in the bin, whereas the bag data are generated from a number of specific vehicles. Thus the average engine size of the bag sample may be quite different from the typical vehicle extrapolated from its neighboring bins. Moreover, the discrepancy at the low VSP valued bins (11, 12, 21, 22, 33, etc) is likely due to the fact that the relationship between road-load and engine load (equation #) is assumed to be linear and constrained to go through the origin. It has been demonstrated from (gamma study reference) that (mechanical) engine friction offsets this relationship, especially at low loads. Accounting for this low speed effect could compensate for this discrepancy. Thus a new method could be proposed, one that fills bins based on bag data, but using PERE. By contrast, the following section uses PERE to fill second-by-second data, which is subsequently binned.

Figure C-11 compares the scaled energy rates, with that determined from extrapolation. Again, note that all of the scaled rates are low at lower energies, but correspond on average at higher energies.

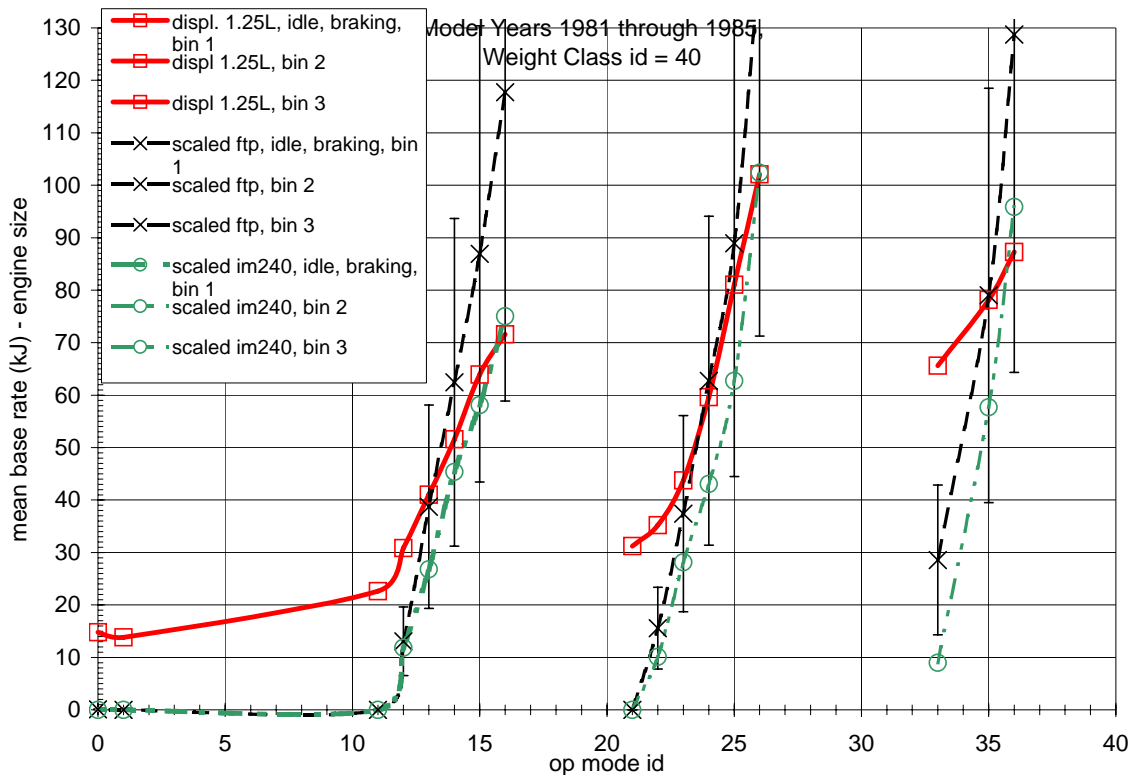


Figure C-10. The red lines are energy rates extrapolated from vehicle weight. The black and green lines are determined from ftp and IM bag results respectively.

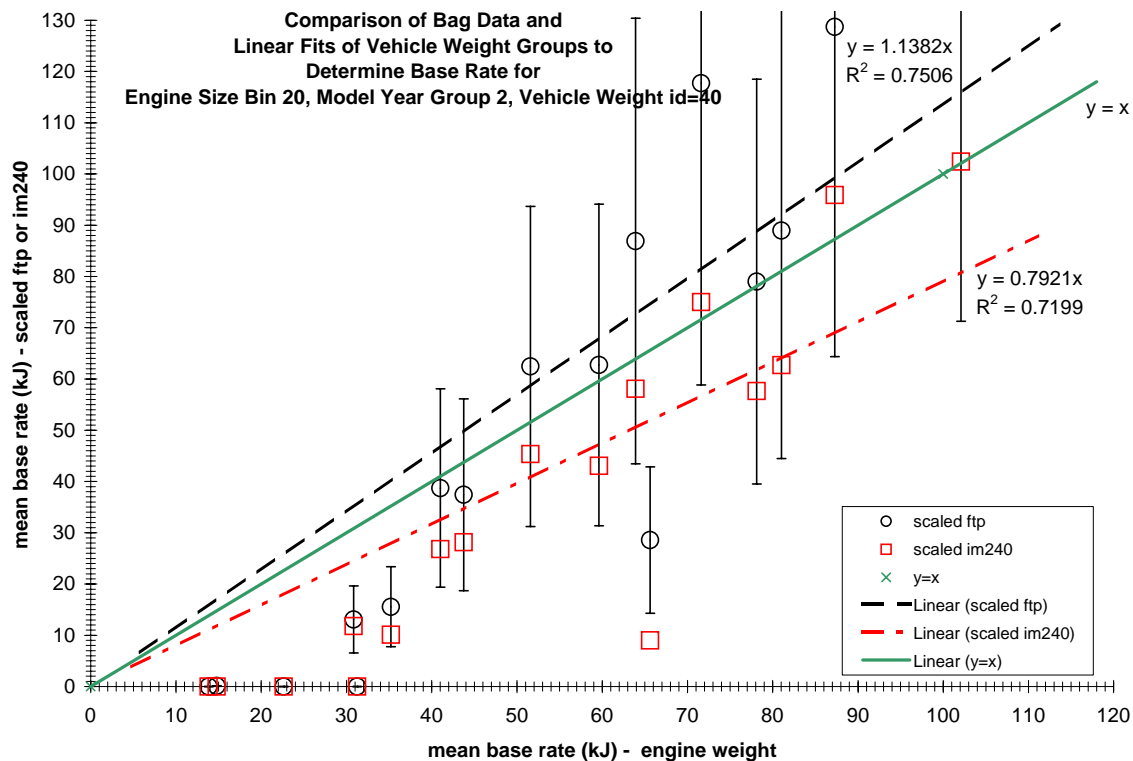


Figure C-11. A comparison of the emissions rates determined from FTP and IM bag data in comparison to the rates from vehicle weight extrapolation.

C.4 Hole Filling Method 3: PERE

This methodology calculates energy rates by employing constraints that are strictly defined by a physical model, PERE. The second-by-second output is binned to obtain emissions rates. There are two methods of employing PERE: a modeled average vehicle (within the source bin) may be calibrated to bag data, or it can predict fuel consumption as a stand-alone (pre-calibrated) model. The latter is possible since PERE is a scalable model.

The model is run for each driving cycle (for which bag data exists). Average masses and TRLHP are determined. Using typical coefficients for the engine efficiency, friction, transmission shift points, transmissions efficiency, etc., second-by-second energy rates can be determined. These energy rates are then binned by operating mode and summed to give a total bag estimate. This comparison (for CO₂) is shown on Figure 12. The PERE estimates are within 4 percent of the bag measurements. The model could in turn be calibrated to the bag data by adjusting some of the parameters in PERE. This would ensure proper correlation to the bag.

Alternately, instead of modeling an average vehicle in the source bin, one could model each vehicle (represented by bag data). This is more cumbersome, but doable, especially if automated within MOVES.

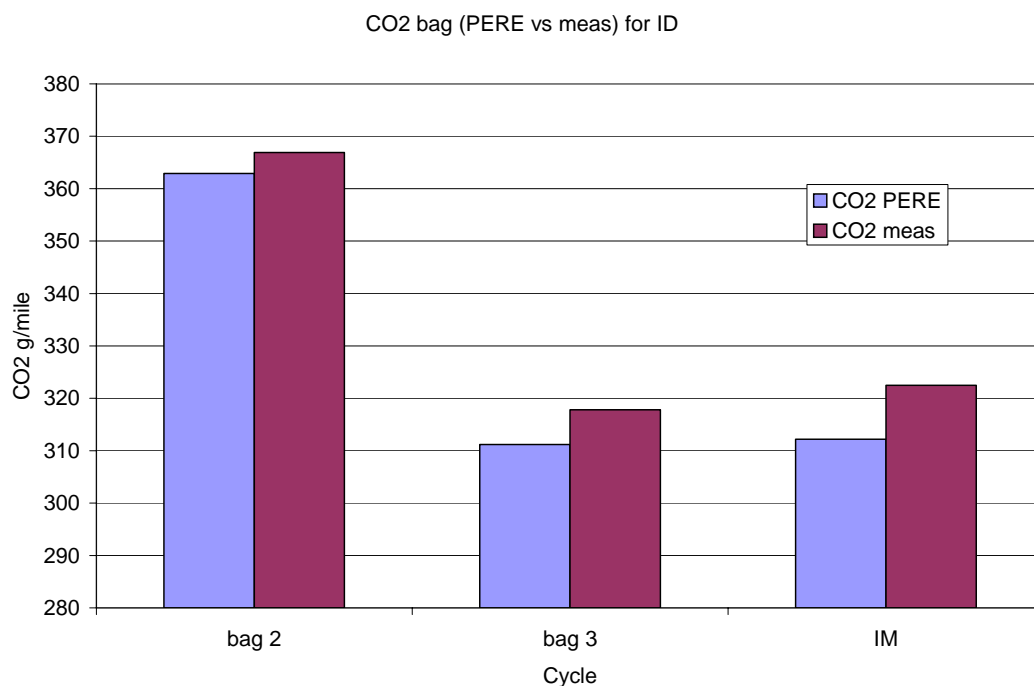


Figure C-12. Comparison of measured CO₂ with that calculated from PERE.

Figure C-13 shows the energy rates by operating mode bin. Figure C-14, compares the rates determined from PERE with those determined from the previous section. The PERE results averaged the FTP and IM rates. The PERE derived values (solid blue diamonds) more closely follow the extrapolated values at low VSP. This is due to the fact that PERE models engine friction explicitly. However, at higher speeds (>50mph), the model predicts significantly lower levels at all 3 VSP levels. As stated in the previous section, a direct comparison between method 3 and method 1 is not appropriate. The “average” vehicles in these two cases are not identical. Therefore, lower values at high speed may be real and correct. However, if there is valid reason to presume that the mode is underpredicting, this could be improved by calibrating PERE more accurately. In this study, PERE was not calibrated, the default values of coefficients were used. Moreover, it is conceivable that these “older” (carbureted) vehicles ran richer, which would increase energy consumption. This too could be calibrated explicitly in PERE.

The uncertainty bars in Figure C-14, derive from the natural variability of PERE within each bin. A better estimate of uncertainty could be achieved through the use of an application which, when linked to Excel, produces Monte Carlo uncertainty estimates. This was not done for this study.

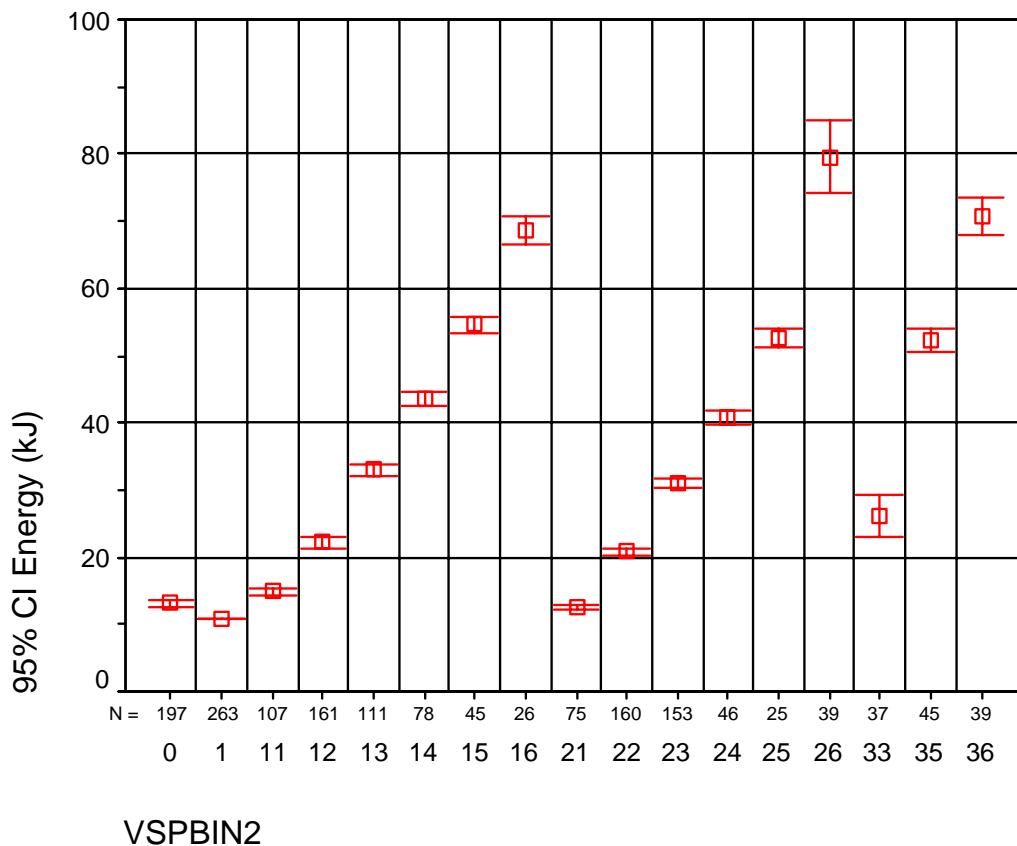


Figure C-13. Energy rates determined from PERE binned by operating mode.

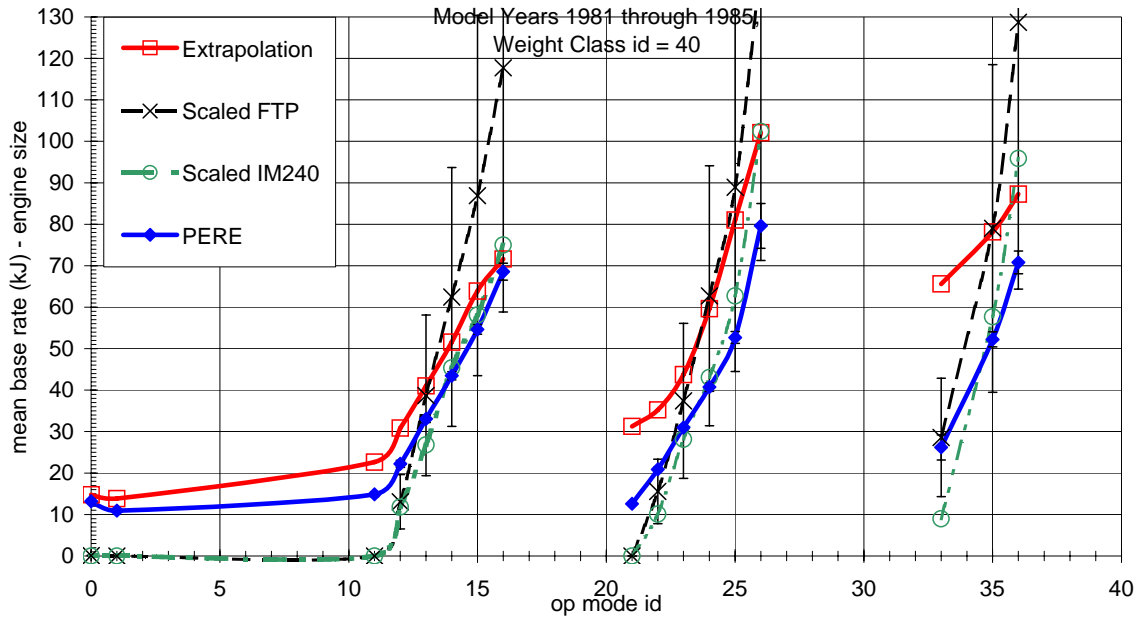


Figure C-14. PERE rates (method 3) compared with methods 1 & 2.

C.5 Conclusion

This study compared three methods for filling data holes in MOVES. The following conclusions can be drawn from the study

- Interpolation and extrapolation from the existing data of surrounding cells (method 1) is possible, however, due to data limitations (sampling issues), there may be unphysical trends visible in the data (see next point).
- A source bin, which contains only a single vehicle, may not be much better than an empty bin.
- Interpolation and extrapolation by vehicle weight is more dependable than that by engine size.
- It may be advantageous to “collapse” the number of engine size categories. This will decrease the overall number of bins and should decrease the number of “holes”.
- Estimating energy rates by scaling to bag data (method 2) gives skewed results. It tends to underpredict at low VSP and overpredict at high VSP compared to method 1. This may be due to the fact that engine load is not necessarily directly proportional to road load.
- Calculating energy rates using PERE (method 3) predicts bag accurately (within 4 percent) even without a calibration step. VSP trends are similar to those obtained from method 1, but is lower at higher speeds. This may be due to older cars running fuel-rich, or due to the different vehicles represented in the two methods.

Any recommended course of action must meet the criteria for integration into MOVES. These are:

- Accuracy
- Ease of use
- Ease of coding and integration
- Ease of updates

According to this study, interpolation/extrapolation by vehicle weight technique of filling holes is probably the simplest to integrate into MOVES. Scaling bag data (method 2) is not recommended due to the systematic errors involved. Of the three methods compared, using bag data to calibrate PERE is probably the most accurate method of filling holes. It has the advantage of employing bag data, thus taking advantage of as much data as possible. It can also estimate rates where bag data is not present. Moreover, PERE is relatively easy to implement and code at this stage (without criteria pollutants). Certainly in data poor regions, where neighboring points are sparse, there is little alternative but to use PERE. An example of data poor regions are in the source bins for light duty diesel vehicles, heavy duty diesel trucks, and motorcycles.

In the case of the unphysical dip in one of the source bins in Figure 5, arguably, PERE would provide a more reliable estimate than the data from that single vehicle. It is possible to use method 1 in situations where neighboring cells are plentiful and method 3, where they are not. However, it is recommended to stay consistent throughout the model and use a single approach, which would simplify the coding process and overall design of MOVES.

Looking ahead to criteria pollutants, the same questions will arise: How will data holes be filled? In that situation, weight will be less important for extrapolation, and bag data may be a more important guide. In such a case, PERE would likely be the most accurate option as it provides another independent check on the results.

Appendix D: Algorithm for Running Energy Hole Filling Using Interpolation / Copying

D.1 Purpose and Scope

The goal of this process was to generate values for the meanBaseRate for sourceBinID \times opModeID cells not populated by data or by the Physical Emission Rate Simulator (PERE). Throughout this discussion, we will refer to these cells as “empty cells.” We generated values for the empty cells by imputing them from neighboring cells having values.

The process of imputation was applied to runningEnergy representing gasoline and diesel vehicles produced in model years to 2010 (shortModYrGrpID = 01 to 05). Values for natural-gas vehicles were generated using the process described in Section 4.

The scope of the imputation analyses covered 689 sourceBins, with 17 opmodes in each sourceBin, giving a total of 11,713 cells. Of these, 2,848 cells contain values estimated from data and 1,624 cells have values estimated by PERE. Thus, following the use of data and PERE, 7,241 cells remained “empty,” i.e., lacked values for the meanBaseRate.

To maintain correspondence between the fleet activity and emissionRate components of MOVES, it is important to provide values for the empty cells. Because MOVES represents 100% of fleet activity, it is preferable to provide meanBaseRate values representing all activity. If cells in the emissionRate table were left empty, the model calculations would effectively assume that activity corresponding to empty cells requires no energy, resulting in underestimation of energy consumption.

D.2 Approach

The approach selected was to impute values from “neighboring sourcebins,” where neighboring sourcebins are defined as set of sourcebins for which the values of all attributes are equal, except for those in the attribute used for imputation. This analysis used three attributes for imputation: weight class (weightClassID), displacement class (engSizeID) and model-year group (shortMdYrGrpID). Considering the sourcebin attributes for runningEnergy as defining a five-way matrix, these three attributes will be referred to as the “dimensions” of imputation, denoted by the labels and symbols *weight* (*wt*, *m*), *disp* (*V*) and *myg* (*Y*), respectively. Table D-1 shows examples of neighboring sourcebins in each of the three imputation dimensions.

Table D-1 Examples of Neighboring Cells By Three Attributes Used for Imputation

SourceBin Description	Source-Bin label ¹	MeanBaseRate in opMode 15 (kJ/SO) ²	Data Source
<i>Neighbors by Weight Class (Gasoline, Conventional, MY 1991-2000, 3.0-3.5 L)</i>			
< 2,000 lb	01-01-00-04-3035-0020	---	empty
2,000 – 2,500 lb	01-01-00-04-3035-0025	196,416	PERE
2,500 – 3,000 lb	01-01-00-04-3035-0030	230,131	data
3,000 – 3,500 lb	01-01-00-04-3035-0035	230,904	data
3,500 – 4,000 lb	01-01-00-04-3035-0040	242,122	data
4,000 – 4,500 lb	01-01-00-04-3035-0045	262,610	data
4,500 – 5,000 lb	01-01-00-04-3035-0050	279,675	data
5,000 – 6,000 lb	01-01-00-04-3035-0060	---	empty
<i>Neighbors by Displacement Class (Gasoline, Conventional, MY 1991-2000, 5,000-6,000 lb)</i>			
2.0 – 2.5 L	01-01-00-04-2025-0050	---	empty
2.5 – 3.0 L	01-01-00-04-2530-0050	---	empty
3.0 – 3.5 L	01-01-00-04-3035-0050	279,675	data
3.5 – 4.0 L	01-01-00-04-3540-0050	239,560	data
4.0 – 5.0 L	01-01-00-04-4050-0050	310,731	data
5.0 L and up	01-01-00-04-5099-0050	315,892	data
<i>Neighbors by Model-Year Group (Gasoline, Conventional, 3.0-3.5 L, 5,000-6,000 lb)</i>			
1986 – 1990	01-01-00-03-3035-0050	---	empty
1991 – 2000	01-01-00-04-3035-0050	279,675	data
2001 – 2010	01-01-00-05-3035-0050	278,776	data
¹ The sourceBin label includes hyphens between the attributes to improve readability. For example, a label of 01-01-00-04-3035-0050 corresponds to a sourceBinID of 1010100043035005000. ² Operating mode 15 represents a VSP range of 9-12 kW/tonne, and a vehicle speed range of 0-25 mph.			

D.3 Methods

We imputed values for empty cells using one of two methods, “interpolation” or “substitution.” The preferred method was to estimate a value by linear interpolation between two neighboring sourcebins, designated as the “high-side” and the “low-side” neighbors to the empty cell. With respect to the *weight* dimension, “high” means “heavier” and “low” means “lighter.” Similarly, with respect to the *disp* dimension, “high” means “larger” and “low” means “smaller,” and with respect to the *myg* dimension, “high” means “later” and “low” means “earlier.” The second option, adopted when two eligible neighbors were not available, was to directly substitute the value of the nearest neighbor into the empty cell.

To allow interpolation, it was necessary to translate the coded values into numeric values in the appropriate units, e.g., lbs for *weight*, L for *disp* or years for *myg*. Thus, for all records, we assigned numeric values to each coded value, to represent a “typical” value of the attribute. For example, for the *weight* dimension, a *weightClass* value of ‘140’ was assigned a midpoint value of 12,000 lbs, calculated as the average of the upper and lower bounds of the class, or $(10,000+14,000)/2$, as shown in Table D-2. Midpoints for *disp* and *myg* dimensions were calculated similarly, as shown in Tables D-3 and D-4.

The performance of imputation involved three major steps:

- 1) Identify and store values of the nearest non-empty neighbor(s) to each cell, in each imputation dimension,
- 2) Assess the “eligibility” of neighbors so identified for use in imputation,
- 3) Perform imputation, following pre-specified orders of precedence for selection of the method and dimension to be used.

D.3.1 Identify Nearest Non-empty Neighbors

The first step was to identify the nearest non-empty neighboring cells to each record in the table, by each of the imputation dimensions, *weight*, *disp* and *myg*. We identified neighbors in three sub-steps: (1) sorting the table to arrange neighboring cells in sequence, (2) storing the *meanBaseRate* and associated fields for 15 “lags” to each record in the table (for a given sort order, defined below) and (3) storing the nearest non-empty lag.

D.3.1.1 Sorting the Table

To arrange neighboring cells in sequence, we sorted the file using the operating mode and *sourceBin* attributes as sort keys. The sort order of the attributes for each imputation dimension differed (as in table D-1 above). The sort orders always placed the *opMode* in first position, and maintained the ordering of the attributes used in the *sourceBinID* label, with the exception that the attribute for the dimension under consideration was placed in the final position. The sort orders used for each dimension are portrayed in Table D-5.

D.3.1.2 Storing Lags

After each sort of the file, the process stored the *meanBaseRate* and identifying information for a maximum of 15 “lags” to each record in the file. The lags are defined as the records preceding each record in the table, in a given sort order. All lags were stored initially, whether or not they contained values for the *meanBaseRate*. For each lag, the values of the *meanBaseRate* and *meanBaseRateCV* were stored, along with attributes needed to identify the lag and perform imputation, including:

- operating mode (opmodeID),
- sourceBin attribute values (fuelTypeID, engTechID, regClassID, shortMdYrGrpID, engSizeID, weightClassID),
- weight class midpoint,
- displacement class midpoint,
- model-year group midpoint,
- dataSourceID

D.3.1.3 Storing Non-empty Lags

After storing up to 15 lags, the process identified the lag “nearest to” each record in the table in the designated imputation dimension for which the meanBaseRate was not empty. In this context, nearest means that the difference (in the imputation dimension) between the cell under consideration and the non-empty neighbor is as small as possible. For example, if the cell under consideration represents engSizeID 2025, and the cell representing engSizeID 2530 is empty, then the cell representing 3035 is the nearest neighbor, if it is not empty. Note that it would be a “high-side” neighbor, as 3035 represents larger engines than 2025 (Table D-1).

The process repeated the steps to identify non-empty neighbors six times, to identify low-side and high-side neighbors in each of the three dimensions. For each dimension, the table was sorted twice. In the first sort, the attribute representing the imputation dimension was sorted in ascending order, to identify low-side neighbors; in the second, the attribute for the imputation dimension was sorted in descending order, to identify high-side neighbors. All attributes except that for the imputation dimension were sorted in ascending order. Any given cell can have a maximum of six non-empty neighbors; corresponding to 0, 1 or 2 non-null neighbors in each dimension.

Table D-2 Midpoints for Values of the weightClassID Attribute

WeightClassID	Class Midpoint (lb)	
	<i>Formula</i>	<i>Value</i>
0020	---	1,800
0025	$(2,000 + 2,500)/2$	2,250
0030	$(2,500 + 3,000)/2$	2,750
0035	$(3,000 + 3,500)/2$	3,250
0040	$(3,500 + 4,000)/2$	3,750
0045	$(4,000 + 4,500)/2$	4,250
0050	$(4,500 + 5,000)/2$	4,750
0060	$(5,000 + 6,000)/2$	5,500
0070	$(6,000 + 7,000)/2$	6,500
0080	$(7,000 + 8,000)/2$	7,500
0090	$(8,000 + 9,000)/2$	8,500
0100	$(9,000 + 10,000)/2$	9,500
0140	$(10,000 + 14,000)/2$	12,000
0160	$(14,000 + 16,000)/2$	15,000
0195	$(16,000 + 19,500)/2$	17,750
0260	$(19,500 + 26,000)/2$	22,750
0330	$(26,000 + 33,000)/2$	29,500
0400	$(33,000 + 40,000)/2$	36,500
0500	$(40,000 + 50,000)/2$	45,000
0600	$(50,000 + 60,000)/2$	55,000
0800	$(60,000 + 80,000)/2$	70,000
1000	$(80,000 + 100,000)/2$	90,000
1300	$(100,000 + 130,000)/2$	115,000
9999	---	148,000

**Table D-3 Midpoints for Values of the engSizeID Attribute
(Engine Displacement)**

engSizeID	FuelTypeID	Class Midpoint (L)	
		<i>Formula</i>	<i>Value</i>
0020	all	---	1.80
2025	all	$(2.0 + 2.5)/2$	2.25
2530	all	$(2.5 + 3.0)/2$	2.75
3035	all	$(3.0 + 3.5)/2$	3.25
3540	all	$(3.5 + 4.0)/2$	3.75
4050	all	$(4.0 + 5.0)/2$	4.50
5099	01	---	6.50
5099	02	---	13.0

**Table D-4 Midpoints for Values of the engSizeID Attribute
(Engine Displacement)**

shortModYrGrpID	Class Midpoint (L)	
	<i>Formula</i>	<i>Value</i>
01	---	1975.0
02	$(1981 + 1985)/2$	1983.0
03	$(1986 + 1990)/2$	1988.0
04	$(1991 + 2000)/2$	1995.5
05	$(2001 + 2010)/2$	2005.5

**Table D-5 Sort Orders for sourceBin Attributes Used to Identify Neighboring
sourceBins by Imputation Dimension**

Attribute	Dimension		
	“Weight” (<i>weight</i>)	“Displacement” (<i>disp</i>)	“Model-year Group” (<i>myg</i>)
opModeID	1 st	1 st	1 st
FuelTypeID	2 nd	2 nd	2 nd
EngTechID	3 rd	3 rd	3 rd
RegClassID	4 th	4 th	4 th
ShortMdYrGrpID	5 th	5 th	7 ^{th 1}
EngSizeID	6 th	7 ^{th 1}	5 th
WeightClassID	7 ^{th 1}	6 th	6 th
¹ To identify low-side neighbors, the attribute used as the seventh sort key was sorted in ascending order; to identify high-side neighbors, the same attribute was sorted in descending order. All other keys were sorted in ascending order.			

D.3.2 Assess Eligibility of Identified Neighbors for Imputation

Following identification and storage of the nearest non-null neighbor or neighbors, the process assessed whether the neighbor(s) would be “eligible” for use in imputation. “Eligibility” was assessed on basis of several criteria, defined in terms of the operating modes and sourceBinID attributes of the neighbors, in relation to those for the cell under consideration, i.e., the “current cell.” In all cases, the process required that values of the operating mode, fuelTypeID, engTechID, and regClassID for the neighbor and the current cell be equal. With respect to the dimensions themselves, additional criteria applied, as follows:

When assessing the eligibility of neighbors by *weight*, the process required that values for the other two dimensions (*myg* and *disp*) also be equal. In the *weight* dimension itself, the weightClassID for a low-side neighbor had to be lower than that for the current cell. Similarly, the weight class for a high-side neighbor had to be greater than that for the current cell. This criterion ensured that “low-side” and “high-side” neighbors were indeed “low” and “high,” as expected. Finally, to preclude the possibility of imputing values from “light-duty” vehicles to “heavy-duty” vehicles and vice versa, we defined an additional criterion related to vehicle weight itself. We defined an additional variable, “duty class” and stipulated that its values for the neighbor and the current cell must be equal. Duty class takes two values: “light-duty” and “heavy-duty.” We defined “light-duty” vehicles as those weighing 8,000 pounds or less (weightClassID ≤ ‘0080’), and “heavy-duty” as all vehicles weighing more than 8,000 lb. Criteria specific to the other two dimensions were similar.

When assessing the eligibility of neighbors by *disp*, the process required that values for the other two dimensions (*weight* and *myg*) be equal. In the *disp* dimension itself, the engSizeID for a low-side neighbor had to be lower than that for the current cell. Similarly, the engSizeID class for a high-side neighbor had to be greater than that for the current cell.

When assessing the eligibility of neighbors by *myg*, the process required that values for the other two dimensions (*weight* and *disp*) be equal. In the *myg* dimension itself, the shortModYrGrpID for a low-side neighbor had to be lower than that for the current cell. Similarly, the shortModYrGrpID class for a high-side neighbor had to be greater than that for the current cell. Table D-6 summarizes the eligibility criteria.

After assessing the eligibility of low-side and high-side neighbors, we assessed the possibility that interpolation would be an option for imputation. If the current cell had eligible low-side *and* high-side neighbors, as defined above, interpolation was possible. However, if only one eligible neighbor was available (either low-side or high-side), its value was substituted into the current cell. If no eligible neighbors were available, no imputation was attempted.

Table D-6 Criteria for Selection of Neighboring SourceBins Containing Values Eligible for Interpolation or Substitution

SourceBin Attribute by Dimension	Low-side neighbor	High-side neighbor
<i>All Dimensions</i>		
opMode	EQUAL	EQUAL
FuelTypeID	EQUAL	EQUAL
engTechID	EQUAL	EQUAL
regClassID	EQUAL	EQUAL
<i>Dimension="Weight"</i>		
shortMdYrGrpID	EQUAL	EQUAL
engSizeID	EQUAL	EQUAL
Duty class	EQUAL	EQUAL
weightClassID	<	>
<i>Dimension="Displacement"</i>		
shortMdYrGrpID	EQUAL	EQUAL
weightClassID	EQUAL	EQUAL
engSizeID	<	>
<i>Dimension="Model-year Group"</i>		
engSizeID	EQUAL	EQUAL
weightClassID	EQUAL	EQUAL
shortMdYrGrpID	<	>

D.3.3 Precedence of Imputation Dimensions and Methods

Zero to two neighbors were potentially available in each of the three imputation dimensions, for a theoretical maximum of six neighbors. To govern the selection of the method and dimension to be used for inputation, we applied one of three selection sequences on a cell-by-cell basis.

The selection sequence differed for each of three groups of cells defined on the basis of fuel type, model-year group and operating mode. The definition of each “precedence group” is presented in Table D-7a. Table D-7b presents the options and precedence for each group. Each option for imputation is identified with a numeric code. The code is used to develop a dataSourceID, which identifies the option used to fill any given cell and on which of two rounds of imputation a cell was filled.

For each cell, the process would evaluate each option in order of precedence, until it identified the option with the highest precedence that was also feasible. It would assign an option ID to each cell, and perform the imputation accordingly.

Table D-7a Precedence Groups for Selection of Imputation Options

Group	fueltypeID	shortModelYrGrpID	OpModeID
Precedence Group 1	01 (gas)	all	All except idle
	02 (diesel)	01-04 (pre 2000)	All except idle
Precedence Group 2	01 (gas)	all	Idle (1)
	02 (diesel)	01-04 (pre 2000)	Idle (1)
Precedence Group 3	02 (diesel)	05 (2000 – 2010)	All opmodes

Table D-7b Options and dataSourceID Values for Imputation of meanBaseRate for Empty Cells

Imputation method	Dimension	Precedence	Option ID	dataSourceID	
				1 st round	2 nd round
Precedence Group 1					
Interpolation	Weight	1	401	4011	4012
	Displacement	2	402	4021	4022
	Model-Year Group	3	403	4031	4032
Substitution	Model-Year Group	4	404	4041	4042
	Weight	5	405	4051	4052
	Displacement	6	406	4061	4062
None eligible		7	400	4001	4002
Precedence Group 2					
Interpolation	Displacement	1	411	4111	4112
	Weight	2	412	4121	4122
	Model-year Group	3	413	4131	4132
Substitution	Model-year Group	4	414	4141	4142
	Displacement	5	415	4151	4152
	Weight	6	416	4161	4162
None eligible		7	410	4101	4102
Precedence Group 3					
Substitution	Model-year Group	1	424	4241	4242
None eligible		2	420	4201	4202

D.3.4 Linear Interpolation

The process performed interpolation in two steps. In the first step, it calculated a slope term between the low-side and high-side neighbors, in the dimension previously assigned for a given cell. The slope terms for the three dimensions are shown in equations D-1, D-2 and D-3, in which the superscripts on all terms represent the imputation dimension. The denominator terms, m , V and Y represent midpoint estimates for high-side and low-side neighboring sourcebins for the *weight*, *disp* and *myg* dimensions, respectively (Tables D-2 – D-4).

In the weight dimension (superscript = *wt*), the slope term $\Delta E/\Delta m$ represents the change in the meanBaseRate per unit change in vehicle weight between the high and low neighbors, as shown in Figure D-1. The additional slope terms $\Delta E/\Delta V$ and $\Delta E/\Delta Y$ represent corresponding changes in the meanBaseRate per unit change in displacement volume or model year for the *disp* and *myg* dimensions, respectively.

$$\left(\frac{\Delta E}{\Delta m} \right)_{\text{high,low}}^{wt} = \frac{E_{\text{high}}^{wt} - E_{\text{low}}^{wt}}{m_{\text{high}}^{wt} - m_{\text{low}}^{wt}} \quad (\text{D-1})$$

$$\left(\frac{\Delta E}{\Delta V} \right)_{\text{high,low}}^{disp} = \frac{E_{\text{high}}^{disp} - E_{\text{low}}^{disp}}{V_{\text{high}}^{disp} - V_{\text{low}}^{disp}} \quad (\text{D-2})$$

$$\left(\frac{\Delta E}{\Delta Y} \right)_{\text{high,low}}^{myg} = \frac{E_{\text{high}}^{myg} - E_{\text{low}}^{myg}}{Y_{\text{high}}^{myg} - Y_{\text{low}}^{myg}} \quad (\text{D-3})$$

In the second step, the process applies the slope term to calculate an estimate for the empty cell by linear interpolation., as shown in equation D-4.

$$E_{\text{empty}}^{wt} = E_{\text{low}}^{wt} + \left(\frac{\Delta E}{\Delta m} \right) (m_{\text{empty}}^{wt} - m_{\text{low}}^{wt}) \quad (\text{D-4})$$

The interpolation is in simple linear form ($y=bx+c$). Using the example of interpolation in the *weight* dimension, the terms are defined as follows:

y = the meanBaseRate for the empty cell (E_{empty}^{wt} , kJ/SHO),

x = the difference between the weight midpoints for the empty cell and its low-side neighbor ($m_{\text{empty}}^{wt} - m_{\text{low}}^{wt}$),

b = the slope term between the high-side and low-side neighbors ($\Delta E/\Delta m$, kJ/SHO-lb),

c = the meanBaseRate for the low-side neighbor (E_{low}^{wt} , kJ/SHO), which acts as a y-intercept.

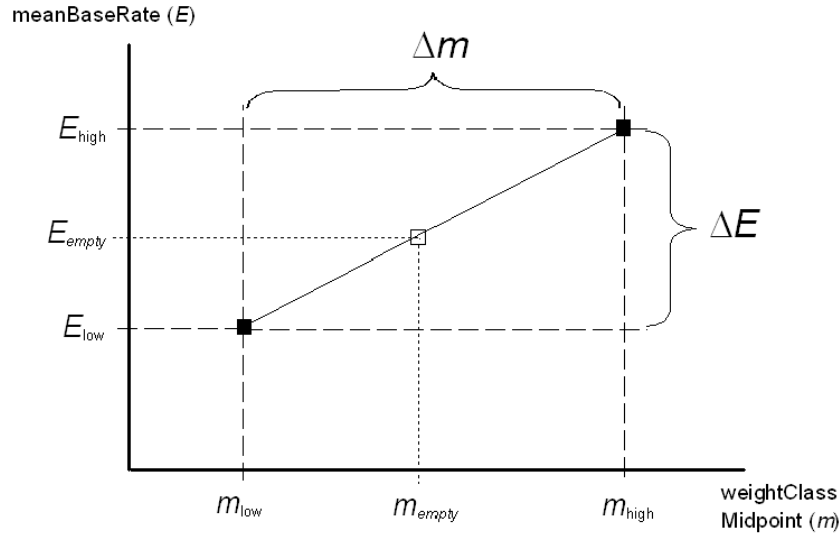
Equations D-5 and D-6 show corresponding interpolations in the *disp* and *myg* dimensions.

$$E_{empty}^{disp} = E_{low}^{disp} + \left(\frac{\Delta E}{\Delta V} \right) (V_{empty}^{disp} - V_{low}^{disp}) \quad (D-5)$$

$$E_{empty}^{myg} = E_{low}^{myg} + \left(\frac{\Delta E}{\Delta Y} \right) (Y_{empty}^{myg} - Y_{low}^{myg}) \quad (D-6)$$

If only one eligible neighbor were available, imputation was performed by substitution rather than interpolation. However, if no eligible neighbor were available, no imputation was attempted.

Figure D-1. Conceptual illustration of imputation by linear interpolation, using the example of interpolation in the vehicle-weight dimension (*wt*).



D.3.5 Rounds of Imputation

To maximize the number of empty cells filled, we performed two rounds of imputation. During first round, only cells filled by data or PERE were available as neighbors. During the second round, the revised file, including initial imputation results,

was resubmitted to the process, during which cells filled during the first round were also available as neighbors.

Specific values of the `dataSourceID` were generated to identify cells filled by imputation. The first three digits indicate the method and dimension of imputation, as indicated by the option ID, and the fourth digit indicates the round of imputation. The value of the fourth digit is either '1' or '2,' as shown in Table D-7 above.

D.4 Results and Examples

D.4.1 Results

Table D-8 shows examples of imputation in the three dimensions. The examples shown include those presented in Table D-1, and show how values were imputed to the empty cells and the identity of the neighbor(s) used for imputation. The table shows several examples of substitution in different dimensions, and one example of interpolation by displacement class. The set of cells differing only by `weightClass` (Example 1) shows that in some cases, a single cell may serve as a neighbor to several empty cells, e.g., the '0050' cell acts as a low-side neighbor to the '0060,' '0070' and '0080' cells. These same cells illustrate how the method avoided imputation by weight between light-duty and heavy-duty vehicles. In imputing a value for the '0090' cell, the method avoided substitution from the '0050' cell, defaulting instead to the next dimension in precedence, and substituting a value by displacement. In addition, Example 3 shows that the same cell used for substitution by weight in Example 1 was also used for substitution by model-year group (*myg*).

Table D-9 summarizes the results of the entire process, showing counts of cells cross-tabulated by data source and operating mode, for all cells included in the scope of this analysis, as defined in D.1 above. Within each operating-mode column, the count represents the number of `sourceBins` represented in that mode. Also within each operating mode, the sum of all rows, representing the total of cells populated by all `datasources`, is identical. This summation represents the disposition of all cells as populated by differing `dataSources`.

Table D-8 Examples of Imputation of meanBaseRate Values to Empty Cells, for Three Groups of Related SourceBins, in Operating Mode 15¹

Low-side Neighbor ²	Source Bin “Current Cell”	High-side Neighbor ³	meanBaseRate (kJ/SFO)	dimension	method	DataSourceID
Example 1: Gasoline, Conventional, MY 1991-2000, 3.0-3.5 L						
	01-01-00-04-3035-0020	01-01-00-04-3035- <u>0025</u>	196,416	weight	subst	4051
	01-01-00-04-3035-0025		196,416			394
	01-01-00-04-3035-0030		230,131			398
	01-01-00-04-3035-0035		230,904			398
	01-01-00-04-3035-0040		242,122			398
	01-01-00-04-3035-0045		262,210			398
	01-01-00-04-3035-0050		279,675			398
01-01-00-04-3035- <u>0050</u>	01-01-00-04-3035- <u>0060</u>		279,675	weight	subst	4051
01-01-00-04-3035- <u>0050</u>	01-01-00-04-3035- <u>0070</u>		279,675	weight	subst	4051
01-01-00-04-3035- <u>0050</u>	01-01-00-04-3035- <u>0080</u>		279,675	weight	subst	4051
	01-01-00-04- <u>3035</u> -0090	01-01-00-04- <u>4050</u> -0090	529,056	disp	subst	4061
	01-01-00-04- <u>3035</u> -0100	01-01-00-04- <u>4050</u> -0100	611,056	disp	subst	4061
01-01-00- <u>03</u> -3035-0140	01-01-00- <u>04</u> -3035-0140		685,872	myg	subst	4041
	01-01-00-04- <u>3035</u> -0160	01-01-00-04- <u>5099</u> -0160	896,544	disp	subst	4061
Example 2: Gasoline, Conventional, MY 1991-2000, 5,000 – 6,000 lb						
01-01-00-04-2025- <u>0045</u>	01-01-00-04-2025- <u>0050</u>		260,034	weight	subst	4051
01-01-00- <u>03</u> -2530-0050	01-01-00- <u>04</u> -2530-0050		237,455	myg	subst	4041
	01-01-00-04-3035-0050		279,675			398
	01-01-00-04-3540-0050		239,560			398
	01-01-00-04-4050-0050		310,731			398
	01-01-00-04-5099-0050		315,892			398
Example 3: Gasoline, Conventional, 3.0-3.5 L, 5,000-6,000 lb						
01-01-00-03- <u>2530</u> -0050	01-01-00-03- <u>3035</u> -0050	01-01-00-03- <u>4050</u> -0050	253,087	disp	interp	4021
	01-01-00-04-3035-0050		279,675			398
	01-01-00-05-3035-0050		278,776			398
¹ Each group, as indicated by the Sourcebin Label, differs in only one sourceBin attribute. ² The “low-side” neighbor used for interpolation or substitution. The attribute that differs between the “low-side” neighbor and the “current cell” is underlined. ³ The “high-side” neighbor used for interpolation or substitution. The attribute that differs between the “high-side” neighbor and the “current cell” is underlined.						

Table D-9 Distribution of Cells Representing Vehicles Manufactured Prior to 2010, for which Values were Derived from Data, Generated by PERE, or Imputed by Interpolation or Substitution, by Operating Mode

DataSourceID ¹	Operating Mode																	Total
	0	1	11	12	13	14	15	16	21	22	23	24	25	26	33	35	36	
1001 (data)	171	170	171	171	171	171	166	156	171	171	171	171	168	159	168	168	154	2,848
2001 (manual) ²	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	119
3001 (PERE)	95	95	95	95	95	95	95	95	95	95	95	95	95	95	98	98	98	1,624
4011	10		10	10	10	10	15	25	10	10	10	10	13	22	10	10	24	209
4111		19																19
4012	4		4	4	4	4	6	9	4	4	4	4	5	8	4	4	8	80
4112		13																13
4021	16		16	16	16	16	16	16	16	16	16	16	16	16	16	16	15	255
4121		7																7
4022	11		11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	176
4122		2																2
4041	105		105	105	105	105	105	105	105	105	105	105	105	105	105	105	105	1,680
4141		106																106
4042	26		26	26	26	26	28	29	26	26	26	26	28	29	26	26	29	429
4142		26																26
4051	115		115	115	115	115	115	115	115	115	115	115	115	115	115	115	115	1,840
4151		152																152
4052	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	16
4152		2																2
4061	83		83	83	83	83	79	75	83	83	83	83	80	76	83	83	77	300
4161		46																46
4062	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	16
4162		0																0
4241	22	22	22	22	22	22	22	21	22	22	22	22	22	22	22	22	20	371
4242	22	22	22	22	22	22	22	23	22	22	22	22	22	22	22	22	24	377
Total	689	689	689	689	689	689	689	689	689	689	689	689	689	689	689	689	689	11,713

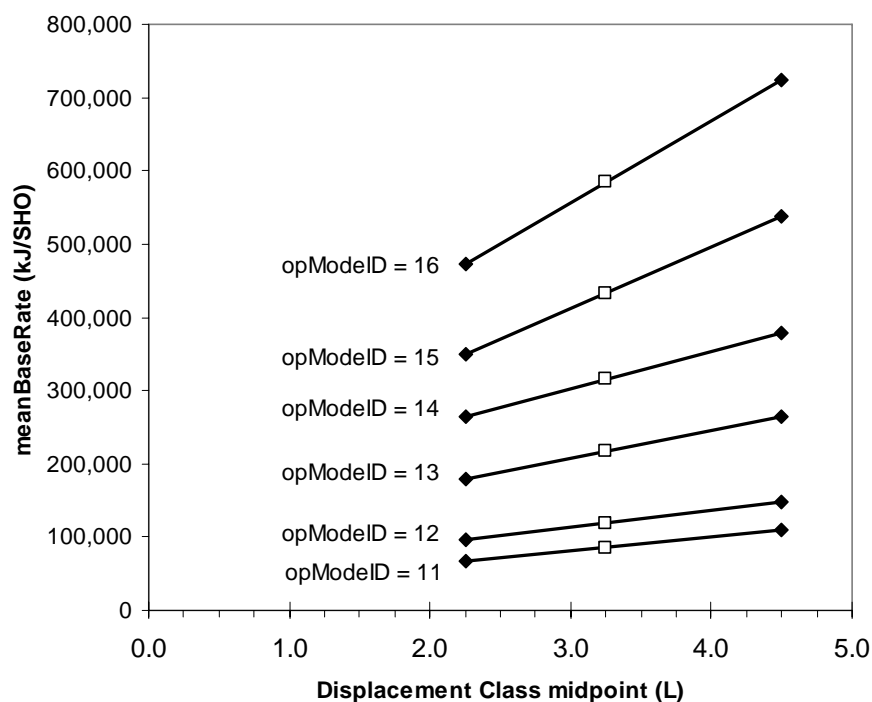
¹ Values for dataSourceID other than data or PERE are defined in table D-1

² Cells for which values were not imputed following the second round of imputation. These cells were subsequently filled “manually.”

D.4.2 Example of Interpolation

Figure D-2 shows an example of interpolation by displacement for six operating modes. In this case, the empty cells (squares) represent gasoline vehicles with displacement in the range of 3.0-3.5 L, weighing from 5,000-6,000 lbs and produced during the period 1986-1990 (01-01-00-03-3035-0060). Estimates for these vehicles were imputed by interpolating between the low-side neighbor (2.0-2.5L, engSizeID=2025), and the high-side neighbor (4.0-5.0 L, engSizeID=4050), both denoted by diamonds. Note that in this case the next lower size class to the empty cells (2.5-3.0 L, engSizeID=2530) is not available to serve as a low-side neighbor. Thus, if the cells for size classes 0020 and 2025 were empty, interpolation would not have been possible in the *disp* dimension, rather, the value for the high-side neighbor would have been substituted. Note that although the imputed rates are not shown in the graph, the two intervening source bins with engSizeID classes 2530 (2.5-3.0 L) and 3540 (3.5-4.0 L) were also interpolated between size classes 2025 and 4050.

Figure D-2. Examples of interpolation by displacement class (disp) in six operating modes. The empty cells (squares) represent sourceBin label 01-01-00-03-3035-0060, which are interpolated between the low-side neighbor 01-01-00-03-2025-0060 and the high-side neighbor 01-01-00-03-4050-0060 (diamonds).



Appendix E: Start Energy Rates

MOVES2004 start energy rates were generated by multiplying the energy content values of 122,893 KJ/gallon for gasoline and 138,451 KJ/gallon for diesel by the fuel consumption values presented in this appendix. Bins with a designation of “N.A.” denote that the bins do not exist, according to fleet data used for bins.

**Table E-1: Fuel Consumed During Engine Starts
For Gasoline-Fueled Light-Duty Cars and Trucks (gallons per start)**

Bins for Light-Duty Gasoline-Fueled

Mdl_Yr Range	Displ. Range	No. of Tests	Mean (gal/start)	Std Dev.	S.E. of Mean	90% Confidence Interval	
Pre-1981	< 2.0	444	0.0252	0.0094	0.0004	0.0245	0.0259
	2.1-2.5	212	0.0315	0.0084	0.0006	0.0306	0.0325
	2.6-3.0	73	0.0309	0.0120	0.0014	0.0286	0.0332
	3.1-3.5	89	0.0372	0.0080	0.0008	0.0358	0.0386
	3.6-4.0	225	0.0372	0.0103	0.0007	0.0360	0.0383
	4.1-5.0	368	0.0441	0.0166	0.0009	0.0427	0.0455
	> 5.0	232	0.0512	0.0230	0.0015	0.0487	0.0537
1981-85	< 2.0	1,993	0.0222	0.0086	0.0002	0.0219	0.0225
	2.1-2.5	1,614	0.0257	0.0102	0.0003	0.0253	0.0261
	2.6-3.0	634	0.0310	0.0099	0.0004	0.0304	0.0317
	3.1-3.5	251	0.0389	0.0076	0.0005	0.0381	0.0397
	3.6-4.0	743	0.0351	0.0155	0.0006	0.0341	0.0360
	4.1-5.0	1,029	0.0409	0.0180	0.0006	0.0399	0.0418
	> 5.0	115	0.0510	0.0202	0.0019	0.0479	0.0541
1986-90	< 2.0	1,888	0.0177	0.0078	0.0002	0.0174	0.0180
	2.1-2.5	1,470	0.0210	0.0096	0.0003	0.0206	0.0214
	2.6-3.0	1,031	0.0259	0.0088	0.0003	0.0255	0.0264
	3.1-3.5	170	0.0224	0.0046	0.0004	0.0219	0.0230
	3.6-4.0	443	0.0248	0.0062	0.0003	0.0243	0.0253
	4.1-5.0	645	0.0308	0.0100	0.0004	0.0302	0.0315
	> 5.0	55	0.0359	0.0107	0.0014	0.0335	0.0383
1991+	< 2.0	224	0.0173	0.0066	0.0004	0.0166	0.0180
	2.1-2.5	322	0.0189	0.0048	0.0003	0.0185	0.0194
	2.6-3.0	131	0.0250	0.0073	0.0006	0.0239	0.0260
	3.1-3.5	133	0.0237	0.0071	0.0006	0.0227	0.0247
	3.6-4.0	92	0.0246	0.0058	0.0006	0.0236	0.0256
	4.1-5.0	57	0.0295	0.0084	0.0011	0.0277	0.0313
	> 5.0	19	0.0346	0.0066	0.0015	0.0321	0.0371

**Table E-2: Fuel Consumed During Engine Starts
For Gasoline-Fueled Heavy-Duty Vehicles (gallons per start)**

Bins for Heavy-Duty Gasoline-Fueled

Mdl_Yr Range	Displ. Range	No. of Tests	Mean (gal/start)	Std Dev.	S.E. of Mean	90% Confidence Interval	
Pre-1981	< 2.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.1-2.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.6-3.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.1-3.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.6-4.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	4.1-5.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	> 5.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
1981-85	< 2.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.1-2.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.6-3.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.1-3.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.6-4.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	4.1-5.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	> 5.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
1986-90	< 2.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.1-2.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.6-3.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.1-3.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.6-4.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	4.1-5.0	3	0.0515	0.0086	0.0050	0.0434	0.0597
	> 5.0	68	0.0464	0.0241	0.0029	0.0415	0.0512
1991+	< 2.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.1-2.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.6-3.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.1-3.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.6-4.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	4.1-5.0	6	0.0449	0.0068	0.0028	0.0403	0.0494
	> 5.0	173	0.0460	0.0185	0.0014	0.0437	0.0483

**Table E-3: Fuel Consumed During Engine Starts
For Diesel-Fueled Light-Duty Cars and Trucks (gallons per start)**

Bins for Light-Duty Diesel-Fueled

Mdl_Yr Range	Displ. Range	No. of Tests	Mean (gal/start)	Std Dev.	S.E. of Mean	90% Confidence Interval	
Pre-1981	< 2.0	3	0.0163	0.0039	0.0023	0.0126	0.0201
	2.1-2.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.6-3.0	4	0.0387	0.0063	0.0032	0.0335	0.0439
	3.1-3.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.6-4.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	4.1-5.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	> 5.0	2	0.0358	0.0002	0.0001	0.0356	0.0360
1981-85	< 2.0	6	0.0165	0.0075	0.0031	0.0114	0.0215
	2.1-2.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.6-3.0	9	0.0344	0.0081	0.0027	0.0300	0.0389
	3.1-3.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.6-4.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	4.1-5.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	> 5.0	1	0.0223	N.A.	N.A.	N.A.	N.A.
1986-90	< 2.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.1-2.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.6-3.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.1-3.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.6-4.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	4.1-5.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	> 5.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
1991+	< 2.0	1	0.0075	N.A.	N.A.	N.A.	N.A.
	2.1-2.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.6-3.0	1	0.0221	N.A.	N.A.	N.A.	N.A.
	3.1-3.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.6-4.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	4.1-5.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	> 5.0	0	N.A.	N.A.	N.A.	N.A.	N.A.

**Table E-4: Fuel Consumed During Engine Starts
For Diesel-Fueled Heavy-Duty Vehicles (gallons per start)**

Bins for Heavy-Duty Diesel-Fueled

Mdl_Yr Range	Displ. Range	No. of Tests	Mean (gal/start)	Std Dev.	S.E. of Mean	90% Confidence Interval	
Pre-1981	< 2.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.1-2.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.6-3.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.1-3.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.6-4.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	4.1-5.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	> 5.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
1981-85	< 2.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.1-2.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.6-3.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.1-3.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.6-4.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	4.1-5.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	> 5.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
1986-90	< 2.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.1-2.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.6-3.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.1-3.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.6-4.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	4.1-5.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	> 5.0	2	0.0397	0.0037	0.0026	0.0354	0.0439
1991+	< 2.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.1-2.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	2.6-3.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.1-3.5	0	N.A.	N.A.	N.A.	N.A.	N.A.
	3.6-4.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	4.1-5.0	0	N.A.	N.A.	N.A.	N.A.	N.A.
	> 5.0	29	0.0269	0.0255	0.0047	0.0191	0.0347

Appendix F: CH₄ & N₂O Rates by Model Year

Table F-1: N₂O and CH₄ Emission Rates for Gasoline-Fueled Motorcycles

Fuel Type	Reg Class	Mdl_Yr Group	N2O Running (g/hr)	N2O Start (g/start)	CH4 Running (g/hr)	CH4 Start (g/start)
Gasoline	MC	Pre-73	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1973	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1974	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1975	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1976	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1977	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1978	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1979	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1980	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1981	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1982	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1983	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1984	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1985	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1986	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1987	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1988	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1989	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1990	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1991	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1992	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1993	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1994	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1995	0.131734	0.014912	1.667984	0.032618
Gasoline	MC	1996	0.103147	0.011676	1.243390	0.024315
Gasoline	MC	1997	0.103147	0.011676	1.243390	0.024315
Gasoline	MC	1998	0.103147	0.011676	1.243390	0.024315
Gasoline	MC	1999	0.103147	0.011676	1.243390	0.024315
Gasoline	MC	2000	0.103147	0.011676	1.243390	0.024315
Gasoline	MC	2001-2010	0.103147	0.011676	1.243390	0.024315
Gasoline	MC	2011-2020	0.103147	0.011676	1.243390	0.024315
Gasoline	MC	2021-2050	0.103147	0.011676	1.243390	0.024315

Table F-2: N₂O and CH₄ Emission Rates for Gasoline-Fueled Passenger Cars

Fuel Type	Reg Class	Mdl_Yr Group	N2O Running (g/hr)	N2O Start (g/start)	CH4 Running (g/hr)	CH4 Start (g/start)
Gasoline	LDV	Pre-73	0.250542	0.028361	3.172287	0.062035
Gasoline	LDV	1973	0.250542	0.028361	3.020166	0.059060
Gasoline	LDV	1974	0.250542	0.028361	3.020166	0.059060
Gasoline	LDV	1975	0.555818	0.062917	2.666425	0.018868
Gasoline	LDV	1976	0.574897	0.065077	2.644317	0.016356
Gasoline	LDV	1977	0.574897	0.065077	2.644317	0.016356
Gasoline	LDV	1978	0.593977	0.067236	2.622208	0.013844
Gasoline	LDV	1979	0.593977	0.067236	2.622208	0.013844
Gasoline	LDV	1980	0.625594	0.070815	2.503794	0.013112
Gasoline	LDV	1981	0.784376	0.088789	1.408570	0.030436
Gasoline	LDV	1982	0.786167	0.088992	1.394812	0.030690
Gasoline	LDV	1983	0.789749	0.089397	1.367297	0.031198
Gasoline	LDV	1984	0.811242	0.091830	1.202202	0.034250
Gasoline	LDV	1985	0.811242	0.091830	1.202202	0.034250
Gasoline	LDV	1986	0.811242	0.091830	1.202202	0.034250
Gasoline	LDV	1987	0.811242	0.091830	1.202202	0.034250
Gasoline	LDV	1988	0.811242	0.091830	1.202202	0.034250
Gasoline	LDV	1989	0.811242	0.091830	1.202202	0.034250
Gasoline	LDV	1990	0.811242	0.091830	1.202202	0.034250
Gasoline	LDV	1991	0.811242	0.091830	1.202202	0.034250
Gasoline	LDV	1992	0.811242	0.091830	1.202202	0.034250
Gasoline	LDV	1993	0.811242	0.091830	1.202202	0.034250
Gasoline	LDV	1994	0.600749	0.100218	0.817028	0.042634
Gasoline	LDV	1995	0.390256	0.108606	0.431854	0.051018
Gasoline	LDV	1996	0.284740	0.112137	0.247517	0.054534
Gasoline	LDV	1997	0.279343	0.112016	0.242012	0.054406
Gasoline	LDV	1998	0.249104	0.109853	0.230393	0.052176
Gasoline	LDV	1999	0.193783	0.105323	0.216592	0.047512
Gasoline	LDV	2000	0.130111	0.100116	0.200625	0.042151
Gasoline	LDV	2001-2010	0.008406	0.090150	0.170263	0.031890
Gasoline	LDV	2011-2020	0.008406	0.090150	0.170263	0.031890
Gasoline	LDV	2021-2050	0.008406	0.090150	0.170263	0.031890

Table F-3: N₂O and CH₄ Emission Rates for Gasoline-Fueled Light-Duty Trucks

Fuel Type	Reg Class	Mdl_Yr Group	N2O Running (g/hr)	N2O Start (g/start)	CH4 Running (g/hr)	CH4 Start (g/start)
Gasoline	LDT	Pre-73	0.286861	0.032472	3.632152	0.071028
Gasoline	LDT	1973	0.283361	0.032076	3.415790	0.066797
Gasoline	LDT	1974	0.283361	0.032076	3.415790	0.066797
Gasoline	LDT	1975	0.660396	0.074755	2.804289	0.089675
Gasoline	LDT	1976	0.714259	0.080852	2.716932	0.092943
Gasoline	LDT	1977	0.687327	0.077803	2.760611	0.091309
Gasoline	LDT	1978	0.687327	0.077803	2.760611	0.091309
Gasoline	LDT	1979	0.714259	0.080852	2.716932	0.092943
Gasoline	LDT	1980	0.714259	0.080852	2.716932	0.092943
Gasoline	LDT	1981	0.848572	0.096056	2.475471	0.098129
Gasoline	LDT	1982	0.875161	0.099065	2.408724	0.096777
Gasoline	LDT	1983	0.928338	0.105085	2.275231	0.094074
Gasoline	LDT	1984	0.981516	0.111104	2.141738	0.091371
Gasoline	LDT	1985	1.034694	0.117124	2.008244	0.088668
Gasoline	LDT	1986	1.087871	0.123144	1.874751	0.085965
Gasoline	LDT	1987	1.327171	0.150232	1.274031	0.073802
Gasoline	LDT	1988	1.327171	0.150232	1.274031	0.073802
Gasoline	LDT	1989	1.327171	0.150232	1.274031	0.073802
Gasoline	LDT	1990	1.327171	0.150232	1.274031	0.073802
Gasoline	LDT	1991	1.327171	0.150232	1.274031	0.073802
Gasoline	LDT	1992	1.327171	0.150232	1.274031	0.073802
Gasoline	LDT	1993	1.327171	0.150232	1.274031	0.073802
Gasoline	LDT	1994	1.129403	0.172129	0.906793	0.076410
Gasoline	LDT	1995	0.905046	0.191016	0.606301	0.080370
Gasoline	LDT	1996	0.792867	0.200460	0.456055	0.082350
Gasoline	LDT	1997	0.792867	0.200460	0.456055	0.082350
Gasoline	LDT	1998	0.638985	0.172150	0.406286	0.075078
Gasoline	LDT	1999	0.462021	0.139594	0.349051	0.066715
Gasoline	LDT	2000	0.523574	0.150918	0.368959	0.069624
Gasoline	LDT	2001-2010	0.023458	0.058910	0.207209	0.045990
Gasoline	LDT	2011-2020	0.023458	0.058910	0.207209	0.045990
Gasoline	LDT	2021-2050	0.023458	0.058910	0.207209	0.045990

**Table F-4: N₂O and CH₄ Emission Rates for
Gasoline-Fueled Buses and Heavy-Duty Trucks**

Fuel Type	Reg Class	Mdl_Yr Group	N2O Running (g/hr)	N2O Start (g/start)	CH4 Running (g/hr)	CH4 Start (g/start)
Gasoline	HDT	Pre-73	0.653748	0.074002	8.277566	0.161870
Gasoline	HDT	1973	0.653748	0.074002	8.277566	0.161870
Gasoline	HDT	1974	0.653748	0.074002	8.277566	0.161870
Gasoline	HDT	1975	0.653748	0.074002	8.277566	0.161870
Gasoline	HDT	1976	0.653748	0.074002	8.277566	0.161870
Gasoline	HDT	1977	0.653748	0.074002	8.277566	0.161870
Gasoline	HDT	1978	0.653748	0.074002	8.277566	0.161870
Gasoline	HDT	1979	0.653748	0.074002	8.277566	0.161870
Gasoline	HDT	1980	0.653748	0.074002	8.277566	0.161870
Gasoline	HDT	1981	0.653748	0.074002	8.277566	0.161870
Gasoline	HDT	1982	0.706594	0.079984	8.038779	0.164551
Gasoline	HDT	1983	0.706594	0.079984	8.038779	0.164551
Gasoline	HDT	1984	0.706594	0.079984	8.038779	0.164551
Gasoline	HDT	1985	0.676812	0.076613	7.302683	0.150156
Gasoline	HDT	1986	0.676812	0.076613	7.302683	0.150156
Gasoline	HDT	1987	1.107280	0.125341	6.345122	0.162482
Gasoline	HDT	1988	1.216107	0.137659	5.945032	0.169359
Gasoline	HDT	1989	1.216107	0.137659	5.945032	0.169359
Gasoline	HDT	1990	1.484948	0.168091	5.373339	0.176430
Gasoline	HDT	1991	1.484948	0.168091	5.373339	0.176430
Gasoline	HDT	1992	1.484948	0.168091	5.373339	0.176430
Gasoline	HDT	1993	1.484948	0.168091	5.373339	0.176430
Gasoline	HDT	1994	1.484948	0.168091	5.373339	0.176430
Gasoline	HDT	1995	1.484948	0.168091	5.373339	0.176430
Gasoline	HDT	1996	1.756906	0.345849	1.564622	0.178185
Gasoline	HDT	1997	1.685840	0.383027	0.945586	0.169328
Gasoline	HDT	1998	1.556483	0.397853	0.475571	0.160324
Gasoline	HDT	1999	1.273626	0.345815	0.465752	0.147875
Gasoline	HDT	2000	0.896482	0.276431	0.452659	0.131276
Gasoline	HDT	2001-2010	0.047909	0.120317	0.423200	0.093929
Gasoline	HDT	2011-2020	0.047909	0.120317	0.423200	0.093929
Gasoline	HDT	2021-2050	0.047909	0.120317	0.423200	0.093929

**Table F-5: N₂O and CH₄ Emission Rates for
Diesel-Fueled Passenger Cars**

Fuel Type	Reg Class	Mdl_Yr Group	N2O Running (g/hr)	N2O Start (g/start)	CH4 Running (g/hr)	CH4 Start (g/start)
Diesel	LDV	Pre-73	0.027897	0	0.023579	0
Diesel	LDV	1973	0.027897	0	0.023579	0
Diesel	LDV	1974	0.027897	0	0.023579	0
Diesel	LDV	1975	0.027897	0	0.023579	0
Diesel	LDV	1976	0.027897	0	0.023579	0
Diesel	LDV	1977	0.027897	0	0.023579	0
Diesel	LDV	1978	0.027897	0	0.023579	0
Diesel	LDV	1979	0.027897	0	0.023579	0
Diesel	LDV	1980	0.027897	0	0.023579	0
Diesel	LDV	1981	0.027897	0	0.023579	0
Diesel	LDV	1982	0.027897	0	0.023579	0
Diesel	LDV	1983	0.023538	0	0.019895	0
Diesel	LDV	1984	0.023538	0	0.019895	0
Diesel	LDV	1985	0.023538	0	0.019895	0
Diesel	LDV	1986	0.023538	0	0.019895	0
Diesel	LDV	1987	0.023538	0	0.019895	0
Diesel	LDV	1988	0.023538	0	0.019895	0
Diesel	LDV	1989	0.023538	0	0.019895	0
Diesel	LDV	1990	0.023538	0	0.019895	0
Diesel	LDV	1991	0.023538	0	0.019895	0
Diesel	LDV	1992	0.023538	0	0.019895	0
Diesel	LDV	1993	0.023538	0	0.019895	0
Diesel	LDV	1994	0.023538	0	0.019895	0
Diesel	LDV	1995	0.023538	0	0.019895	0
Diesel	LDV	1996	0.023028	0	0.019464	0
Diesel	LDV	1997	0.023028	0	0.019464	0
Diesel	LDV	1998	0.023028	0	0.019464	0
Diesel	LDV	1999	0.023028	0	0.019464	0
Diesel	LDV	2000	0.023028	0	0.019464	0
Diesel	LDV	2001-2010	0.023028	0	0.019464	0
Diesel	LDV	2011-2020	0.023028	0	0.019464	0
Diesel	LDV	2021-2050	0.023028	0	0.019464	0

**Table F-6: Proposed N₂O and CH₄ Emission Rates
For Diesel-Fueled Light-Duty Trucks**

Fuel Type	Reg Class	Mdl_Yr Group	N2O Running (g/hr)	N2O Start (g/start)	CH4 Running (g/hr)	CH4 Start (g/start)
Diesel	LDT	Pre-73	0.036326	0	0.030703	0
Diesel	LDT	1973	0.036326	0	0.030703	0
Diesel	LDT	1974	0.036326	0	0.030703	0
Diesel	LDT	1975	0.036326	0	0.030703	0
Diesel	LDT	1976	0.036326	0	0.030703	0
Diesel	LDT	1977	0.036326	0	0.030703	0
Diesel	LDT	1978	0.036326	0	0.030703	0
Diesel	LDT	1979	0.036326	0	0.030703	0
Diesel	LDT	1980	0.036326	0	0.030703	0
Diesel	LDT	1981	0.036326	0	0.030703	0
Diesel	LDT	1982	0.036326	0	0.030703	0
Diesel	LDT	1983	0.031443	0	0.026576	0
Diesel	LDT	1984	0.031443	0	0.026576	0
Diesel	LDT	1985	0.031443	0	0.026576	0
Diesel	LDT	1986	0.031443	0	0.026576	0
Diesel	LDT	1987	0.031443	0	0.026576	0
Diesel	LDT	1988	0.031443	0	0.026576	0
Diesel	LDT	1989	0.031443	0	0.026576	0
Diesel	LDT	1990	0.031443	0	0.026576	0
Diesel	LDT	1991	0.031443	0	0.026576	0
Diesel	LDT	1992	0.031443	0	0.026576	0
Diesel	LDT	1993	0.031443	0	0.026576	0
Diesel	LDT	1994	0.031443	0	0.026576	0
Diesel	LDT	1995	0.031443	0	0.026576	0
Diesel	LDT	1996	0.032094	0	0.027127	0
Diesel	LDT	1997	0.032094	0	0.027127	0
Diesel	LDT	1998	0.032094	0	0.027127	0
Diesel	LDT	1999	0.032094	0	0.027127	0
Diesel	LDT	2000	0.032094	0	0.027127	0
Diesel	LDT	2001-2010	0.032094	0	0.027127	0
Diesel	LDT	2011-2020	0.032094	0	0.027127	0
Diesel	LDT	2021-2050	0.032094	0	0.027127	0

**Table F-7: Proposed N₂O and CH₄ Emission Rates
For Diesel-Fueled Buses and Heavy-Duty Trucks**

Fuel Type	Reg Class	Mdl_Yr Group	N2O Running (g/hr)	N2O Start (g/start)	CH4 Running (g/hr)	CH4 Start (g/start)
Diesel	HDT	Pre-73	0.095981	0	0.081124	0
Diesel	HDT	1973	0.095981	0	0.081124	0
Diesel	HDT	1974	0.095981	0	0.081124	0
Diesel	HDT	1975	0.095981	0	0.081124	0
Diesel	HDT	1976	0.095981	0	0.081124	0
Diesel	HDT	1977	0.095981	0	0.081124	0
Diesel	HDT	1978	0.095981	0	0.081124	0
Diesel	HDT	1979	0.095981	0	0.081124	0
Diesel	HDT	1980	0.095981	0	0.081124	0
Diesel	HDT	1981	0.095981	0	0.081124	0
Diesel	HDT	1982	0.095981	0	0.081124	0
Diesel	HDT	1983	0.095981	0	0.081124	0
Diesel	HDT	1984	0.095981	0	0.081124	0
Diesel	HDT	1985	0.095981	0	0.081124	0
Diesel	HDT	1986	0.095981	0	0.081124	0
Diesel	HDT	1987	0.095981	0	0.081124	0
Diesel	HDT	1988	0.095981	0	0.081124	0
Diesel	HDT	1989	0.095981	0	0.081124	0
Diesel	HDT	1990	0.095981	0	0.081124	0
Diesel	HDT	1991	0.095981	0	0.081124	0
Diesel	HDT	1992	0.095981	0	0.081124	0
Diesel	HDT	1993	0.095981	0	0.081124	0
Diesel	HDT	1994	0.095981	0	0.081124	0
Diesel	HDT	1995	0.095981	0	0.081124	0
Diesel	HDT	1996	0.095981	0	0.081124	0
Diesel	HDT	1997	0.095981	0	0.081124	0
Diesel	HDT	1998	0.095981	0	0.081124	0
Diesel	HDT	1999	0.095981	0	0.081124	0
Diesel	HDT	2000	0.095981	0	0.081124	0
Diesel	HDT	2001-2010	0.095981	0	0.081124	0
Diesel	HDT	2011-2020	0.095981	0	0.081124	0
Diesel	HDT	2021-2050	0.095981	0	0.081124	0

Table F-8: N₂O and CH₄ Emission Rates for Alternative-Fueled Vehicles

Fuel Type	Reg Class	Mdl_Yr Group	N2O Running (g/hr)	N2O Start (g/start)	CH4 Running (g/hr)	CH4 Start (g/start)
CNG	LDV	1996-2000	1.083132	0.427926	11.116180	2.563660
CNG	LDV	2001-2010	1.083132	0.427926	11.116180	2.563660
CNG	LDT	1996-2000	1.083132	0.427926	11.116180	2.563660
CNG	LDT	2001-2010	1.083132	0.427926	11.116180	2.563660
CNG	Buses (HDV)	1985-1988	1.552808	0.613488	151.004912	34.825390
CNG	HDVs (ALL)	1989-2000	2.846815	1.124727	117.109077	27.008190
CNG	HDVs (ALL)	2001-2010	2.846815	1.124727	117.109077	27.008190
LPG	LDV	1982-1984	1.456956	0.575618	7.406733	1.708172
LPG	LDV	2001-2010	1.456956	0.575618	7.406733	1.708172
LPG	LDV	1977-1989	1.456956	0.575618	7.406733	1.708172
LPG	LDT	2001-2010	1.456956	0.575618	7.406733	1.708172
LPG	HDVs	1979-2000	1.437785	0.568044	1.313509	0.302927
LPG	HDVs	2001-2010	1.437785	0.568044	1.313509	0.302927
Ethanol	HDVs	1989-1993	3.489026	1.378453	25.285053	5.831346
Ethanol	HDVs	2001-2010	3.489026	1.378453	25.285053	5.831346
Methanol	HDVs	1992	2.079996	0.821770	7.856731	1.811952
Methanol	HDVs	2001-2010	2.079996	0.821770	7.856731	1.811952
Electric	ALL	ALL	0	0	0	0

Appendix G: Comments on Emission Analysis Plan

In the fall of 2002 EPA published planning documentation for MOVES in the form of two reports: “Draft Design and Implementation Plan for EPA’s Multi-Scale Motor Vehicle & Equipment Emission System (MOVES)”, and “Emission Analysis Plan for MOVES GHG”. These reports subsequently underwent two paths of review: public review, in which comments were solicited from model stakeholders, and independent formal peer review, in which comments were solicited from peer reviewers paid by the Agency according to Agency peer review guidelines.

This appendix provides a summary of written comments pertaining to energy and emission analysis issues received as a result of both review paths (comments pertaining to the design plan are responded to in a separate document, the MOVES2004 Software Design Reference Manual). The commenter of a specific comment is identified in parentheses, identified according to the bolded designation in the list below. Each comment number captures a unique sentiment, although variation on the general idea may have been submitted by several commenters as reflected in commenter identifications.

Comments were received from the following parties:

Public Review:

- Alliance of Automobile Manufacturers (**AAM**)
 - Review prepared by Tom Darlington, AIR
- Engine Manufacturer’s Association (**EMA**)
 - Review prepared by Tom Darlington, AIR
- David **Roden**, AECOM Consult
 - Review prepared for U.S. DOT with respect to MOVES design applicability to TRANSIMS
- **U.S. DOT** – FHWA
- Peter **McClintock**, Applied Analysis
 - written comments in response to 11/2002 workshop
- Wayne Elson, **EPA Region 10**
- Donald **Stedman**, Professor, University of Denver
- **Natural Propane Gas Association / Propane Vehicle Council**
- Phyllis Jones, **North Carolina DAQ**

Formal Peer Review:

- Administered by Southwest Research Institute
- Reviewers:
 - Marc **Ross**, Professor Emeritus, University of Michigan
 - Ted **Russell**, Professor, Georgia Tech
 - Michael **Replogle**, Transportation Director, Environmental Defense
 - Janet **Buckingham**, Principle Analyst, Southwest Research Institute (Emission Analysis Plan only)

Comments are subdivided into subareas as appropriate.

Fuel-based emission rates and inventories:

1. **Comment:** Producing fuel-based emission estimates will be useful (Russell). Fuel-based emission estimates will reduce variability (Stedman).

***EPA Response:** These comments are more relevant to the yet-to-be-developed criteria pollutant version of MOVES, but as MOVES2004 estimates energy consumption (easily converted to fuel consumption) it does provide a foundation for investigating this approach.*

2. **Comment:** An alternative method for generating emission inventories would be to use fuel sales rather than VMT, and remote sensing device (RSD) data used for direct emission factors. The proposed “bottom-up” approach should be checked against top-down fuel sales data (Stedman, Russell, U.S. DOT)

***EPA Response:** The primary purpose for developing MOVES2004 as an energy consumption model is to make the comparison of bottom-up fuel consumption to top-down fuel sales, as suggested. The results of this comparison will be published in a separate document, “MOVES2004 Validation Results”. As mentioned under comment 1, MOVES2004 could provide the foundation for the inventory methodology provided, but the efficacy of this approach would need to be investigated further. One drawback of the suggested approach is it does not allow any breakdown of emission inventory to sub-regional levels, e.g. roadway type.*

Binning/analysis approach:

3. **Comment:** The binning approach is more problematic for some variables than others. VSP binning (and the use of VSP as the driving variable that determines emissions) is a good approach, but vehicle odometer would be better served through a linear function (Ross, Russell). Linear functions are preferable to binning, particularly for uncertainty estimation (Russell)

***EPA Response:** The main benefit of the binning approach is that it allows easier processing of raw test data into emission rates, which serves the use case of having a model which can be easily updated based on new data. While we think this approach has proven feasible, we can see the merit of using functions instead of bins from the standpoint of model runtime performance, so it is something we would consider as model development progresses. We agree some variables can't be binned, hence the inclusion of adjustment factors for temperature and air conditioning.*

4. **Comment:** Additional binning by speed, engine displacement and odometer is recommended, as is investigation of other variables (Replogle). Where other operating variables also examined besides cycle speed (Buckingham)?

EPA Response: As detailed in this report, we did end up binning by speed and engine displacement; odometer doesn't turn up as very important with regard to fuel consumption, but clearly some mechanism (or vehicle age) will need to be added for criteria pollutants. We did investigate other variables than cycle speed (and ended up choosing instantaneous speed), as discussed in Appendix A.

5. **Comment:** Allowing source bins to vary by pollutant is supported, to allow MOVES to be credible across a wide range of applications (Replogle)

EPA Response: none required – this supports our current design.

6. **Comment:** Direct measurement of acceleration should be used in VSP calculation, rather than indirect methods (Ross). Using an aggregate statistic for acceleration/deceleration rates (e.g. “jerk”) appears warranted, particularly for HC (Replogle)

EPA Response: Most test programs don't include a direct measure of acceleration, other than the difference in speed measurement from one second to the next. We're not sure how much this would improve the precision of the model. We do agree that the binning approach for HC might be improved, but in general HC emissions are so varied (especially second-by-second) there may always be more uncertainty than with other pollutants.

7. **Comment:** Evaluations presented in emission analysis plan do not consider sensitivity to changes in control systems over the years. The approach should be analyzed by comparing two or more datasets with vehicles that use control technologies of different vintages (Ross). The proposal should be evaluated on second-by-second data from for non-Tier 1 light duty gasoline vehicles, light-duty gasoline trucks, light-duty diesel trucks, heavy-duty gasoline and heavy-duty diesel vehicles (AAM, EMA).

EPA Response: The proof-of-concept binning evaluation presented in Appendix A was designed to address this comment – we purposely included trucks, and a range of model years and control technologies, in the light-duty evaluation, and extended to analysis to include heavy-duty diesel. We feel the binning approach proved effective across the range of these vehicles.

8. **Comment:** EPA should not rely on CO₂ performance in the selection of an appropriate emissions modeling approach for any group of vehicles (AAM, EMA)

EPA Response: To address this comment, the proof-of-concept binning evaluation focused not only on CO₂ (fuel), but HC, CO and NO_x as well.

9. **Comment:** The proposed binning approach based on VSP was based on Tier 1 vehicles and may not be appropriate for LEV or Tier 2 vehicles. Tier 2 vehicles no longer differentiate based on weight class, so the MOVES design should ensure that emissions are equivalent for all passenger vehicle weight categories. (AAM, EMA)

EPA Response: *We don't think anything about newer technology vehicles will affect the appropriateness of the binning approach, but can investigate it as data of these vehicles becomes available. For energy consumption, the weight classes are for loaded weight, and the importance of this on fuel consumption won't change for Tier 2. Accounting for the provisions of Tier 2 will be handled in the criteria model.*

10. **Comment:** Concern that large number of VSP bins brought on by criteria that no more than 10 percent of total emissions be allowed in a given bin may have sparse emission data; was analysis performed with looser cutpoint (e.g. 15 percent)? (Buckingham)

EPA Response: *The original analysis didn't consider a looser cutpoint, but the final binning analysis presented in Appendix A did relax this criteria, and some of the bins account for more than 10 percent of the time.*

11. **Comment:** EPA has not adequately characterized which variables are most important in explaining variability in HC, CO, and NO_x emissions. These analyses should be conducted on test data from a variety of vehicle types and technologies before selecting a binning strategy for each of these pollutants.

EPA Response: *We still consider everything we've published on HC, CO and NO_x as preliminary, although the binning analysis did consider a broader range of vehicles and vehicle operation, with good result.*

12. **Comment:** The "Modeling Dataset" used by NCSU uses an uneven mixture of data from onboard vehicles, twin-roll dynamometers, and single-roll dynamometers. Due to the limitations of twin-roll dynamometers and onboard instrumentation, and differences in time-alignment methods, the quality of the second-by-second data is suspect (AAM, EMA)

EPA Response: *We did substantial quality checking of second-by-second data, including time alignment where necessary, as discussed in Sections 3 and 4 of this report.*

13. **Comment:** NCSU should have done an apples-to-apples comparison by comparing the same vehicles driven on different cycles rather than validating on an independent sampling of vehicles and cycles. In addition, NCSU used the same estimates for frontal area and rolling resistance for all vehicles in the "Modeling Dataset". The emissions correlations could have been improved by using vehicle-specific estimates for these inputs (AAM, EMA).

EPA Response: *The binning analysis from Appendix A was designed to address this comment – we did the "apples-to-apples" comparison suggested, to remove the added uncertainty of vehicle-to-vehicle variability. We also used vehicle specific estimates for road load coefficient as suggested.*

14. **Comment:** How will EPA estimate VSP coefficients by vehicle use type, and how will uncertainty in these values affect uncertainty of the emission results (AAM, EMA).

EPA Response: The method for doing this in the development of emission rates is discussed in Appendix B. An algorithm relating road load coefficients to vehicle weight was required, since reliable data generally isn't available from test programs.

Use of Physical Model (PERE)

15. **Comment:** The use of both physically and empirically-based analysis is supported. Both have advantages and limitations which can be traded off (Replogle, Ross)

EPA Response: no response required

16. **Comment:** Engine-out emissions have become relatively stable in recent years, so modeling engine-out distinctly from tailpipe could be useful for modeling emission variability and modeling new technologies (Ross)

EPA Response: PERE employs this approach.

17. **Comment:** Calibration of physical model to empirical data should include inventory-weighted comparison (Russell). The method for calibrating the physical model has not been adequately described and requires more detail (Buckingham). Physical model provides more opportunity for “gaming”, which should be managed by calibration process (Replogle).

EPA Response: Any physical model developed will be calibrated to known data, whether that data is from chassis dynamometer, on-road, remote sensing, or some combination has yet to be determined.

18. **Comment:** Given the number of source and operating mode bins it would be helpful to estimate how many might be flagged as not having enough data (i.e. less than 40 seconds based on the criteria presented in the emission analysis plan) (Buckingham). Not enough analysis presented on the 40 second criteria (AAM, EMA)

EPA Response: This quality check has been done and is described in some of the MOVES documentation. The statistical criteria for bin adequacy has been expanded.

19. **Comment:** What method for estimating emission rates will be used for bins with little data, but no bag data (Buckingham)? EPA currently does not plan to use the MSOD bag data (other than to “calibrate” the PERE), a huge source of EPA data which has been relied upon in developing the MOBILE series of models. EPA should explore more options for using these data to “anchor” the MOVES model, especially for older vehicles (AAM, EMA).

EPA Response: *Hole-filling methods are addressed in Section 4. It is true that for the binning approach, bag data is not used directly, except to calibrate PERE; although bag data is used for generated adjustment factors, start rates, and CH₄ / N₂O emission rates. We will investigate whether we need to employ bag data for criteria pollutants.*

20. **Comment:** EPA plans to use the PERE model to fill “data-gaps.” The development work on PERE has focused on warmed-up operation from Tier 1 light duty vehicles only; therefore, this work has not progressed to the point where a decision can be made to rely on this model for this purpose (AAM, EMA).

EPA Response: *The handling of cold start in PERE is discussed further in the PERE documentation.*

Time Weighting of Emissions:

21. **Comment:** Weighting by time rather than by vehicle will reduce the influence of high emitters with small amounts of data (i.e. RSD) (Stedman, Ross). The one-second weighting approach proposed for MOVES is accepted as valid (Buckingham, Replogle, Russell)

EPA Response: *MOVES takes the time-weighted approach, so these comments support the chosen approach*

Age and Odometer:

22. **Comment:** Age is a better indicator of vehicle deterioration than odometer and should be considered in MOVES (McClintock). The correlation between odometer and age is not static, and to assume so is a considerable source of uncertainty (Stedman)

EPA Response: *This is more relevant to criteria pollutants, and we will be looking into the question of age versus odometer.*

23. **Comment:** More emphasis should be placed on older vehicles. It would be challenging to do an analysis of policies addressing older vehicles in MOVES unless the user provided data directly. Information on older vehicles will tend to be inaccurate and location-dependent (Ross)

EPA Response: *This is more relevant to criteria pollutants, and we will be looking into the question of age versus odometer.*

Uncertainty and Sensitivity:

24. **Comment:** In general the inclusion of uncertainty is supported (Russell, EPA Region 10). How will uncertainty be used in a policy context, i.e. the application of uncertainty tolerances in determining conformity? (U.S. DOT)

EPA Response: *We are still working out how to include uncertainty in MOVES – it is NOT included in MOVES2004. As discussed in the MOVES2004 Software Design Reference Manual, we have investigated both a propagation of error method and Monte Carlo simulation. We believe the latter is most feasible for the MOVES design, but need to consider it would be implemented in terms of model runtime performance. The MOVES database does have placeholders for coefficient of variation (CV), and these were calculated by the binner program and used to populate the EmissionRate table. The policy context of uncertainty would need to be addressed with official guidance.*

25. **Comment:** The equation given for calculating uncertainty is an approximation; a more pure form would explicitly account for uncertainty and sensitivity of each term (Russell)

EPA Response: *We would attempt to do this with a Monte Carlo simulation approach.*

26. **Comment:** Monte Carlo methods for generating uncertainty are time consuming; analytical methods as proposed in the emission analysis plan are faster and can identify sensitivities (Buckingham, Russell). Use of such methods requires an assumption of normality, but this is appropriate (Buckingham).

EPA Response: *We investigated propagation of error, but found it impractical due to certain aspects of the MOVES design. Model runtime performance is the biggest concern with Monte Carlo.*

27. **Comment:** MOVES does not address location dependence of activity or emission information well; specifically, operating mode distributions, information on older vehicles, and vehicle age distribution (Ross). Uncertainty of vehicle activity and characteristic components are not considered, and could be considerable, resulting in serious underestimation of overall uncertainty (AAM, Replogle). Will MOVES consider uncertainty of the model computation process and the input data supplied by the user? (U.S. DOT)

EPA Response: *While local areas can certainly customize any input if data is available, we will not have the resources to develop location-specific defaults – we focus on national average defaults. Although uncertainty is not employed in MOVES2004 we have designed the database to include Coefficient of Variation (CV) for all inputs, including activity. Thus uncertainty of input data could be accounted for, but not the model computation process.*

28. **Comment:** The uncertainty correction for different averaging times proposed in the emission analysis plan appears justified and reasonable (Buckingham)

EPA Response: *no response necessary*

29. **Comment:** EPA should perform a comprehensive sensitivity analysis for the model and publish the results so this task is not left to the users (U.S. DOT)

EPA Response: *We do plan to perform sensitivity analysis, but it will be after the model is made public.*

High Emitters and Tampered Vehicles:

30. **Comment:** Representative vehicle recruitment will continue to be an issue (Buckingham). The success of using a continuous distribution to define variability in fleet will depend largely on the ability to recruit high emitters (Ross). EPA voluntary programs and IM programs tend to leave out the grossest emitters (Stedman). How will high emitter data be selected, collected and included in bins (U.S. DOT)?

EPA Response: *We will be investigating this issue as we develop the criteria pollutant version of MOVES; the issue of representative sampling is long-standing issue for MOBILE as well as MOVES, and we plan to look at IM program and RSD data to help determine how to best represent the in-use fleet.*

31. **Comment:** The single distribution method is not recommended for characterizing high emitters, because the shape can easily be skewed, masking the true variation of normal and high emitting groups. Malfunction categories may not be easy to implement. The discrete emitter category approach is recommended (Buckingham).

EPA Response: *We will be investigating this issue as we develop the criteria pollutant version of MOVES.*

32. **Comment:** EPA's proposed methods of calibrating emitter distributions in MOVES to local data contradicts its preliminary decisions on the use of IM240 and remote sensing data in developing emission rates for VSP bins (AAM, EMA).

EPA Response: *We didn't use IM240 (program) or RSD data for MOVES2004, but will be looking at it as we develop the criteria pollutant version of MOVES.*

33. **Comment:** Treatment of tampering needs to be more explicitly addressed; new surveys are needed (EPA Region 10)

EPA Response: *We will be investigating this issue as we develop the criteria pollutant version of MOVES.*

34. **Comment:** The "unrepresented" distribution method is preferable because it captures I/M and other strategies as well as local data (NC DAQ)

EPA Response: *We will be investigating this issue as we develop the criteria pollutant version of MOVES.*

On-Board Measurement (PEMS):

35. **Comment:** EPA should not use PEMS until data accuracy issues are resolved (AAM, EMA)

EPA Response: We didn't end up using PEMS data directly in MOVES2004 (it was used in proof-of-concept evaluations, e.g. Appendix A, but not the model itself); however, we believe the shootout program and PEMS programs conducted since have demonstrated the accuracy of the approach.

36. **Comment:** On-Board measurement may be costly and inefficient; flexibility should be retained in data collection (Ross)

EPA Response: As we move into more on-board data collection, we would likely supplement data analysis with other sources of data to ensure representativeness. Ultimately we believe PEMS will be much more cost effective and efficient than laboratory testing, especially for heavy duty vehicles.

Data Collection

37. **Comment:** Data proposed for MOVES GHG appears to be a good cross section of labs, companies and studies with no bias toward a particular study. The EPA MSOD format and a definitive set of units is suggested as a requirement for future studies to reduce the merging effort. Is there a recommended method for cross-checking data quality (Buckingham)?

EPA Response: We did do substantial QA on the data, as described in Section 3 and 4. We do envision compiling our data requirements for the benefit of other test organizations, so that their data might be more easily incorporated into MOVES.

38. **Comment:** IM240 data should not be used to determine VSP bin emission rates in MOVES until the uncertainties in fuels, test temperatures, and vehicle preconditioning are resolved or addressed. (AAM, EMA). IM240 data has fuel inconsistencies but would be sufficient as a validation tool for CO₂ data. (NC DAQ)

EPA Response: We did not use IM240 program data for MOVES2004, and agree that these issues would need to be addressed in order to include it in MOVES in the future.

39. **Comment:** Use of the US06 cycle for model and inventory development, including its use for the calibration of the PERE, is inappropriate. The US06 cycle is a severe cycle and is not representative of real-world driving. We believe that the use of the US06 driving cycle in the development of MOVES would affect inventory development assessments performed using the model. (AAM, EMA)

EPA Response: *While we did exclude certain cycles in the development of MOVES (Section 4), we decided to keep the US06 data because it contains real-world aggressive driving not found on other cycles.*

40. **Comment:** Some vehicle classes/ages are poorly represented in the Mobile Source Observational Database (MSOD). For example, the entire Tier 1 national vehicle population over 50,000 miles is represented by only 50 vehicles, although they were tested multiple times. In addition, the MSOD apparently does not include any second-by-second data for NLEV or Tier 2 vehicles. When will this data be incorporated into the MSOD? (U.S. DOT)

EPA Response: *We added to the Tier 1 and NLEV dataset with the data sources in Section 4, primarily the New York IPA dataset. Tier 2 vehicles are just starting to come on the market, so it will likely be the next iteration of MOVES before we have in-use data to analyze.*

41. **Comment:** Concern with using limited dataset of diesel buses to represent heavy-duty vehicles (EMA)

EPA Response: *We added to the heavy-duty dataset with the data sources in Section 4, primarily the WVU, CRC E-55 and CE-CERT datasets.*

Remote Sensing Data

42. **Comment:** RSD can be helpful in “filling the gaps” for determining emitter distributions and deterioration with age and odometer (McClintock). Only RSD captures the gross emitters (Stedman). Remote sensing should be used if not directly, at least as a verification step (Russell).

EPA Response: *We will be investigating this issue as we develop the criteria pollutant version of MOVES.*

43. **Comment:** RSD sampling is not akin to 1 second on on-board data, because the vehicle is moving relatively rapidly in the course of a 1 second RSD measurement (Ross)

EPA Response: *We will be investigating this issue as we develop the criteria pollutant version of MOVES.*

Fuel Effects

44. **Comment:** Oxygenate is important for fuel economy (Stedman)

EPA Response: *On an energy consumption basis there is no differentiation between oxygenate use; it can be accounted for in post-processing conversion to energy or fuel.*

45. **Comment:** Binning fuel parameters is not a good approach. It would result in a very high number of bins; and emission response should be continuous. A better approach is to use continuous corrections within each bin (Russell)

***EPA Response:** We agree; fuel effects will be treated as a continuous function in future versions of MOVES.*

46. **Comment:** MOVES should add the functionality to deal with fuel properties and reformulation issues, i.e. as the EPA Complex Model does (Russell)

***EPA Response:** We will be investigating this issue as we develop the criteria pollutant version of MOVES.*

47. **Comment:** MOVES GHG should include LPG as a fuel type (Natural Propane Gas Association / Propane Vehicle Council)

***EPA Response:** LPG is included as a fuel type in MOVES2004*

Methane

48. **Comment:** High emitters drive methane, and should be accounted for in MOVES (Stedman). Proposed methane approach likely too simplistic, should add complexity of CO₂ or criteria pollutants (Russell).

***EPA Response:** We decided to take the simplistic approach for now, since there is relatively little CH₄ data available for stand-alone analysis. A more complex approach is merited when we add HC to MOVES, to take advantage of the correlation between total HC and CH₄.*

Temperature Effects

49. **Comment:** Low and high temperature effects should be included in the analysis (Russell)

***EPA Response:** We did include temperature effects, as discussed in Section 9*

Model Accuracy

50. **Comment:** For binning analysis, CO₂ should be able to be modeled within 10 percent in average driving; the possibility of achieving anything near this level of accuracy for the criteria pollutants is unclear (Ross). Validation results of VSP and average speed bins presented in emission analysis plan (e.g. prediction of Bag 3 emissions within 10 percent, prediction of most UCC cycles within 10-20 percent) were quite good (Buckingham).

EPA Response: *We met this criteria in the revised binning analysis presented in Appendix A*

51. **Comment:** The validation results performed by EPA for HC, CO, and NO_x of the UCC dataset (light duty vehicles only)—the best validation test of the VSP approach provided by the EPA—indicate that there are serious concerns with using VSP and speed to characterize HC, CO, and NO_x (AAM, EMA).

EPA Response: *We think the binning analysis (Appendix A) demonstrates that the binning approach can perform well for HC, CO and NO_x as well as fuel consumption.*

General Emission-Related Comments

52. **Comment:** EPA's efforts in improving emission for MOVES should be as robust for nonroad sources as for on-road sources (AAM, EMA)

EPA Response: *This is the plan, based on large-scale data collection of nonroad data; but the specific implementation details will need to wait until off-road is implemented in MOVES, currently planned for 2007*

53. **Comment:** Pre-Tier 0 vehicles should be split up to account for non-catalyst controls implemented in the 1970's (AAM, EMA)

EPA Response: *We think the binning analysis (Appendix A) demonstrates that the binning approach can perform well for HC, CO and NO_x as well as fuel consumption.*

54. **Comment:** Will EPA consider incorporating fugitive dust in the PM emission factor estimation process? (U.S. DOT) It would be worth considering (Replogle)

EPA Response: *We will be investigating this issue as we develop the criteria pollutant version of MOVES.*

Appendix H: Pre-Publication Peer Review Comments

Professor Lawrence Caretto of California State University Northridge was contracted to provide formal peer review on a pre-publication version of this document. His comments are included verbatim in this Appendix. Responses to substantive (i.e. non-editorial) comments have been added following each comment, in italics to differentiate it from the original comments. Editorial changes suggested by Dr. Caretto have generally been made in the final version of this report.

REPORT REVIEW

SEPTEMBER 13, 2004 DRAFT EPA REPORT: "MOVES2004
ENERGY AND EMISSION INPUTS"

Review Prepared for U. S. Environmental Protection Agency
Office of Transportation and Air Quality
Assessment and Standards Division

OCTOBER 17, 2004

L. S. CARETTO
7805 Cowper Avenue
West Hills, CA 91304

Introduction

The U. S. Environmental Protection Agency (EPA) is developing a new computer model, known as MOVES, to model emissions from mobile sources. They have prepared a series of reports describing the development of this model. This review discusses the report “MOVES2004 Energy and Emission Inputs” that describes the data that are inputs for the model, including the procedures used to fill missing data items.

In requesting this review,^c EPA stated its primary interests as (1) report clarity, (2) overall methodology, (3) appropriateness of the data sets selected, (4) data analyses conducted, including the statistical approaches used, the models selected, and appropriateness of the resulting conclusions. EPA also asked the following specific questions:

- Are variables important to energy consumption or GHG emission production captured, given available data?
- Has the effectiveness of the modal binning structure chosen for characterizing emission rates been adequately demonstrated with regard to greenhouse gas pollutants and HC, CO and NOx?
- Are methods for “hole filling” using PERE and interpolation/copying sound?
- Is the method for generating advanced technology rates using PERE sound?

The following discussion section of this review first considers the five items listed as “primary interests” by EPA. This is followed by answers to the questions in the bulleted list above. The final part of the discussion section contains miscellaneous comments on the report, mainly technical questions.

For the most part, the comments in this review deal with changes to the report that can readily be done. The comments that refer to improvements that are more exploratory in nature or depend on data not available to EPA are identified as such when they are made.

^c Letter from Chester J. France to Larry Caretto, September 16, 2004.

Discussion

Scope of the report

The report that is the subject of this review describes the techniques used to provide the data the MOVES model requires for energy consumption rates and emission rates of CO₂, CH₄ and N₂O for a range of vehicles and operating conditions that encompass all activities of the onroad vehicle fleet. Emission rates for CO₂, the main greenhouse gas, are determined from energy consumption rates. Thus, most of the focus of the report is on obtaining energy-rate data and making suitable approximations where such data are missing.

The approach of the MOVES model is to assign data to bins. There are two kinds of bins. Source bins describe the kinds of vehicles (cars, trucks, etc.) and basic vehicle parameters used to classify their emissions performance (weight, engine displacement, fuel type, etc.). Operating condition bins are defined in terms of vehicle specific power (VSP) and vehicle speed. The set of operating condition bins describes all possible emission rates (and energy use) for a vehicle in a source bin. The set of source bins, appropriately weighted, will describe the entire vehicle fleet.

The report really deals with two separate but related topics. The first is the development of specific input data for greenhouse gases. The second is a more general analysis of the binning approach for all modeled quantities, and an extension of previous work to use operating condition bins that are defined by a combination of vehicle specific power (VSP) and vehicle speed. This second topic, which is more fundamental to the overall development of MOVES, deals not only with greenhouse gas elements, but also with criteria pollutants.

Where existing data do not provide information for one or more bins, the necessary bin information is found by interpolation or by modeling using a model called PERE (physical emission rate simulator). The approaches described here for greenhouse gases can also be applied to some degree to other pollutants. However, the use of PERE for modeling criteria pollutant rates is not discussed in the report being reviewed.

Comments on areas of primary interest to EPA

Report clarity

As noted in the Scope section above, the report deals with both general and specific topics. The general topics of how to bin the existing data sets and what to do when there are bins that have no data have implications beyond the present report. The specific topics of obtaining the data for energy rates and greenhouse gas emissions provide examples of how these general topics might be addressed in extending MOVES to criteria pollutants. The introduction to the report should point out the significance of general topics that might have applications in MOVES beyond the establishment of data for greenhouse gas emissions.

This report, like other reports on the details of emission model development, is aimed at a limited audience with the technical background necessary to understand the information presented. Although it potentially has reasonable clarity, the draft provided for review appears to have been hastily assembled to meet a deadline. There are some instances of internal authors' notes for items to be added or revised in a final draft. There are also incomplete cross references to other parts of the text (such as "see section 3-?") and simple errors that could be found by using the spelling and grammar checker. Some basic editorial changes have been made to the soft copy sent to this reviewer. This edited version has been sent to EPA as an email attachment.

The differences in colors used in some charts in the report, particularly those in Appendix C, did not show up well in the printed version. Uses of color should be augmented by different fill effects or other features to clarify charts for readers of a version that is not printed in color. Figure C-9 is a good example of this practice. This figure uses a dashed red line and a solid black line so that the differences between these two lines clearly show up on the black-and-white copy.

Some additional comments on report clarity are included in the Miscellaneous Comments section starting on page 169.

Overall methodology

There are two overall methodologies used in the report. The first is the modification of earlier binning strategies to use a combination of vehicle speed and VSP to define operating mode bins. The second is the development of the MOVES data set for greenhouse gases. The latter requires methods for data analysis and selection, placing data in bins, and filling missing bin data using the physical emission rate estimator (PERE) or interpolation.

The report describes the quality control methods used for the data that are eventually placed in bins. The procedures for data quality control appear sound, and the fact that quality controls were used to ensure the accuracy of the data in the model is perhaps more important than the specific details of the quality-control procedures.

The report describes procedures that were used to create data for bins that had no available data. Appendix C of the report contains a discussion of three proposed methods, of which two were selected. This "proof-of-concept" appendix shows the potential errors in the processes ultimately used to provide approximate data for the model in cases where actual data were not available. The proof-of-concept appendices give appropriate comparisons between measured and modeled results, indicating the level of accuracy to be expected from the model. The agreement between modeled and actual results is usually good, showing that the methodology used here is sound.

Appropriateness of data sets selected

The data sets used in this report were appropriate. This reviewer is not aware of other data sets, which were not examined by EPA, for consideration in this study. EPA used appropriate quality control measures for the selection of data sets to be used and for the selection of individual data points from those data sets.

Data analyses conducted, including the statistical approaches used, the models selected, and appropriateness of the resulting conclusions

The main data analyses conducted here are for the methods for “hole filling” obtaining approximations for missing data bins in a model which requires data for every bin. The basic approach of MOVES, using a combination of source bins and operating condition bins, is based on other analyses published in previous reports and is not a point of this review. The discussion of these analyses is covered in response to specific questions, below.

The report also contains an analysis of fuel rates due to starts. This analysis is similar to the one done in the development of start emission data for MOBILE6.^d Start emissions for light-duty vehicles are defined as the difference between the initial 505 second cold-start portion of the federal test procedure (FTP) for light-duty vehicles and a similar cycle run with no start, called the HR505. An initial regression analysis shows that the HR505 results, which are not a conventional start of the FTP, are virtually the same as the hot-start results, HS505, which are available for all FTP rests. The regression equation for these two different types of tests is $HR505 = 1.0095 HS505 - 0.002$. Although this strongly suggests the near equality of the two results, it would be helpful to add statistical tests for a zero intercept and a slope of 1. Although this analysis was done for the light-duty FTP, there was no mention of a similar analysis for the heavy-duty FTP. Presumably the same assumption – that start emissions were the difference between the cold-start and hot-start portions of the heavy-duty FTP – was made for those vehicles.

***EPA Response:** The statistical test suggested was performed and have been added to the report.*

The lack of data for some source classes requires some significant approximations, as described in the report. In addition, MOVES2004 does not account for different soak times – the time between engine shut-off and restart. Presumably this will be added in subsequent versions of MOVES.

***EPA Response:** soak time effects will be added to subsequent versions of MOVES*

To a first-order approximation, the start energy is the energy required to increase the average temperature of the engine and drive train from its initial value to its value at fully warmed operation. Thus, the result of the analysis – which is to be added in the final draft – that the start energy for hybrids depends on the engine displacement seems a reasonable one. However, the assumption that the start energy for a fuel cell and an electric vehicle is the same as that for a gasoline engine is not reasonable. Both of these operate at a much lower temperature than a gasoline engine.

^d Ed Glover and David Brzezinski, “Final Determination of Hot Running Emissions from FTP Bag Emissions,” Report EPA420-R-01-059, USEPA, Assessment and Standards Division, Office of Transportation and Air Quality, November 2001.

EPA Response: *We will need to evaluate this when more data becomes available, particularly on fuel cell vehicles.*

The regression analyses for the effect of temperature on energy consumption for both starts and running conditions are appropriate. However, no uncertainty analysis is presented for the regression results for the temperature analyses. Table 9-1, which shows the regression results for the effect of temperature on diesel start emissions, has coefficients that do not match the results shown in equation 9-6 which is actually used to compute the temperature effect.

EPA Response: *The regression performed in Table 9-1 isn't meant to produce the coefficients used in the model, simply to assess the relative importance of different variables.*

Specific questions

Are variables important to energy consumption or GHG emission production captured, given available data?

Yes. The combined binning approach (source bins and operating condition bins) is used for energy consumption which is directly related to CO₂ emissions. Here the source bin variables of vehicle type, weight, fuel, and engine displacement will provide the essential variables for determining differences from the different source types. The operating bin boundaries should capture the variations due to vehicle operations.

Data for CH₄ and N₂O are limited to bag data so that operating mode bins were not used for these species. (Presumably all bins were filled with the same average rate shown in Appendix F for CH₄ and N₂O.) It is possible to get second-by-second data for these species using the method discussed in Appendix C for disaggregating bag data into bins. This method was rejected in favor of other methods where data were available. However, since only aggregate bag data are available for CH₄ and N₂O, this approach could be used if approximate binned data were required for these species. Such a task should be reserved for future versions of the model and not considered for immediate application.

The emission rates for CH₄ and N₂O in Appendix F are of the order of 1 g/hr. Emissions of CO₂ are of the order of 6000 g/hr. The global warming potentials for N₂O and methane, given in Table 8-1, are, respectively, 320 and 21. This means that the error in CO₂ equivalent emissions from ignoring these species completely is about 5%. The error from assuming that the emission rate in all bins is the same should be much less than this.

EPA Response: *For future versions we are considering restructuring CH₄ and N₂O to be based on ratios to HC and NO_x, respectively; if this were the case then having binned emissions would be more possible, although it isn't clear that having modal binned rates would be necessary for these pollutant, given the relatively small contribution to total GHG emissions as noted by the reviewer.*

As noted above, this report contains information on how the data for greenhouse gas emissions were selected and a general discussion of how the overall approach for modeling not only greenhouse gas emissions, but also criteria pollutants, would be accomplished. The variables that affect energy consumption can be directly related to the emissions of CO₂, the dominant greenhouse gas emitted from mobile source operations. Although the variables that affect greenhouse gas emissions are captured by the model, these emissions are relatively simple to model. The ultimate goal of MOVES is to model not only greenhouse gases, but also criteria pollutants. Some of the aspects of criteria pollutant modeling are discussed in this report. However, the fact that the appropriate behavior of greenhouse gas emission is captured does not mean that the extension to criteria pollutants can be done without some changes in the approaches used here.

***EPA Response:** We plan to use the same general binning approach for criteria pollutants, but the specific variables used to define source bins or for operating mode bins may vary depending on the pollutant. Our goal would be to harmonize bin definitions across pollutants as much as possible for simplicity, but will analyze each pollutant to assess how realistic this is.*

Has the effectiveness of the modal binning structure chosen for characterizing emission rates been adequately demonstrated with regard to greenhouse gas pollutants and HC, CO and NO_x?

Appendix A compares model predictions with observed results for both average data over several trips and for individual trips. These comparisons show good agreement, particularly for the average data. The comparison of average results for both high-speed and low-speed conditions shows that good comparisons for the average results are not due to cancellation of biased results at high and low speeds. Results for individual trips show significant scatter, but this is to be expected from a model that is intended to predict results for a vehicle fleet.

The method for obtaining trips in this analysis is not clear. It is easy to see how trips can be determined from the onboard data sets, but how were trips determined from cycle data?

***EPA Response:** For the light-duty dataset, trips were defined by each test cycle (for the lab data) or each key on / key off (for the on-board data). For the heavy-duty dataset, trips were defined as key on / key off for the trailer data and for buses, we simply defined all operation as a single trip they were generally idling between excursions.*

The use of one set of trips to derive the binned data for the model, and a separate set of trips for comparison with the model results, is a good approach to give confidence in the model results.

The HC emissions data in Appendix A apparently are for exhaust HC only. How will evaporative emissions be handled in MOVES? If an approach to evaporative emissions is currently under development, it would be interesting to add a brief statement to the treatment of such emissions to this appendix.

EPA Response: *Evaporative emissions are beyond the scope of this report, but will be considered for the criteria pollutant version of MOVES.*

Are methods for “hole filling” using PERE and interpolation/copying sound?

The approach to hole filling is discussed in section 4.3.3 which is supported by appendices C and D. Appendix C provides a comparison of three possible methods. Two of these – calculations from the physical emission rate simulator (PERE) and interpolation – were actually used to obtain data for MOVES. Scaling of bag data was rejected as being less accurate. Section 11.1 provides a set of codes that are used to identify the sources of binned information in MOVES: actual data, PERE results, or a variety of interpolation/copying processes. (Copying data from neighboring bins was done in cases where PERE was not run and it was not possible to find appropriate data for interpolation or extrapolation.)

Obviously, one would like to use data instead of approximations; however, the binning approach selected for MOVES would require a large amount of data. Table D-9 shows that there are 11,322 combinations of source bins and operating condition bins for vehicles manufactured prior to 2010. Only 2,634 (23%) of these combinations were filled with data. The remaining bins were determined by PERE (1,275 or 11%) or by interpolation/copying (7,413 or 65%). However, the number of bins that are filled by PERE and interpolation/copying shows the need for a sound practice for filling these empty bins.

EPA Response: *it is worth noting that although the actual number of bins filled with data is relatively low, the proportion of the fleet covered is much higher, as presented in Table 4-6.*

The proof-of-concept discussion in Appendix C shows reasonable agreement between procedures used for hole filling and actual data. It is likely that the data available for bins comes from measurements on vehicles and operating conditions that comprise a large fraction of actual driving. Thus, the hole filling may represent a fraction of overall vehicle and driving conditions that is less than 65%.

As noted above, there are basically two kinds of bins used in MOVES: source bins and operating condition bins. The discussion of hole-filling approaches is focused on missing source categories. Apparently, there is no use of interpolation among operating condition bins. This seems reasonable since the results shown for these bins are highly nonlinear. It would be helpful to include a statement that no interpolation was done across operating condition bins, if this is in fact the case.

Some improvements are obviously possible, but this would be a longer term effort that could be done as MOVES is expanded to include criteria pollutants. The results from Figure D-3 – showing the mean energy consumption rate as a function of vehicle weight – are a good example of interpolated data that could be improved. The interpolated data between points “0330” and “0800” would extrapolate to a negative energy rate for weights below 18,000 pounds. The PERE result that the lower weight point “0160” has a lower energy rate than the higher weight point “0330” seems unrealistic. There is no

information on the number of data points that go into the “0330” data cell. A future hole-filling approach could try to provide some combination of statistical and modeling results to fill holes with data that are more realistic physically. (It is not apparent how such a method might work, but it would be worth exploring for future MOVES hole filling.)

EPA Response: *This example also illustrates uncertainty that can be introduced into the imputation by the combination of data sources. In this case, the value for weightClasses 0160 and 0800 were generated by PERE, whereas that for 0330 was generated from data. We expect the meanBaseRate to follow a positive linear trend with weight. In this case, however, apparent scatter around the trend results in a decreasing local trend between 0160 and 0330, followed by a steep increasing trend between 0330 and 0800.*

Although Figure 14 on page 117 compares PERE rates with extrapolation and scaled bag data, there did not appear to be any attempt to use a combination of PERE with interpolation/extrapolation to fill a single hole. This combination could perhaps yield a better result than either used separately. Such an approach is intended to be used in regions that are relatively rich with data so that only a few holes have to be filled. This should be contrasted with the process in data-poor areas where it appears that PERE was used to fill a few holes and interpolation/extrapolation was then used based on PERE results instead of actual data.

EPA Response: *Indeed a combination of extrapolation/interpolation and PERE would probably provide a smoother transition between the points. There are several reasons why this was not done: 1) It is very labor intensive; 2) PERE is already calibrated to a number of vehicle types and changing the calibration so that it matches results from these test programs, could throw it off on others. A fixed calibration was maintained throughout for consistency. 3) It is possible that the data is more limited than the model in that only a single driving cycle is included in the test program(s) that generated the rates for that particular source bin. In this case, it is very possible that the IM240 driving schedule is the only one defining the rates for several source bins - or at least dominating the mean value. PERE runs a number of cycles so that the higher power bins are filled, which the IM240 cycle may not.*

There is a shift in the details of the “interpolation” between the proof of concept in Appendix C and the actual application in Appendix D. Appendix D uses linear interpolation of two data points for the data actually used in MOVES. The proof of concept in Appendix C develops a linear regression equation from four data points. This gives a small difference between the method used in the proof of concept and the method used to actually get the data.^e

EPA Response: *It is true the methodologies shifted between the initial assessment in Appendix C and actual application in Appendix D. As noted we would expect a small difference in the results.*

^e In numerical analysis, interpolation usually refers to methods in which an n^{th} order polynomial is used to fit $n+1$ data points exactly, as is done in Appendix D.

The development of PERE is the subject of a previous report not under review here. These comments refer simply to its use for the purpose of filling missing data in the bins structure of MOVES.

Is the method for generating advanced technology rates using PERE sound?

Regulatory mobile source emission models face the challenge of providing emission results for vehicles that have not yet been built. These results are required for analyses used to support state implementation plans and transportation conformity analyses. In MOBILE6 and earlier models, the emissions of future vehicles, under actual operation, were scaled based on the future emission standards. This approach assumes that the technology for meeting future emission standards will be similar to the technology used for current vehicles. This assumption seems more tenuous for advanced technology vehicles and provides the justification for the use of a model such as PERE.

The background for using the PERE model for advanced technology vehicles is provided in another report.^f That report provides some comparisons of modeled fuel consumption results with measured results for fuel economy over the federal test procedure (FTP) cycle. These results show good agreement for conventional vehicles and some early advanced technology vehicles including hybrids and one fuel cell vehicle.

The approach for using PERE to model advanced technology vehicles appears sound. Of course, when such vehicles do enter the fleet, their actual data can be used in MOVES to replace the modeled results. An important data element, which is not part of this report, is a good estimate of the fleet distribution of advanced technology vehicles in future years.

Miscellaneous Comments

The comments in this section are classified as technical comments, report clarity comments, and general comments.

Technical comments

The finite-difference equations for acceleration shown on pages 95 and 96, $a_t = v_t - v_{t-1}$, use backward differences. A more accurate approach would be to use central differences, computing $a_t = v_{t+1} - v_{t-1}$. The definition of acceleration used in the report would be a central difference if it referred to the acceleration at the midpoint: $a_{t+1/2} = v_t - v_{t-1}$; this definition implies that the calculation of vehicle-specific power would use the midpoint velocity, $v_{t+1/2} = (a_t + a_{t-1})$ for consistency.

The assumption of the same heating value for gaseous and liquid hydrogen in Table B-14, pp 97-98, will lead to a small error in the fuel consumption. Less energy is available for combustion because of the need to heat and evaporate the liquid hydrogen.

^f Edward Nam, "Advanced Technology Vehicle Modeling in PERE," USEPA report 420-D-04-002, Assessment and Standards Division, Office of Transportation and Air Quality, March 2004.

Figure C-5, page 107, is called a linear fit of engine size bins and the text says that this linear regression is done using Excel. However, the plot shows curves, not linear, fits to the points. This appears to be due to a plotting error in which the Excel plot option of a smoothed curve was selected rather than simply plotting the points and a linear trend line.

The factors relating the relative global warming potential of different greenhouse gases shown in Table 8-1 appear to be those for a 100-year time frame. Is this correct? Can users substitute global warming potentials for a different time frame selected?

EPA Response: We did use 100 year time frame GWPs, in accordance with IPCC guidance. Users can change these values in the MOVESDefault database; the MOVES2004 Software Design Reference Manual has more details on how to accomplish this.

Starting in Table D-1, page 120, the energy rate is stated in units of kJ/SHO. However, the meaning of SHO is never defined. The energy rate data in this table are of the order of 10^5 kJ/SHO, compared to the expected values (of the order of 10 kJ/s) shown in other parts of the report. On page 15 the energy rate units are listed as kJ/hr. This is less intuitive than kJ/s which is the same as kW. The energy input in kJ/s should be the same as the engine power output in kW divided by the engine efficiency.

EPA Response: SHO stands for source hours operating, e.g. KJ per hour. Activity is expressed in terms of hours, so it is easiest to express energy and emission rates per hour as well.

Comment on report clarity

Both simple classifications of bin boundaries and numeric program ID codes for bins are used in the report. For the ease of the reader, data should be presented in conventional terms. In tables 4-12 and 4-13, for example, a conventional description is used for the fuel type and the model year. However, program codes are used for engine size and for weight. When the code structure is first introduced in Table 4-2, there should be a reference to Appendix B in which the individual parts of the bin ID code are specified.

General comment

During the development of MOBILE6, EPA made data sets used in the generation of data for that model available to users who could then check the parameters in the model. Does EPA intend to do so with the data for MOVES? If so, a statement could be added to the report giving readers directions for obtaining the data files.

EPA Response: All data used to generate the energy and emission rates in MOVES will be made available, through request to mobile@epa.gov.

References

- ¹ *Mobile Source Observation Data User Guide and Reference*, U.S. EPA Office of Transportation and Air Quality, EPA Report EPA-420-B-04-004, February 2004 (available online at <http://www.epa.gov/otaq/models/msod/420b04004.pdf>)
- ² Gerber, W. and P. Henson, *Mobile Source Observation Data (MSOD) Database Update*, Report prepared for U.S. EPA Office of Transportation and Air Quality by Eastern Research Group, Inc., December 2002 (available online at <http://www.epa.gov/otaq/models/ngm/r02033.pdf>)
- ³ Gerber, W. and P. Henson, *Mobile Source Observation Data (MSOD) Database Update*, Report prepared for U.S. EPA Office of Transportation and Air Quality by Eastern Research Group, Inc., April 2003
- ⁴ Limsakul, B., *Mobile Source Observation Data (MSOD) Database Update*, Report prepared for U.S. EPA Office of Transportation and Air Quality by Eastern Research Group, Inc., August 2003
- ⁵ *User Manual for VALDATA Data Checking Software for MSOD Database*, EPA 2C-S026-NTSA, Britton Information Systems, Inc., November 2002
- ⁶ *User Manual for EFLOADData Loading Software for MSOD Database*, EPA 2C-S026-NTSA, Britton Information Systems, Inc., November 2002
- ⁷ DeFries, T., W. Gerber, and S. Kishan, *Determination of Important Parameters for CO₂ and CH₄ Emission Factor Modeling*, Report prepared for U.S. EPA Office of Transportation and Air Quality by Eastern Research Group, Inc., November 2001 (available online at <http://www.epa.gov/otaq/ngm.htm>)
- ⁸ Koupal, J. and J. Kremer, *Air Conditioning Correction Factors in MOBILE6*, U.S. EPA Office of Transportation and Air Quality, EPA Report No. EPA420-R-01-055, November 2001 (available online at <http://www.epa.gov/otaq/models/mobile6/r01055.pdf>)
- ⁹ Carlson, T. and T. Austin, *Development of Speed Correction Cycles*, Report prepared for U.S. EPA Office of Transportation and Air Quality by Sierra Research, Inc., MOBILE6 Report No. M6.SPD.001, 1997 (available at <http://www.epa.gov/otaq/models/mobile6/m6spd001.pdf>)
- ¹⁰ Warila, J.E., E. Glover, J. Koupal and R. Giannelli. *Running Energy Consumption Rates within the MOVES Modal Framework*. U.S. EPA Office of Transportation and Air Quality, Proceedings of 14th CRC On-road Vehicle Emissions Workshop, San Diego, CA, March 29-31, 2004. Coordinating Research Council, Alpharetta, GA.

-
- ¹¹ Nam, E. and R. Giannelli, *Fuel Consumption Modeling of Conventional and Advanced Technology Vehicles in the Physical Emission Rate Estimator (PERE)*, U.S. EPA Office of Transportation and Air Quality, EPA Report No. EPA420-P-05-001 February 2005
- ¹² Chon, D. and J. Heywood, *Performance Scaling of Spark-Ignition Engines: Correlation and Historical Analysis of Production Engine Data*, SAE 2000-01-0565, 2000
- ¹³ Wang, M., *GREET 1.5 – Transportation Fuel Cycle Model Volume 1: Methodology, Development, Use and Results*, Argonne National Laboratory Center for Transportation Research, ANL/ESD-39, Vol 1. August 1999 (available online at <http://www.transportation.anl.gov/software/GREET/publications.html>)
- ¹⁴ Lim, H. *Study of Exhaust Emissions and Idling Heavy Duty Diesel Trucks and Commercially Available Idle Reducing Devices*, SAE 2003-01-0290, September 2002
- ¹⁵ *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990 – 2001*, U.S. EPA Office of Transportation and Air Quality, EPA Report Number 430-R 03-004, April 2003 (available online at: <http://yosemite.epa.gov/oar/globalwarming.nsf/content/ResourceCenterPublicationsGHGEmissionsUSEmissionsInventory2003.html>)
- ¹⁶ Heywood, J. *Internal Combustion Engine Fundamentals*, McGraw-Hill, Inc., New York, New York, 1988
- ¹⁷ *Direct and Indirect Emissions from Mobile Combustion Sources*, EPA Climate Leaders Greenhouse Gas Inventory Protocol, December 2003. (available online at <http://www.epa.gov/climateleaders/pdf/mobilesourceguidance.pdf>)
- ¹⁸ Browning, L. *Update of Methane and Nitrous Oxide Emission Factors for On-Highway Vehicles*, Report prepared for U.S. EPA Office of Transportation and Air Quality by ICF Consulting, EPA Report No. 420-P-04-016, February 2004 (available online at <http://www.epa.gov/otaq/models/ngm.htm>)
- ¹⁹ See reference note 15
- ²⁰ See reference note 15 (Annex N)
- ²¹ Koupal, J., *Air Conditioning Activity Effects in MOBILE6*, U.S. EPA Office of Transportation and Air Quality, EPA Report No. EPA420-R-01-054, November 2001 (available online at <http://www.epa.gov/otaq/models/mobile6/r01054.pdf>)
- ²² See reference note 8
- ²³ Hart, C., J. Koupal and R. Giannelli, *EPA's Onboard Emissions Analysis Shootout:*

Overview and Results, U.S. EPA Office of Transportation and Air Quality, EPA Report No. EPA420-R-02-026, October 2002 (available online at <http://www.epa.gov/otaq/models/ngm/r02026.pdf>).

- ²⁴ Frey, C., Unal, A., Chen, J., Li, S., Xuan, C., *Methodology for Developing Modal Emission Rates for EPA's Multi-Scale Motor Vehicle & Equipment Emission System*, Report prepared for U.S. EPA Office of Transportation and Air Quality by North Carolina State University Computational Laboratory for Energy, Air and Risk, August 2002 (available online at <http://www.epa.gov/otaq/models/ngm/r02027.pdf>)
- ²⁵ Koupal, J. *Emission Analysis Approach for EPA's Multi-scale Motor Vehicle & Equipment Emission System (MOVES)*, U.S. EPA Office of Transportation and Air Quality, Proceedings of 13th CRC On-road Vehicle Emissions Workshop, San Diego, CA, April 7-9, 2003. Coordinating Research Council, Alpharetta, GA.
- ²⁶ Nam, E. *Speed Anomalies in VSP Based Modeling*, Ford Motor Company, Proceedings of 13th CRC On-road Vehicle Emissions Workshop, San Diego, CA, April 7-9, 2003. Coordinating Research Council, Alpharetta, GA.
- ²⁷ Petrushov, V.A. *Coast Down Method in Time-Distance Variables*, SAE Paper No. 970408, 1997
- ²⁸ See reference note 1
- ²⁹ Joy, R. and J. Lee, *I/M Lookup Table Update*. Report prepared for U.S. EPA Office of Transportation and Air Quality by Sierra Research, Inc, October 2000 (available online at <http://www.epa.gov/otaq/regs/im/finalrpt.pdf>)
- ³⁰ See reference note 27
- ³¹ Barth, M., G. Scora and T. Younglove. *A Modal Emission Model for Heavy Duty Diesel Vehicles*. Transportation Research Board Annual Meeting, January, 2004. Washington, D.C.
- ³² Gillespie, T.D. *Fundamentals of Vehicle Dynamics*. Society of Automotive Engineers. Warrendale, Pennsylvania, 1992
- ³³ Kish, L. *Survey Sampling*. John Wiley & Sons, New York, New York, 1965
- ³⁴ *IM240 and Evap Technical Guidance*, U.S. EPA Office of Transportation and Air Quality, EPA Report No. 420-R-00-007, April 2000 (available online at <http://www.epa.gov/otaq/regs/im/r00007.pdf>)