![ERG — Eastern Research Group, Inc.]

Final Report for:

Modifying Link-Level Emissions Modeling Procedures for Applications within the MOVES Framework

Under Federal Highway Administration Contract Number:

DTFH61-09-C-00028

Submitted to:

Federal Highway Administration U.S. Department of Transportation

Submitted By:

Eastern Research Group, Inc.

Publication Number:

FHWA-HEP-11-006

September 2010

![ERG logo](www.erg.com)

Final Report for:

Modifying Link-Level Emissions Modeling Procedures for Applications within the MOVES Framework

Submitted By:

Scott Fincher
Cindy Palacios
Sandeep Kishan
Diane Preusse
Heather Perez

Eastern Research Group, Inc.

Points of Contact:

Principal Investigator	Sandeep Kishan, P.E.	Ofc: 512-407-1830 Fax: 512-409-0089	3508 Far West Blvd, Suite 210, Austin, TX 78731	sandeep.kishan@erg.com
Technical Contact	Scott Fincher	Ofc: 512-407-1826 Fax: 512-409-0089	3508 Far West Blvd, Suite 210, Austin, TX 78731	scott.fincher@erg.com
Business Contact	Linda Taylor	Ofc: 781-674-7229 Fax: 781-674-2851	110 Hartwell Ave Lexington, MA 02421	linda.taylor@erg.com

September 2010

Table of Contents

List of Tables

List of Figures

1.0   Introduction

Emission inventories for on-road mobile sources are compiled across the country in support of State Implementation Plans (SIPs) and Conformity Demonstrations as mandated by the Federal Clean Air Act. Each local agency responsible for developing these inventories has created its own approach for combining vehicle activity, area specific fleet information, local regulations, and EPA's MOBILE6 emissions modeling software. All of these approaches will soon have to be redesigned in order to work with EPA's replacement for MOBILE6, entitled the Motor Vehicle Emission Simulator (MOVES) model. The primary goal of this study is to ease the transition to the MOVES model by providing both a road map for developing updated emissions inventories, as well as an analysis of the likely impacts of the transition on the emissions inventories themselves.

The adoption of EPA's new MOVES model represents a fundamental change in the on-road vehicle emission inventory process. MOVES uses real world driving cycle data and the measured emissions from associated drive cycles to model a wide range of possible driving patterns and their resultant emissions. With this approach, emissions due to any particular driving cycle may be estimated. The resulting estimates are based on real world driving and are well-suited for integration with transportation models. Effective utilization of the new MOVES model within the existing inventory development framework may require extensive modification to on-road data collection and processing requirements. However, successful integration of the MOVES model into this process offers the potential for significant improvements in both the accuracy and precision of the resulting emissions estimates. With this in mind, another goal of this study was to evaluate the effects of using alternative drive cycles within the MOVES model upon vehicle emissions.

This study was designed specifically to answer the following questions:

- What is the best approach for creating an emissions inventory with MOVES based on activity data obtained from a travel demand model (TDM)?

- What is the best approach for creating an emissions inventory with MOVES based on activity data obtained from the Highway Performance Monitoring System (HPMS)?

- What will be the likely impacts on an emission inventory from developing and implementing different drive cycles derived from real world testing data within MOVES?

- What will be the likely impacts on an emission inventory after conversion from a MOBILE6 basis to a MOVES basis?

We hope that the findings of this study will help agencies responsible for developing on-road emission inventories to effectively use data processing tools provided by EPA to convert existing MOBILE6-based vehicle characteristic and activity data to be compatible with MOVES, and to develop additional such tools where necessary. We have also set out to help identify those factors responsible for the largest changes in emission estimates when migrating from MOBILE6 to MOVES, using real world data and case studies. In this way, agencies will be informed as to which types of on-road activity data may be most important to develop and refine for use within MOVES.

In addition to the key questions above, we also kept the following questions in mind:

- How can users best transition from MOBILE6 to MOVES?

- How substantial will the emissions change likely be when switching to MOVES as a primary model?

- What information must be collected to successfully run MOVES using both link-level and non-link-level sources of activity data?

- How difficult is it to run MOVES?

- What potential surprises will users encounter when switching to MOVES?

- How can alternate drive cycle data sets best be incorporated for both HPMS-based and TDM-based inventories in different areas?

- How will output emissions differ when modeling TDM data versus HPMS data?

The version of MOVES used in this study was MOVES2010, originally released in December of 2009. Although we have attempted in this report to cover as much ground as possible in comparison of MOBILE6 to MOVES, the reader should understand that the information provided here is not all-encompassing. For example, this study only examines a specific ozone season day, for a single year, in the eight-county Houston non-attainment area, and its results should not be assumed to be applicable to other regions of the country. Similarly, the data used for the calculation of alternate drive cycles input to the model is derived from data collected in Kansas City, and should be considered specific to that area. In addition, the study only examines impacts from CO, NOx, and VOC emissions; other outputs from MOVES (such as PM10, CO2, or total energy consumption) were not modeled.

Section 2 discusses methodologies for preparation of the inputs necessary to perform a MOVES model run, and details the inputs used in the model runs we developed for this project. Section 3 focuses on the successful execution of MOVES, and processing of output from the model. Section 4 discusses the development of drive cycles based upon the Kansas City Emissions Study previously performed for EPA by ERG, and the steps necessary for integration of those drive cycles into MOVES. Section 5 presents summaries of the emissions inventories for CO, NOx, and VOC emissions calculated using MOVES for all modeled scenarios, and compares those outputs both to the TTI MOBILE6-based inventory, and in the case of alternate drive cycles, to each other. Finally, Section 6 summarizes our conclusions from this study, and attempts to answer the study questions set forth here.

2.0   Preparation of MOVES Inputs

In order to answer the first two study questions listed previously, ERG began the process of re-creating, using MOVES, the emissions inventory for the 8-county Houston-Galveston-Brazoria (HGB) ozone nonattainment area[1] prepared by the Texas Transportation Institute (TTI) using MOBILE6 and other associated tools. We selected this inventory as the subject for the analysis based on the level of detail utilized in the current modeling, as well as our long standing relationship with the Texas Commission on Environmental Quality (TCEQ) and the Texas Transportation Institute (TTI), the parties responsible for developing the inventory for the Houston region.

TTI's inventory uses link-level travel demand model outputs obtained from the Houston-Galveston Area Council (H-GAC), which have been processed and combined with MOBILE6 factors for each of the model's 28 vehicle classes by TTI. TTI processing activities include application of vehicle miles traveled (VMT), volume, speed, time of day, and seasonal adjustment factors to the link-level H-GAC activity data to obtain hourly allocations, as well as preparation and execution of episode-specific MOBILE6 runs in order to develop final mass emissions estimates. The emissions inventory files, as well as all of the supporting data and documentation, used to support the 2006 Houston-Galveston-Brazoria ozone nonattainment SIP is posted on TCEQ's public ftp server[2].

The inventory scenario selected for re-creation was an ozone season non-school weekday. Each county in the HGB area was modeled independently. Data provided by TTI included tab delimited summary files containing travel demand model (TDM) based VMT, VMT mix, hours of operation, vehicle speed, and other parameters on an hourly basis for each county, which were used for development of the first set of MOVES model runs prepared for this analysis. HPMS-based VMT was also provided, and was used to develop the second set of MOVES modeling runs. In addition, TTI provided MOBILE6 input files, along with associated MOBILE6 VMT by hour, speed VMT, VMT by facility, registration distribution, and inspection and maintenance (I/M) inputs.

---

[1] *2005/2006 On-Road Mobile Source, Ozone Episode Modeling Emission Inventories for the HGB Eight-Hour Nonattainment Area* (TTI, 2007)
[2] These files are available at the following URL: ftp://ftp.tceq.state.tx.us/pub/OEPAA/TAD/Modeling/Mobile_EI/.

The TDM-based VMT data was processed using SAS[3] programs developed by ERG to obtain VMT estimates usable in MOVES. ERG developed other spreadsheet tools and methodology for conversion of HPMS-based VMT into MOVES format[4]. MOVES spreadsheet tools developed by EPA specifically for conversion of MOBILE6 inputs into MOVES formats (available at http://www.epa.gov/otaq/models/moves/tools.htm) were also used extensively. The methodology used in preparing inputs for execution in MOVES generally involved four steps:

- Pre-processing of VMT Data

- Application of EPA Converter Tools to VMT Data and MOBILE6 files

- Pre-Processing of Other Required Data

- MOVES Model Setup

The general data flow for the inputs required to the MOVES model is presented in Figure 2-1.

Figure 2-1. MOVES Input Data Flow

---
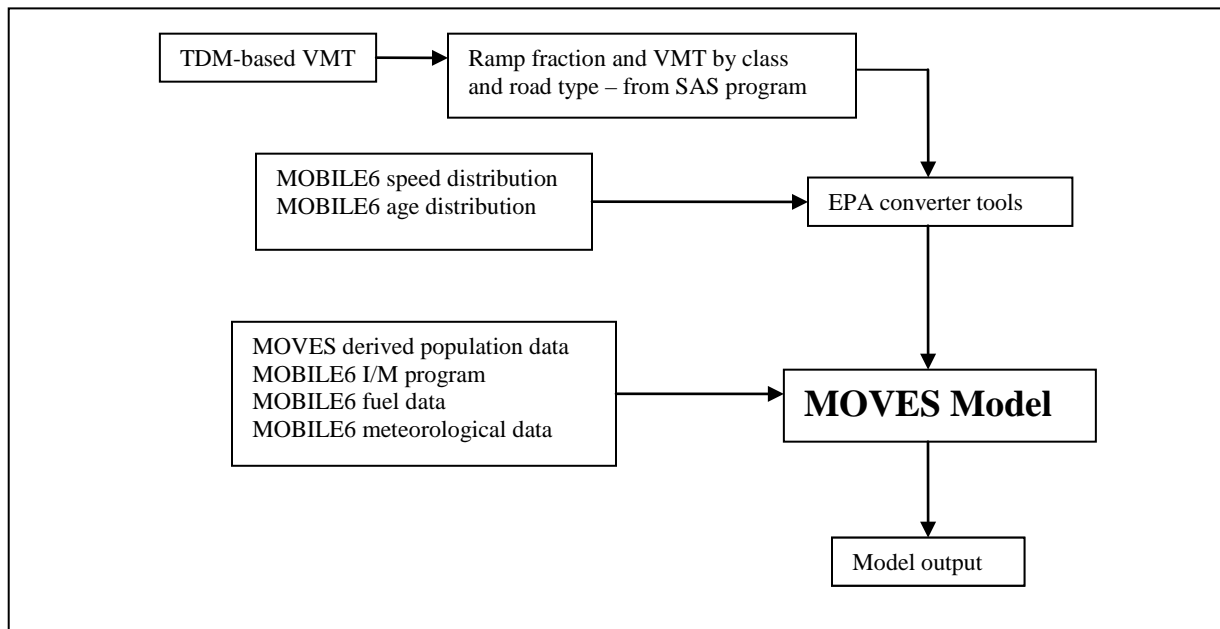
[3] SAS, or the Statistical Analysis System, is an integrated system of software products, developed by SAS Institute Inc., that ERG used in its data analysis during the course of this study. More information is available at www.sas.com.

[4] These tools are provided electronically in Appendix A, and described in more detail in Section 2.1.

A description of the processes used in analysis of the data provided by TTI, along with documentation of the methodology employed in creating MOVES input and output files, is provided below.

2.1   VMT Pre-Processing

The first step in preparing data for input to MOVES is to ensure that the activity data that will be used is representative and complete. Correct application of VMT, whether obtained from a travel demand model, HPMS, or other sources, is crucial to ensure model outputs correctly estimate emissions for a given area. For the purposes of this study, we prepared activity data derived from both TDM and HPMS. In this section, we discuss procedures we undertook in preparing our activity data, prior to subsequent processing in EPA's VMT spreadsheet tools, which will be discussed later.

Processing of TDM-based VMT data

In the case of TDM-based VMT, TTI provided tab-delimited summary files that included VMT, VMT mix, hours of operation, vehicle speed, and other parameters on an hourly basis for each county of interest. An example of the tab-delimited VMT summary data we received is presented in Table 2-1.

Table 2-1. Example Harris County Tab-delimited 24-hour VMT (gasoline vehicles)

ROADWAY TYPE	LDGV	LDGT1	LDGT2	LDGT3	LDGT4	HDGV2b	HDGV3	HDGV4	HDGV5	HDGV6	HDGV7	HDGV8a	HDGV8b
Urban Interstate	5234029	425653	1416995	423979	194975	47905	16793	6649	2387	6478	2131	2472	426
Urban Other Freeway	2325426	189130	629612	188396	86638	21108	7399	2930	1052	2855	939	1089	188
Toll Roads	3554228	289074	962325	287950	132419	32149	11270	4462	1602	4348	1430	1659	286
Ramps (Fwy/Toll/Frnt)	364309	32879	109453	33570	15438	4340	1521	602	216	587	193	224	39
Urban Prin. Art.	1648401	148801	495356	151937	69871	19876	6967	2759	990	2688	884	1026	177
Urban Other Art.	2948941	266194	886158	271797	124991	35163	12326	4880	1752	4755	1564	1814	313
Urban Collector	213299	24812	82600	26893	12367	4218	1479	585	210	570	188	218	38
Local (Cent. Conn.)	3902947	454054	1511540	492125	226314	76860	26942	10668	3829	10394	3419	3966	684
Rural Interstate	12238642	995312	3313381	991404	455917	111089	38941	15419	5534	15023	4942	5733	988
Rural Other Freeway	2521654	205085	682724	204286	93945	22873	8018	3175	1140	3093	1018	1180	203
Rural Prin. Art.	6175292	557389	1855542	569097	261710	73626	25808	10219	3668	9957	3275	3799	655
Rural Other Art.	15704998	1417602	4719178	1447400	665616	186714	65449	25914	9302	25250	8306	9635	1661
Rural Major Col.	801542	93251	310433	101070	46479	15690	5500	2178	782	2122	698	810	140
Rural Collector	473582	55099	183422	59718	27463	9224	3233	1280	460	1247	410	476	82
Local (Intrazonal)	138475	16113	53642	17464	8031	2729	956	379	136	369	121	141	24
TOTALS	58245767	5170448	17212360	5267087	2422176	663563	232601	92097	33059	89736	29518	34241	5902

It became apparent to us that for the purposes of this study, using a combination of the tab-delimited summary files and MOBILE6 inputs provided by TTI would be the most efficient way to create representative MOVES input files. We arrived at this conclusion after perusing the MOVES conversion tools made available by EPA, which rely primarily on existing MOBILE6 inputs for use in creating MOVES input files.[5] It is important to note that EPA's conversion tools are designed to accept activity data in NMIM[6] format, and data available in that format may require little to no additional processing on the part of the user. However, since activity data for a given region may be available in a variety of formats, users interested in preparing TDM-based VMT for input to MOVES for their own region may need to develop their own methods for processing VMT. In doing so, the primary consideration should be:

- to appropriately calculate VMT, in millions of miles, by HPMS road type and MOBILE6 vehicle type. An example of properly formatted VMT data is presented in Figure 2-2.

- to appropriately calculate the percentage of Vehicle Hours Traveled (VHT) apportioned to ramps[7]. MOVES does not have a separate roadway type for ramps. Instead, it requires ramp fraction as an input. If VHT are not available, defaults ramp fractions can be used, although this is not preferred.

- to appropriately calculate VMT distributed by hour, if not already available from existing MOBILE6 inputs.

Each of these inputs are required for EPA's VMT processor, which is further described in Step 2.

---

[5] Because many users of MOVES are likely to already have MOBILE6 input files and (other information based on MOBILE6 vehicle types) available to them, and because EPA is encouraging usage of their conversion tools, we feel this process reflects the application of what will likely be a common way of creating MOVES runspecs.

[6] The National Mobile Inventory Model is, according to EPA, "a free, desktop computer application developed by EPA to help you develop estimates of current and future emission inventories for on-road motor vehicles and nonroad equipment". More information is available at http://www.epa.gov/otaq/nmim.htm.

[7] A procedure for calculating VHT is presented in section 3.6 of EPA's "Technical Guidance on the Use of MOVES2010 for Emission Inventory Preparation in State Implementation Plans and Transportation Conformity," EPA-420-B-09-042.

Figure 2-2. Example Input of VMT to EPA Conversion Tool – Harris County

HARR-vmt-converter-veh28-20100209.xls

Enter Base Year: 2006

BaseYear	vClass28	RoadType	VMT (million miles)		vClass	Vehicle Class	Description
				Note: The vClass field must be populated as follows:			
2006	1	11	12.238642		1	LDGV	Light-Duty Gasoline Vehicles (Passenger Cars)
2006	1	13	8.6969459		2	LDGT1	Light-Duty Gasoline Trucks 1 (0 - 6,000 lbs. GVWR, 0 - 3,750 lbs. LVW)
2006	1	15	15.704997		3	LDGT2	Light-Duty Gasoline Trucks 2 (0 - 6,000 lbs. GVWR, 3,751 - 5,750 lbs. LVW)
2006	1	17	0.8015422		4	LDGT3	Light-Duty Gasoline Trucks 3 (6,001 - 8,500 lbs. GVWR, 0 - 5,750 lbs. ALVW)
2006	1	19	0.4735822		5	LDGT4	Light-Duty Gasoline Trucks 4 (6,001 - 8,500 lbs. GVWR, greater than 5,751 lbs. ALVW)
2006	1	21	0.1384752		6	HDGV2b	Class 2b Heavy-Duty Gasoline Vehicles (8,501 - 10,000 lbs. GVWR)
2006	1	23	5.2340291		7	HDGV3	Class 3 Heavy-Duty Gasoline Vehicles (10,001 - 14,000 lbs. GVWR)
2006	1	25	5.8796547		8	HDGV4	Class 4 Heavy-Duty Gasoline Vehicles (14,001 - 16,000 lbs. GVWR)
2006	1	27	1.6484011		9	HDGV5	Class 5 Heavy-Duty Gasoline Vehicles (16,001 - 19,500 lbs. GVWR)
2006	1	29	2.9489411		10	HDGV6	Class 6 Heavy-Duty Gasoline Vehicles (19,501 - 26,000 lbs. GVWR)
2006	1	31	0.2132992		11	HDGV7	Class 7 Heavy-Duty Gasoline Vehicles (26,001 - 33,000 lbs. GVWR)
2006	1	33	3.9029468		12	HDGV8a	Class 8a Heavy-Duty Gasoline Vehicles (33,001 - 60,000 lbs. GVWR)
2006	2	11	0.9953118		13	HDGV8b	Class 8b Heavy-Duty Gasoline Vehicles (>60,000 lbs. GVWR)
2006	2	13	0.7624738		14	LDDV	Light-Duty Diesel Vehicles (Passenger Cars)
2006	2	15	1.4176015		15	LDDT12	Light-Duty Diesel Trucks 1 and 2 (0 - 6,000 lbs. GVWR)
2006	2	17	0.0932514		16	HDDV2b	Class 2b Heavy-Duty Diesel Vehicles (8,501 - 10,000 lbs. GVWR)
2006	2	19	0.0550985		17	HDDV3	Class 3 Heavy-Duty Diesel Vehicles (10,001 - 14,000 lbs. GVWR)
2006	2	21	0.0161134		18	HDDV4	Class 4 Heavy-Duty Diesel Vehicles (14,001 - 16,000 lbs. GVWR)
2006	2	23	0.4256532		19	HDDV5	Class 5 Heavy-Duty Diesel Vehicles (16,001 - 19,500 lbs. GVWR)
2006	2	25	0.4782044		20	HDDV6	Class 6 Heavy-Duty Diesel Vehicles (19,501 - 26,000 lbs. GVWR)
2006	2	27	0.1488007		21	HDDV7	Class 7 Heavy-Duty Diesel Vehicles (26,001 - 33,000 lbs. GVWR)
2006	2	29	0.2661943		22	HDDV8a	Class 8a Heavy-Duty Diesel Vehicles (33,001 - 60,000 lbs. GVWR)
2006	2	31	0.0248124		23	HDDV8b	Class 8b Heavy-Duty Diesel Vehicles (>60,000 lbs. GVWR)
2006	2	33	0.4540539		24	MC	Motorcycles (Gasoline)
2006	3	11	3.3133812		25	HDGB	Gasoline Buses (School, Transit, and Urban)
2006	3	13	2.538266		26	HDDBT	Diesel Transit and Urban Buses
2006	3	15	4.7191779		27	HDDBS	Diesel School Buses
2006	3	17	0.3104329		28	LDDT34	Light-Duty Diesel Trucks 3 and 4 (6,001 - 8,500 lbs. GVWR)
2006	3	19	0.1834223				
2006	3	21	0.0536415				
2006	3	23	1.4169945				

Sheet tabs: HPMSvTypeYear / MonthVMTFraction / HourVMTFraction / RoadTypeDistribution / RoadType / Instructions / **Import vClass28 VMT** / Import CountyVMTMon

For this study, ERG used a SAS program to process the VMT summary files provided by TTI, aggregated by MOBILE6 vehicle type and H-GAC roadway type. Table 2-2 presents the mapping ERG used for converting H-GAC roadway types to HPMS roadway types; a similar mapping will likely need to be developed for roadway types in other areas. The SAS program was written solely for use with the TDM output provided by TTI, and was intended as both a QC measure (to verify MOBILE6 inputs provided by TTI were consistent with the TDM outputs) and as a tool to generate ramp fractions and VMT for use in the EPA conversion tools. ERG's SAS program produced appropriately formatted VMT, ramp fractions, and other QC outputs[8] for each county, which were subsequently used as input to the EPA converter tools discussed later.

---

[8] ERG's SAS program also independently calculated MOBILE6-format facility VMT, speed VMT, and VMT by hour files. These files were used for QC purposes, and were checked against similar files already provided by TTI, The result was that ERG's facility VMT and VMT by hour files for Harris County very closely resembled those provided by TTI. Speed VMT differed in that ERG's speeds were not as widely distributed across the speed range as TTI's – this was later determined to be caused by ERG's use of the VMT summary, whereas TTI was using link-level VMT as a basis for its speed distribution, which is more accurate. Subsequently, VMT by hour and speed VMT from TTI were used as input to EPA's converter tools for all subsequent MOVES runs.

Table 2-2. H-GAC Road Types Mapped to HPMS Road Types

H-GAC Road Type Code	H-GAC Road Type Description	HPMS Road Type Code	HPMS Area Type	HPMS Road Type Description
1	Urban Interstate Freeways	23	Urban	Interstate
2	Urban Other Freeways	25	Urban	Other Freeways and Expressways
3	Toll Roads	25	Urban	Other Freeways and Expressways
4	Ramps (Frwy/Toll/Frontage)	Special Case, no direct link, must map by ramp fraction		
5	Urban Principal Arterials	27	Urban	Other Principal Arterial
6	Urban Other Arterials	29	Urban	Minor Arterial
7	Urban Collectors	31	Urban	Collector
8	Locals (Centroid Connectors)	33	Urban	Local
10	Rural Interstate Freeways	11	Rural	Interstate
11	Rural Other Freeways	13	Rural	Other Principal Arterial
12	Rural Principal Arterials	13	Rural	Other Principal Arterial
13	Rural Other Arterials	15	Rural	Minor Arterial
14	Rural Major Collectors	17	Rural	Major Collector
15	Rural Collectors	19	Rural	Minor Collector
40	Local (Intrazonals)	21	Rural	Local

Processing of HPMS-based VMT data

In the case of HPMS-based VMT, a different set of procedures must be used to arrive at activity data that can be used as format to the EPA converter tools, and subsequently, to the MOVES model itself. The HPMS VMT, also provided by TTI, is simplified and streamlined relative to the TDM-based VMT discussed above. The TDM based-VMT was provided for a number of specific vehicle types and road types, across all 24 hours of a given ozone season weekday. The HPMS VMT used in this analysis, however, is organized only by road types and area types, as seen in Table 2-3, which are less specific that what is available in the TDM-based VMT, a sample of which is presented previously in Figure 2-2.

Table 2-3. Harris County HPMS VMT Example

County	Harris								
Data Type	Area Type	Interstate	Freeway	Principal Arterial	Minor Arterial	Major Collector	Minor Collector	Local	Total
	Rural (<5,000)	681,075.940	0.000	886,330.630	776,455.580	531,412.352	133,798.075	492,808.352	3,501,880.929
	Small Urban (5,000-49,999)	0.000	0.000	177,295.000	81,002.190	50,515.680	0.000	14,614.040	323,426.910
VMT	Large Urban (50,000-199,999)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Urbanized (200,000+)	30,049,085.890	21,045,993.960	18,072,516.123	19,274,509.064	8,199,444.036	0.000	3,651,528.589	100,293,077.662
	Total	30,730,161.830	21,045,993.960	19,136,141.753	20,131,966.834	8,781,372.068	133,798.075	4,158,950.981	104,118,385.501

Since MOVES requires VMT data formatted by source type, and the HPMS data provided was aggregated only by road type, it is necessary to convert the VMT basis for use in the model. This can be accomplished by extracting default VMT ratios from within MOVES itself , which can then be used to define the relationship between VMT by source type and VMT by road type, and thereby calculate representative VMT by HPMS source type.

In order to do this, ERG exported the HPMSVtypeYear table from the MOVES MySQL database and used the VMTGrowthFactor for the year of interest to project the 1999 HPMSBaseYearVMT to 2006. The general formula is shown here. A sample of the values calculated is shown in Table 2-4.

$$HPMSBaseYearVMT_{2006} = HPMSBaseYearVMT_{1999} * VMTGrowthFactor_{2006}$$

Table 2-4. Default HPMS VMT, Growth Factors, and 2006 Projected VMT

HPMSVtypeID	yearID	VMTGrowthFactor	HPMSBaseYearVMT
10	1999		10600000000
20	1999		1.56864E+12
30	1999		9.00735E+11
40	1999		7657000000
50	1999		70273700000
60	1999		1.32358E+11
10	2006	1.119	11861400000
20	2006	0.98953	1.55222E+12
30	2006	1.0398	9.36584E+11
40	2006	1.007	7710599000
50	2006	1.007	70765615900
60	2006	1.034	1.36858E+11

Next, ERG calculated the VMT fraction by HPMSVtypeID using the HPMSBaseYearVMT. The VMT fraction is equal to the HPMSBaseYearVMT for each HPMSVtypeID, divided by the sum of HPMSBaseYearVMT for all HPMSVtypeIDs, as illustrated by this equation. An example of the calculation is shown in tabular form in Table 2-5.

$$VMTfraction_{HPMSVtypeID} = \frac{HPMSBaseYearVMT_{HPMSVtypeID}}{sum(HPMSBaseYearVMT)}$$

Table 2-5. Default MOVES VMT Fractions

HPMSVtypeID	yearID	HPMSBaseYearVMT	baseYearOffNetVMT	VMT Fraction
10	2006	11861400000	0	0.004367237
20	2006	1.55222E+12	0	0.571508987
30	2006	9.36584E+11	0	0.344840023
40	2006	7710599000	0	0.002838958
50	2006	70765615900	0	0.026055122
60	2006	1.36858E+11	0	0.050389674
TOTAL		2.716E+12		

ERG then applied the VMT fraction by HPMSVtypeID to the HPMS total VMT provided for each county by TTI, as shown in this equation, and in Table 2-6.

$$LOCALVMT_{HPMSVtypeID} = HPMSTotalVMT_{county} * VMTfraction_{HPMSVtypeID}$$

Table 2-6. HPMS VMT by Source Category

County		Brazoria	Chambers	Fort Bend	Galveston	Harris	Liberty	Montgomery	Waller
Total VMT - HPMS		5,811,028.161	2,565,380.290	8,382,649.628	5,941,534.855	104,118,385.501	2,369,353.010	10,079,341.994	1,880,278.930
HPMSVtypeID	**VMT Fraction**								
10	0.004367237	25378.13745	11203.62384	36609.01799	25948.09112	454709.6702	10347.52624	44018.87574	8211.623797
20	0.571508987	3321054.817	1466137.891	4790759.597	3395640.566	59504593.02	1354106.538	5760434.532	1074596.306
30	0.344840023	2003875.083	884645.7974	2890673.088	2048879.014	35904186.42	817047.7457	3475760.522	648395.4289
40	0.002838958	16497.2635	7283.006285	23797.98822	16867.76649	295587.6986	6726.493117	28614.82618	5338.032462
50	0.026055122	151407.0453	66841.29538	218410.9553	154807.4131	2712817.194	61733.78077	262618.4812	48990.89615
60	0.050389674	292815.8144	129268.6764	422398.9816	299392.0041	5246491.498	119390.9257	507894.7567	94746.64222

Finally, ERG formatted the local VMT data by HPMS vehicle type for import into MOVES according to the MOVES template for VMT by vehicle type. (Note that baseYearOffNetVMT is always set to zero for this purpose.) The HPMS-based VMT derived in this way negates the use of the first VMT converter tool described under Step 2 below, in that at this point we have effectively calculated total VMT by county and top level HPMS source category.

When using HPMS-based VMT as a source of activity data for MOVES, users must not only assign appropriate source types to the VMT described above, but also derive road type distributions for the data.[9] In order to perform this calculation for the HPMS data, ERG mapped the Houston-Galveston Area Council (H-GAC) road types listed in the TTI tab files to the standard HPMS road types. In cases where the mapping was not transparent, ERG used its best engineering judgment to assign the H-GAC road types to the most appropriate HPMS road type category. Table 2-7 shows the mapping.

Table 2-7. H-GAC Road Types Mapped to HPMS Road Types

H-GAC Road Type Code	H-GAC Road Type Description	HPMS Road Type Code	HPMS Area Type	HPMS Road Type Description
1	Urban Interstate Freeways	23	Urban	Interstate
2	Urban Other Freeways	25	Urban	Other Freeways and Expressways
3	Toll Roads	25	Urban	Other Freeways and Expressways
4	Ramps (Frwy/Toll/Frontage)	Special Case, no direct link, must map by ramp fraction		
5	Urban Principal Arterials	27	Urban	Other Principal Arterial
6	Urban Other Arterials	29	Urban	Minor Arterial
7	Urban Collectors	31	Urban	Collector
8	Locals (Centroid Connectors)	33	Urban	Local
10	Rural Interstate Freeways	11	Rural	Interstate
11	Rural Other Freeways	13	Rural	Other Principal Arterial
12	Rural Principal Arterials	13	Rural	Other Principal Arterial
13	Rural Other Arterials	15	Rural	Minor Arterial
14	Rural Major Collectors	17	Rural	Major Collector
15	Rural Collectors	19	Rural	Minor Collector
40	Local (Intrazonals)	21	Rural	Local

ERG assigned the HPMS VMT provided by TTI to the appropriate road types. The VMT for the urban HPMS road types (codes 23, 25, 27, 29, 31, and 33) was obtained by summing the

---

[9] For TTI-based VMT, the calculation of road type distribution was handled by the EPA converter tool, as discussed in the following section.

HPMS VMT for small urban areas, large urban areas, and urbanized areas. Then, the HPMS road types were mapped to the MOVES road types. Table 2-8 illustrates the HPMS road type to MOVES road type mapping, which was derived from calculation sheets contained in the EPA converter tools discussed below in Section 2.2.

Table 2-8. HPMS Road Types Mapped to MOVES Road Types

HPMS ROAD TYPE CODE	HPMS AREA TYPE	HPMS ROAD TYPE DESCRIPTION	MOVES ROAD TYPE ID	MOVES ROAD TYPE DESCRIPTION
11	Rural	Interstate	2	Rural restricted access
13	Rural	Other Principal Arterial	3	Rural unrestricted access
15	Rural	Minor Arterial	3	Rural unrestricted access
17	Rural	Major Collector	3	Rural unrestricted access
19	Rural	Minor Collector	3	Rural unrestricted access
21	Rural	Local	3	Rural unrestricted access
23	Urban	Interstate	4	Urban restricted access
25	Urban	Other Freeways and Expressways	4	Urban restricted access
27	Urban	Other Principal Arterial	5	Urban unrestricted access
29	Urban	Minor Arterial	5	Urban unrestricted access
31	Urban	Collector	5	Urban unrestricted access
33	Urban	Local	5	Urban unrestricted access

Once these mapping schemes were completed, ERG summed the VMT by MOVES road type. The next step in the process was to obtain the MOVES default VMT fraction by vehicle type, using the same methodology described above in the HPMS-based VMT Conversion section. After projecting 2006 VMT from 1999 base VMT using factors supplied in the HPMSVtypeYear table in MOVES, ERG calculated a VMT fraction by HPMS type using default VMT ratios, an example of which is shown in Table 2-9.

Table 2-9. Example MOVES Default VMT Fraction by Vehicle Type ID

HPMSVtypeID	VMT Fraction
10	0.004367237
20	0.571508987
30	0.344840023
40	0.002838958
50	0.026055122
60	0.050389674

ERG then multiplied the default VMT fraction (by vehicle type) with the total VMT (by MOVES road type) to obtain the VMT aggregated by HPMS Vehicle type and MOVES road type, according to the following equation.

$$VMT_{HPMSvehicletypeandMOVESroadtype} = VMTFractionbyVehicleType_{MOVESdefault} * VMT_{MOVESRoadType}$$

Following the equation for RoadTypeVMTFraction provided below, ERG then divided the VMT by HPMS vehicle type and MOVES road type with the sum of all VMT for a given HPMS vehicle type in order to obtain the VMT distribution by road type that is required for input to MOVES:

$$RoadTypeVMTFraction = \frac{VMT_{HPMSvehicletypeandMOVESroadtype}}{[Sum(VMT_{HPMSvehicletypeandMOVESroadtype})]_{HPMSvehicletype}}$$

All road type distribution calculations performed in this section are provided for reference electronically in Appendix A in the "Calculate HPMS VMT by HPMS vehicle type" spreadsheet.

2.2   Use of EPA Spreadsheet Converter Tools

Because EPA understands that many users of MOVES will already be familiar with MOBILE6 and NMIM, and likely have on hand previously performed analyses for their areas of interest that use those models, the agency has developed a number of converter tools that allow for transition of older data to MOVES. These spreadsheet tools were most recently updated in February 2010, and are freely available at EPA's website[10]. ERG made extensive use of these tools in adapting TTI's previously developed MOBILE6 files into a form usable in MOVES, and we expect that others users will want to do the same. Each of the converter tools used during this study have been included electronically in Appendix A, and are discussed in additional detail below.

VMT Converter Tool

Having processed the TDM-based VMT using SAS, ERG used the first of several EPA spreadsheet tools employed for this effort, the VMT Converter Tool, to prepare inputs for MOVES. The specific converter used[11] was based on translation of specific VMT by the 28 MOBILE6 vehicle types, listed under "Set 2" on EPA's Tools for MOVES page. This tool is necessary to convert both VMT and road type fractions from a MOBILE6 vehicle type basis to a MOVES source type basis[12.] Note that although we used only one of the VMT converter tools

---

[10] http://www.epa.gov/otaq/models/moves/tools.htm
[11] http://www.epa.gov/otaq/models/moves/tools/vmt-converter-veh28-20100209.xls
[12] It is important to note that MOBILE6 vehicle types are developed primarily on a vehicle weight basis, while MOVES source types focus more on usage classification of a particular vehicle. This distinction arises often while converting information from a MOBILE6 basis to a format useable in MOVES.

EPA provides on the Tools for MOVES website, there are seven other such VMT converters available. These converters are designed to assist users with conversion of VMT on a number of vehicle type bases (either 28, 16, 12, or 8 vehicle types), and based on either specific VMT or VMT/road type fractions. Users of MOVES are encouraged to closely examine EPA's available converter tools to determine which is best suited for their own application.

A separate converter spreadsheet was created for each of the eight counties of interest for this effort. This particular tool requires as input ramp fractions, hourly VMT fractions, monthly VMT fractions, and VMT aggregated by vehicle class and road type, each of which are described below:

- The input ramp fraction calculated using SAS was input for each of the three road types (11, 23, and 25) specified in the Import Ramp Fractions sheet. This was done primarily because the ramp VMT provided by TTI was not associated with a particular road type. Therefore, ramp fractions were assumed to apply equally to each road type. Users with a more specific association between road type and ramp VMT could use more detailed fractions.

- VMT by hour were copied directly from TTI's MOBILE6 inputs for each county into the Import HourlyVMTFractions sheet.

- VMT aggregated by class and road type obtained from SAS processing were copied directly into the Import vClass28 VMT sheet. An example screenshot of this sheet was previously presented in Figure 2-2.

- Note that information was NOT supplied for the Import CountyVMTMonthAllocation sheet, because in re-creating the TTI inventory for the HGB area, modeling was only necessary for a single month (July) and single day. If a user wished to populate this tab for use in MOVES, he could using derive it from an NMIM county database, if available, per the SQL script provided on the Instructions sheet in the converter tool.

The output from the converter tool includes VMT by HPMS source type, monthly VMT fractions by source type, hourly VMT fractions by source type and vehicle type, road type distribution by source type, and a converted ramp fraction. The ramp fraction and road type distribution produced here were directly imported into MOVES, as described later. The monthly fractions produced, meanwhile, were ignored in favor of those produced by EPA's Average Annual Daily (AAD) VMT converter tool, because using the default monthly weighting provided in this tool for VMT in July (1.0871 in a non-leap year) would have over-estimated VMT for the scenario we were modeling. The calculated VMT itself, as shown in Figure 2-3, was in turn used as input to the AADVMT tool.

Figure 2-3. VMT Output from EPA Conversion Tool – Harris County

AADVMT Calculator Tool

The AADVMT tool[13] is the second of EPA's tools ERG used to develop MOVES inputs. This tool is necessary to convert Average Annual Daily VMT, such as that provided to us by TTI, into annual VMT, which is required as input to MOVES whether a user is performing an annual analysis or not. The AADVMT tool calculates annual VMT based on provided AAD VMT, and weights the VMT appropriately across months and days (weekend or weekday) of interest. As before, a separate AADVMT spreadsheet was created for of the eight counties to be modeled.

To use the AADVMT tool, the HMPS VMT obtained from the HPMSvTypeYear output sheet in the VMT converter tool shown above was copied (per guidance in the Instructions sheet) to the Import HPMS AADVMT and Factors sheet in the AADVMT calculator. In addition, all of the monthly and weekend-day adjustment factors on that sheet were changed from their default values to 1.0 to reflect that, for this inventory, we are ultimately only attempting to model emissions for a single day in a single month (thus, our daily VMT did not need to be re-weighted

---

[13] http://www.epa.gov/otaq/models/moves/tools/aadvmtcalculator_hpms.xls

on an monthly or weekend-day basis for conversion to annual VMT). An example of the inputs to the AADVMT tool is shown in Figure 2-4. The outputs from the AADVMT calculator (in the HPMSVTypeYear, monthVMTFraction-calculated, and dayVMTFraction-calculated sheets) were ultimately used for inputs to MOVES through the County Data Manager, which is described in more detail later in this document.

Figure 2-4. AADVMT Input – Harris County

Age Distribution

Yet another converter tool[14] provided by EPA facilitates the transformation of MOBILE6-formatted registration distribution data into MOVES-compatible vehicle age distribution data files. ERG input TTI's MOBILE6 registration data to the tool, which expands registration data across thirty-one years, applies the registration distribution to the vehicle count for the calendar year of interest, and maps the total vehicle counts, by age, to one of the thirteen appropriate MOVES source types. The vehicle counts are then renormalized for each of the thirteen source types. The outputs of the converter tool, formatted appropriately for import into MOVES to populate the SourceTypeAgeDistribution table, include source type, year, vehicle age, and vehicle distribution fraction.

Speed Distribution

EPA also provides a tool[15] that converts MOBILE6-formatted speed distribution files to the format required by MOVES. (We used the MOBILE6-based version of the tool, but EPA provides an NMIM-based version as well for users with NMIM-formatted speed distributions.) ERG input TTI's MOBILE6 speed distribution data to the tool, which expands MOBILE6 fourteen-bin, VMT-based speed distribution files to the MOVES sixteen-bin, time-based speed distribution format. The outputs of the speed distribution converter tool, formatted appropriately for import directly into MOVES to populate the AvgSpeedDistribution table, include source type, road type, hour and day, average speed bin, and speed distribution fraction.

2.3 Preparation of Other Inputs

In addition to the various outputs from the converter tools described in the previous section, other inputs are also necessary for calculating emissions inventories at the county level in MOVES. These include vehicle populations, inspection and maintenance (I/M) program parameters, fuel characteristics, and ambient meteorological conditions. These model inputs were developed by ERG, are provided electronically as part of the MOVES County Data Manager input databases included in Appendix A, and are described in the section that follows.

---

[14] http://www.epa.gov/otaq/models/moves/tools/reg-distrib-converter-veh16-20100209.xls
[15] http://www.epa.gov/otaq/models/moves/tools/averagespeedconverter_mobile6.xls

Vehicle Population

EPA's *Technical Guidance*[16] prescribes the use of state motor vehicle registration data for developing vehicle populations in MOVES, which are used to calculate both start and evaporative emissions. ERG obtained 2006 registration data for each county of interest in this study from the Texas Department of Motor Vehicles. This data was aggregated by MOBILE6 vehicle class. In order to prepare the data for modeling in MOVES, ERG converted the vehicle populations provided to MOVES source types by using the mapping ratios available in Table A.1 of the Appendix in the *Technical Guidance*.

Although usage of registration data to develop vehicle populations is suggested by EPA in the Technical Guidance, users should understand that the vehicles registered in a county for a given point in time do not necessarily correspond to the VMT driven over that same period in the county. For example, commuters living in Fort Bend county may drive a significant portion of their vehicle miles in Harris county. In this example, using registration data as a population surrogate may lead to overestimation of start and evaporative emissions in Fort Bend county, while underestimating those same emissions in Harris County. While a better alternative for determining representative source populations does not currently exist, it is an important issue for users to be aware of.

Ideally, users will be able to develop population data files for import into MOVES using state and local resources, such as state motor vehicle registration data, local transit authorities, and/or data provided by other stakeholders. However, it has been our experience that this data can sometimes be difficult to obtain at the level of disaggregation required for import into MOVES. If adequate resources for development of vehicle populations are lacking, users can follow section 3.3 of EPA's Technical Guidance to calculate local vehicle population based on their VMT data.

Because we originally had difficulty obtaining registration data ourselves, ERG went through the process of developing populations surrogates based on VMT. As described in the EPA Technical Guidance, the first step in deriving local population estimates based on local VMT data is to perform a MOVES modeling run using MOVES default population and VMT data. To do this, ERG selected "National" as the modeling domain, and "County" as the geographic selection type. We then selected the county of interest, along with the 2006 year, for the evaluation. For the cases being evaluated in this exercise, the selections for time span were a

---

[16] "Technical Guidance on the Use of MOVES2010 for Emission Inventory Preparation in State Implementation Plans and Transportation Conformity," EPA-420-B-10-023. United States Environmental Protection Agency. April 2010. Available at http://www.epa.gov/otaq/models/moves/420b10023.pdf

24-hour period for a weekday in the month of July. All vehicle and fuel types (except the placeholder fuel type) were selected, as well as all road types for this modeling scenario. Since the only interest here is obtaining the default population to VMT ratio, selection of a particular pollutant is not relevant. However, the model required the selection of at least one pollutant, so Oxides of Nitrogen was selected for these runs. Finally, output data selections were made, which are important to ensure the required information for calculating the default MOVES population to VMT ratios is present. For the evaluations undertaken here, the following selections were made:

- 24-Hour Day,

- Output VMT data,

- Output Population,

- Output Time Units as Days, and

- Output VMT units as Miles.

Once the model runs were complete, ERG exported the ActivityType table and the MOVESActivityOutput table from the output database generated by the model. Using this data, ERG calculated the MOVES default population to VMT ratio for each source type by dividing the MOVES default population by the MOVES default VMT in the outputs.

Next, ERG obtained the local VMT data, by source type, for the area of interest (which was obtained by summing the VMT by county as calculated in EPA's VMT Converter Tool spreadsheet). Finally, the local vehicle population, by source type, was calculated by applying the default MOVES population to default MOVES VMT ratio to the local VMT, by source type, by county. This calculated population is what can ultimately be used for input to MOVES if other sources of population data are unavailable.

I/M Programs

Information on I/M programs in the HGB area was also provided by TTI in the form of inputs to MOBILE6. Note that only five of the eight counties modeled are affected by I/M – Brazoria, Fort Bend, Galveston, Harris, and Montgomery. Although there are no converter tools provided by EPA to adapt MOBILE6 I/M information into a format usable by MOVES, doing so manually is a fairly straightforward process. However, there are a few issues to be aware of when doing so.

As specified in the MOVES User's Guide[17], the most straightforward approach to making changes to an I/M program in MOVES is to begin with the MOVES default I/M information for the particular county being modeled, and adapt the information in the defaults as necessary. This can be most easily done by exporting the default I/M values using the feature provided in the County Data Manager into an Excel Spreadsheet.

In order to most accurately represent the inventory modeled by TTI, ERG changed the test standards from default values to match those specified in MOBILE6 inputs. In doing so, we added new records to reflect the I/M parameters specified in MOBILE6 inputs and included a value of Y for the useIMyn field as suggested in the MOVES User's Guide. We changed the useIMyn field to N for existing records in the IM table. Specifically, we changed the testStandardsID from 24 to 23 for the exhaust I/M program affecting light duty vehicles, represented by source types 21, 31, and 32 in MOVES (passenger cars, passenger trucks, and light commercial trucks, respectively). This reflected the ASM 2525 Final Cutpoints being used in the MOBILE6 inputs, as opposed to the ASM 2525/5015 Phase-in Cutpoints in place in the MOVES defaults. Similarly, the testStandardsID was changed from 43 to 45 for an evaporative program affecting light-duty vehicles, which represented an OBD Evap and Gas Cap program, as opposed to only an Evap program. In addition, ERG also added an I/M program consisting of a two-mode, 2500 RPM/Idle Test for heavy duty source types 41, 42, 43, 51, 53, 54, 61, 62, in addition to the program that already existed for heavy-duty source type 52.

One of the challenges of applying I/M programs as specified in MOBILE6 is that MOBILE6 vehicle types do not map precisely with MOVES source types. In particular, MOVES source types 31 and 32 can encompass MOBILE6 source types LDGT1 all the way up to HDGV5. In section 3.10.6 of the Technical Guidance, EPA recommends calculating a compliance factor based not only on compliance rate and waiver rate (both provided in MOBILE6 inputs), but also on regulatory class coverage. Calculating the regulatory class coverage is now fairly straightforward, as the information in Table A.3 of the April 2010 *Technical Guidance* Appendix (Gasoline I/M Regulatory Coverage Adjustments) provides the proper conversion factors, along with example calculations.

Once users have appropriately modified the default I/M program for the modeling area of interest and appropriately calculated conversion factors for each I/M program, pollutant, source type, and model year combination, the spreadsheet created can be imported back into MOVES as described later.

---

[17] Available both at http://www.epa.gov/otaq/models/moves/420b09041.pdf and through the MOVES graphical interface

Fuel Formulation and Supply

MOVES uses two tables to characterize information with respect to fuels. The first, Fuel Supply, lists fuel formulations, along with their respective market shares, on a monthly and yearly basis. The second table, Fuel Formulation, lists a number of descriptive parameters for the various fuel formulations. In order to populate these two tables, ERG initially culled information on fuels used in the preparation of TTI's inventory from their MOBILE6 input files, as well as the accompanying 2005/2006 On-Road Mobile Source, Ozone Episode Modeling Emissions Inventories for the HGB Eight-Hour Nonattainment Area report. Reid vapor pressures, oxygenate weight percentages, and sulfur content for both gasoline and diesel fuels were specified in the MOBILE6 runs.

MOBILE6 oxygenates are listed in terms of weight percent, and must be converted to volume percent for use in MOVES. This was done using ratios specified on page 166 of the User's Guide to MOBILE6.1 and MOBILE6.2[18.] Sulfur content and RVP provided in MOBILE6 were directly input to the fuel formulation sheet. The other fields required in the fuel formulation sheet (e.g., aromatic content, e200, cetane index, and so forth) were populated using the existing fuel formulation defaults in MOVES. The fuel supply sheet was populated with two new fuel formulation IDs for July 2006, one representing gasoline and one diesel, each with 100% market share. We felt this most accurately reflected the data available in the TTI MOBILE6 inputs. Note that per the Technical Guidance, creating some sort of single average fuel to be representative of a given area is discouraged – rather, multiple fuels in use for a given area should be input, and market share adjusted appropriately for each of those fuels. In this case, however, ERG created a single fuel, since that most closely mirrored the inputs used in TTI's MOBILE6 input files.

This initial attempt to create a representative fuel supply and formulation uncovered a couple of notable bugs in MOVES2010. The first bug involved creation of new fuel formulation IDs for input to MOVES. When a user creates and imports new fuel formulation IDs, MOVES does not currently "zero out" the market share of existing fuel formulation IDs, which leads to overestimation (approximately two to four times) of pollutants in model calculations. To avoid this problem, EPA suggests that users should currently adjust data associated only with existing fuel formulation IDs, and not create new IDs.

The second bug that was discovered is related to fuel type IDs. ERG's initial import contained a fuel type ID of 12, which represents E10 fuel. The ethanol volume percentage associated with this fuel, as calculated from the MOBILE6 inputs, was 9.28%. This is less than

---

[18] EPA420-R-030-010, August 2003, available at http://www.epa.gov/otaq/models/mobile6/420r03010.pdf

the range of 10-20% that MOVES expects for that particular fuel type ID. Therefore, MOVES did not handle calculation of VOC emissions correctly, and underestimated them by approximately an order of magnitude[19.] In this particular case, the workaround suggested by EPA was to use a different fuel type ID of 13, which corresponds to E8 fuel and has an expected ethanol range that includes the 9.28% specified.

Because of the issues surrounding emissions calculations related to fuel formulations in MOVES, EPA[20] recommended that ERG use the default Fuel Formulation and Fuel Supply provided in MOVES for the HGB area instead of importing or adjusting custom formulations, until the existing bugs are fixed. The most straightforward way to do this is to export MOVES default fuelformulation and fuelsupply tables via the interface, and then reimport them, as suggested in the MOVES User's Guide. In most cases, modeling with the default Fuel Formulation and Fuel Supply for a given area should provide representative outputs.

Meteorological Data

MOVES provides a template for importing hourly temperature and relative humidity data. ERG simply formatted the hourly temperature and relative humidity data from the MOBILE6 input files according to the template in an Excel spreadsheet. Although this is a fairly straightforward conversion that ERG performed manually, EPA also provides two meteorological data converter tools on their website (one for MOBILE6-formatted data, and one for NMIM-formatted data) for modelers to use.

2.4   MOVES Model Setup

The procedures used for setting up and configuring MOVES input files (also known as run specifications, or "runspecs") and input databases were fairly similar across all modeling scenarios performed, differing only in external inputs provided to the model. A separate MOVES input runspec and associated database was created for each county of interest, using default MOVES drive cycles, for both TDM-based VMT, as well as HPMS-based VMT. The same runs were re-created using ERG's drive cycles developed from the Kansas City Emissions Study (discussed in detail in Section 3 below). A total of 32 MOVES runs were performed.

Model Option Selection

The first step in setting up these runs was to make appropriate selections for each model option on the submenus listed in the main MOVES interface (see Figure 2-5). It is very

---

[19] MOVES tables of interest related to this bug include etohbin and fuelsubtypeID.
[20] Per multiple conversations with Sean Hillson

important to note that model option selections for all of the submenus pictured below must be made before entering data via the County Data Manager, or the user may experience difficulties.

Figure 2-5. MOVES Interface Example
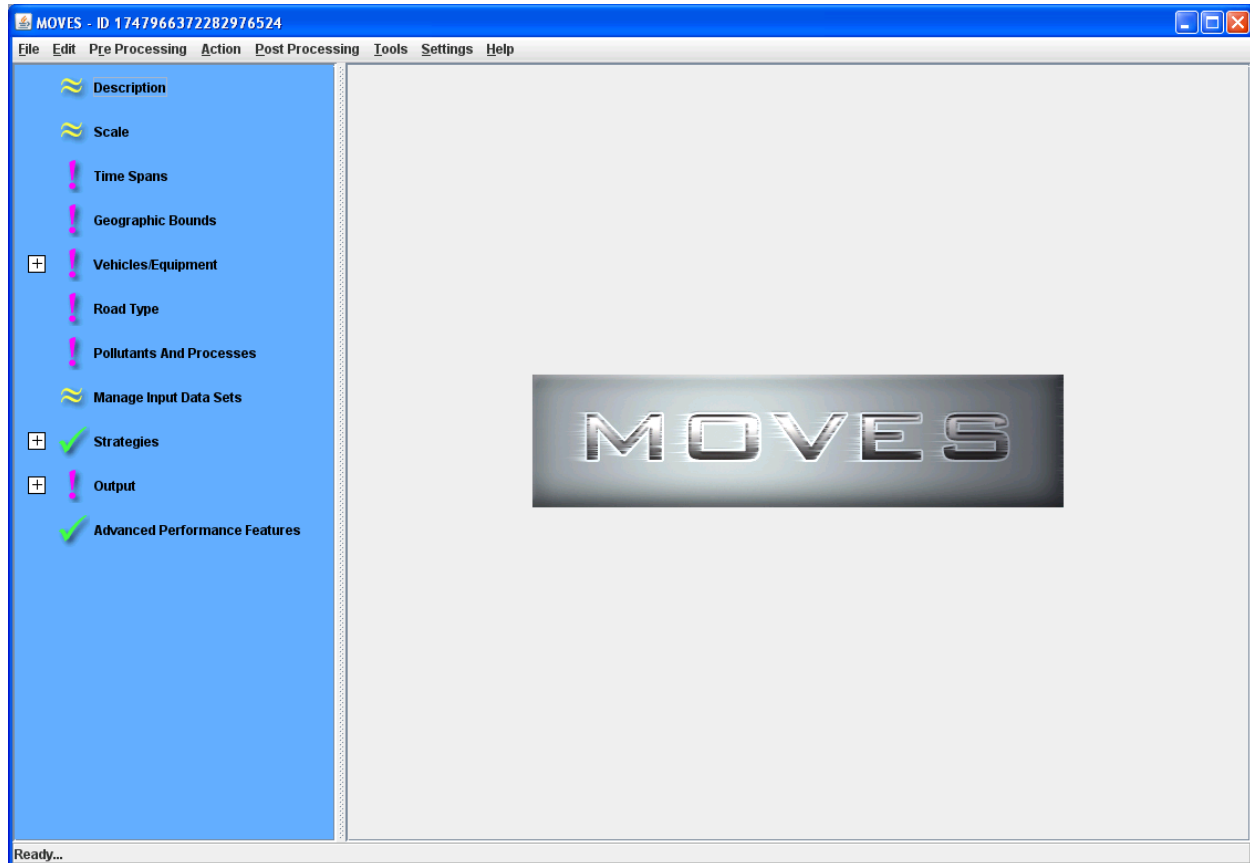
Selections were made for each of the submenus pictured, and they are described below.

- **Description** – A text description was entered in this field. A typical description might consist of text like "USDOT MOVES Runs, Harris County 2006, TDM VMT Basis".

- **Scale** – A County scale, along with an Inventory calculation type, was used for all MOVES performed for this analysis.

- **Time Spans** – Hourly aggregation was selected for a July weekday in 2006. All 24 hours of the day were modeled.

- **Vehicles/Equipment -On Road Vehicles**: Compressed natural gas (CNG), Diesel, Electricity, and Gasoline fuel types were selected, and combined with all available source use types, for a total of 52 fuel/source combinations. Note that this produces warning messages for certain fuel/source combinations that are

currently not present in the MOVES database (for example, diesel motorcycles, gasoline combination long-haul trucks, and gasoline intercity buses, etc.).

- **Road Type** – All five available road types were selected.

- **Pollutants and Processes** - Carbon Monoxide, Oxides of Nitrogen, Volatile Organic Compounds (VOCs), Non-Methane Hydrocarbons, Total Gaseous Hydrocarbons, and Methane were all selected. The latter three pollutants are required, as indicated by the interface during selection, for calculation of VOCs. All process types were selected by checking the box on the far left side of the interface for each pollutant.

- **Manage Input Datasets**: No selection was made in this submenu.

- **Strategies**: No selection was made in this submenu.

- **Output – General Output**: In this submenu, it is important to create an output database with appropriate descriptive nomenclature, particularly when many different such databases will be residing on a single server. For this study, we selected units of tons for mass, million BTU for energy, and miles for distance. In the Activity section, we checked boxes for distance traveled and population, as both are important when performing QC on outputs.

- **Output -Output Emissions Detail**: A time period of 24-hours was selected, along with a location of County, for ease of comparison with the daily emissions calculated on a by-county basis in TTI's inventory. For future detailed analysis, we selected calculation of emissions by emissions process, both by road type and source use type.

- **Advanced Performance Features**: No selection was made in this submenu.

- **Geographic Bounds**: Having fully populated all of the model options in each of the above submenus, ERG then returned to the Geographic Bounds submenu. For each run performed, a single county of interest was selected. We then clicked on the Enter/Edit Data button to begin importing county-specific information into an input database via the County Data Manager.

County Data Manager

The next step in the process is to create an input database, using appropriate descriptive nomenclature, that will store county-specific data for the model. This database must then be populated. We did so using the processed data described in the previous section, for each county and VMT basis modeled. Before a MOVES run can be performed, data must be imported into each of the tabs shown in the County Data Manager, pictured in Figure 2-6.

Figure 2-6. County Data Manager Example
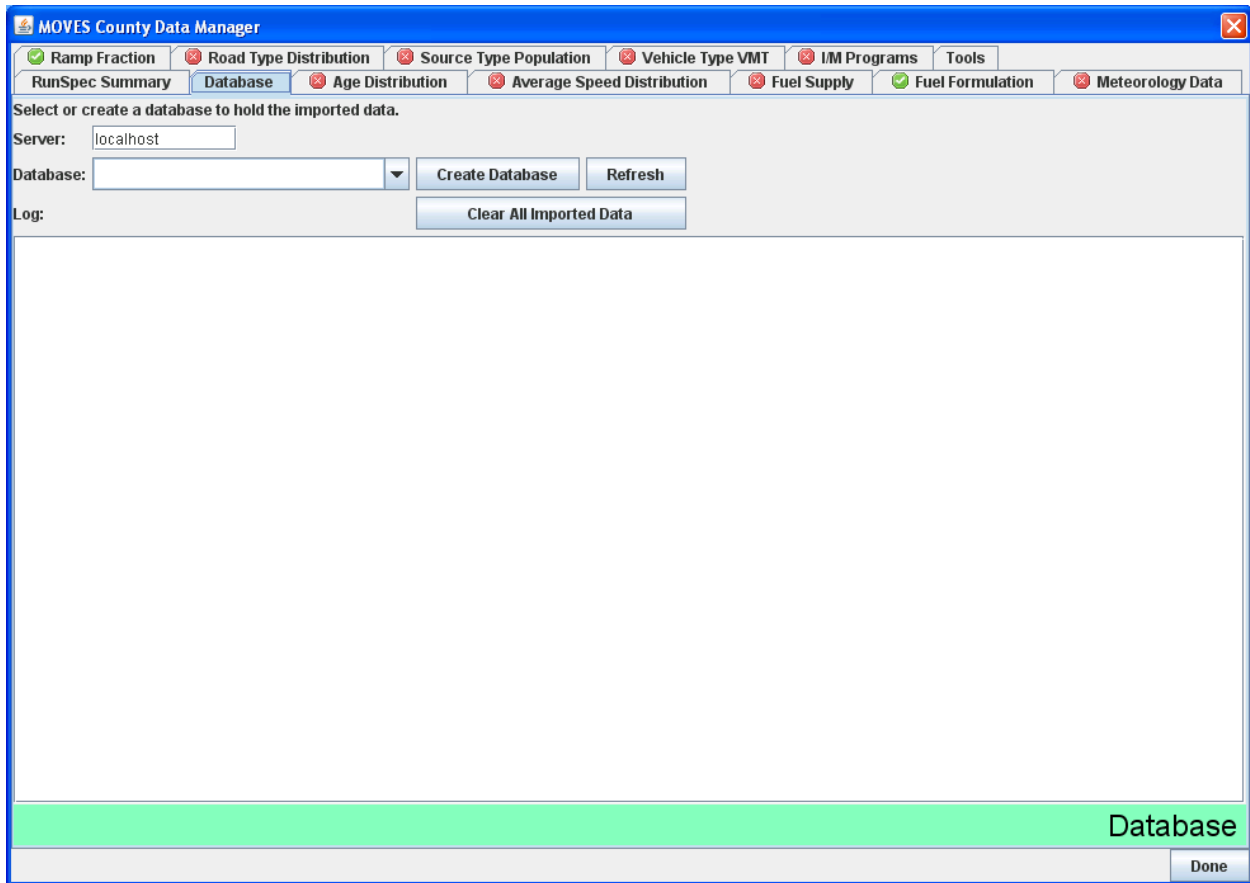
ERG used a number of different spreadsheets to populate the County Data Manager. These spreadsheets are listed in Table 2-10, and have been provided electronically in Appendix A for further reference. The naming convention used here is not particularly meaningful with respect to MOVES; any filename can be used during the import process. Because the County Data Manager requires a number of different inputs to be properly populated, users may find it helpful to use a table such as the one shown here for QC purposes in creating their own MOVES input databases.

Table 2-10. Summary of Inputs to County Data Manager

County Data Manager Input	TTI VMT Basis Data Source	HPMS VMT Basis Data Source
Ramp Fraction	[County]-vmt-converter-veh28-20100209.xls RoadType tab	Defaults Used
Road Type Distribution	TTIRoadTypeDist.xls [County] tab	HPMS_RoadTypeDist.xls MOVES[County] tab
Source Type Population	TXDMV_Converted_Pop.xls [County] tab	TXDMV_Converted_Pop.xls [County] tab
HPMSVTypeYear	[County]aadvmtcalculator.xls HPMSVTypeYear tab	[County]-HPMSaadvmtcalculator.xls HPMSVTypeYear tab
monthVMTFraction	[County]aadvmtcalculator.xls monthVMTFraction-calculated tab	[County]-HPMSaadvmtcalculator.xls monthVMTFraction-calculated tab
dayVMTFraction	[County]aadvmtcalculator.xls dayVMTFraction-calculated tab	[County]-HPMSaadvmtcalculator.xls dayVMTFraction-calculated tab
hourVMTFraction	[County]-vmt-converter-veh28-20100209.xls HourVMTFraction tab	[County]-HPMSaadvmtcalculator.xls hourVMTFraction-default tab
I/M Programs	Revised_IM.xls [County]IM	Revised_IM.xls [County]IM
Age Distribution	RegData_HGBarea.xls [County]MOVES tab	RegData_HGBarea.xls [County]MOVES tab
Average Speed Distribution	DOT_TTI_Speed [County] tab	DOT_TTI_Speed [County] tab
Fuel Formulation	Defaults Used	Defaults Used
Fuel Supply	FuelSupply_HGBArea.xls [County] tab	FuelSupply_HGBArea.xls [County] tab
Meteorology Data	MetTemplate_HGBarea.xls [County] tab	MetTemplate_HGBarea.xls [County] tab

There are a few things to keep in mind when using the County Data Manager. First of all, MOVES sometimes will return errors during an import of data directly from certain sheets in the EPA VMT tool. This can be resolved by copying data from the VMT tool into a blank spreadsheet, and then importing that sheet instead. Secondly, Fuel Formulation data should be imported before Fuel Supply data; if not imported in that order, MOVES will return an error regarding unknown formulations. Finally, note that after importing all four of the VMT sheets required for the Vehicle Type VMT tab, the red X in the interface may not change to a green

check, even when you've imported valid data. This is a known bug and is documented in the current MOVES2010 Errata/Information Sheet[21].

The above procedures describe the methodology for creating a single MOVES run – in this case, for a particular county and VMT basis. However, as previously discussed, 32 different MOVES runs were set up and executed for this analysis, and it would have been fairly tedious to set up all 32 of those runs manually within the MOVES interface. Fortunately, MOVES provides a tool to assist with import of numerous sets of county-specific data. This tool, which generates an XML file to assist with importing, can be accessed on the Tools tab of the County Data Manager by clicking the "Generate XML Importer File" button.

The XML generator was used to create an XML file, which was in turn used to import data into the input database for a given county outside of the MOVES interface, at the command line. When generating MOVES runs for multiple counties, ERG found that it is a fairly straightforward process to run the County Data Manager for a single county, use the XML generator tool to prepare an XML importer template, and then alter that template in a text editor to produce importers for multiple counties. These XML files can then be called from the command line, or in a batch file, to create multiple input databases at once. The procedure for doing so is documented in Appendix C of the MOVES User's Guide. ERG has provided example XML importers generated for this analysis, as well as the batch file created to perform multiple imports at once, in Appendix A of this document.

In addition to creating multiple input databases using the XML generator, it was also necessary for ERG to create the 32 MOVES runspecs described earlier. We found the most efficient way to do this was to start by creating a runspec for a given county using the MOVES graphical interface, and save that file, which is stored by MOVES in XML format. Next, ERG edited the XML runspec manually in a text editor for each county of interest. This is a simple process which involved changing a few references to the county modeled, as well as input and output database paths, within the XML runspec file. These runspec files are also included for reference in Appendix A.

Having created model runspecs and input databases for the 32 model runs, ERG prepared to execute each of the runs in MOVES. Figure 2-7 presents a summary of inputs necessary for preparing county-level MOVES runs, and issues to keep in mind while doing so.

---

[21] EPA-420-B-09-043, available at http://www.epa.gov/otaq/models/moves/420b09043.pdf

Figure 2-7. MOVES Input Preparation Quick Reference

Inputs Needed For County-Level Inventory Calculations in MOVES

- **Activity data** in the form of TDM-based or HPMS based VMT. If possible, obtain data disaggregated by source type and road type. Remember that activity data may require a significant amount of pre-processing, even prior to use of EPA's converter tools.

- **Ramp Fractions**. Defaults may be used, but users are encouraged to calculate these based on available VHT when possible.

- **EPA Converter Tools**. While not strictly required, many users will find these helpful in the process of preparing the following required inputs to the model:

  - **Annual VMT**

  - **Road Type Distributions**

  - **Month, Day and Hour VMT Fractions**

  - **Age Distribution**

  - **Speed Distribution**

- **Source Type (Vehicle) Populations**. These are needed for calculation of emission from vehicle starts and evaporative processes.

- **I/M Program Information**. This will not apply to all areas, but if it is required, program test information, along with applicability to particular source types and model years, will be necessary.

- **Fuel Formulation and Supply**. This includes not only physical characteristics of fuels to be modeled, but also information on fuel market share for a given area.

- **Ambient Meteorological Data**.

Issues to Keep in Mind During Model Input Development

- Review the latest version of the MOVES Errata before beginning calculations with the model

- Does the total VMT output from the VMT converter match the total input across vehicle/source types? If not, this may indicate a problem with VMT processing.

- Are you receiving any warnings from the MOVES interface during import of information to the County Data Manager? If so, check your inputs carefully, as MOVES may still allow you to perform calculations even when there is an error in an input file.

- Using the XML Importer File Tool to generate scripts for multiple counties can be a big time saver if you need to set up multiple MOVES runs. When using XML files to import information into multiple County Data Manager databases, closely check the syntax of input files before execution to avoid errors.

- Double check your model options to ensure you are modeling exactly what you intend to model with respect to time span, source type, road type, pollutants. Ensure your input and output database are named appropriately to avoid confusion.

3.0　Model Execution and Output Processing

      Individual MOVES runs can be executed through the model's graphical use interface (GUI), and multiple runs can be submitted in series through the use of batch files. The execution of MOVES runs can be accomplished using a single computer, or if a user so desires, multiple machines working in tandem. The outputs from such model runs can be analyzed via the GUI, or via external MySQL tools included with the installation of the model. The follow section discusses model execution and output processing in additional detail.

3.1　Model Execution

      After selections have been made for model options in MOVES, and a county input database has been created (through either the County Data Manager or the external XML importer script process described above), MOVES can be executed for a county scale analysis. For an individual run, this is done simply by clicking on the Execute command under the Action dropdown menu in the main MOVES interface.

      In this case, ERG had 32 different runs to execute. We found that using the Multiple Runspec Creator, available under the Tools dropdown menu and pictured in Figure 3-1, was very helpful in generating batch files to allow for execution of multiple model runs, in sequence, outside of the MOVES interface.

Figure 3-1. Multiple RunSpec Creator
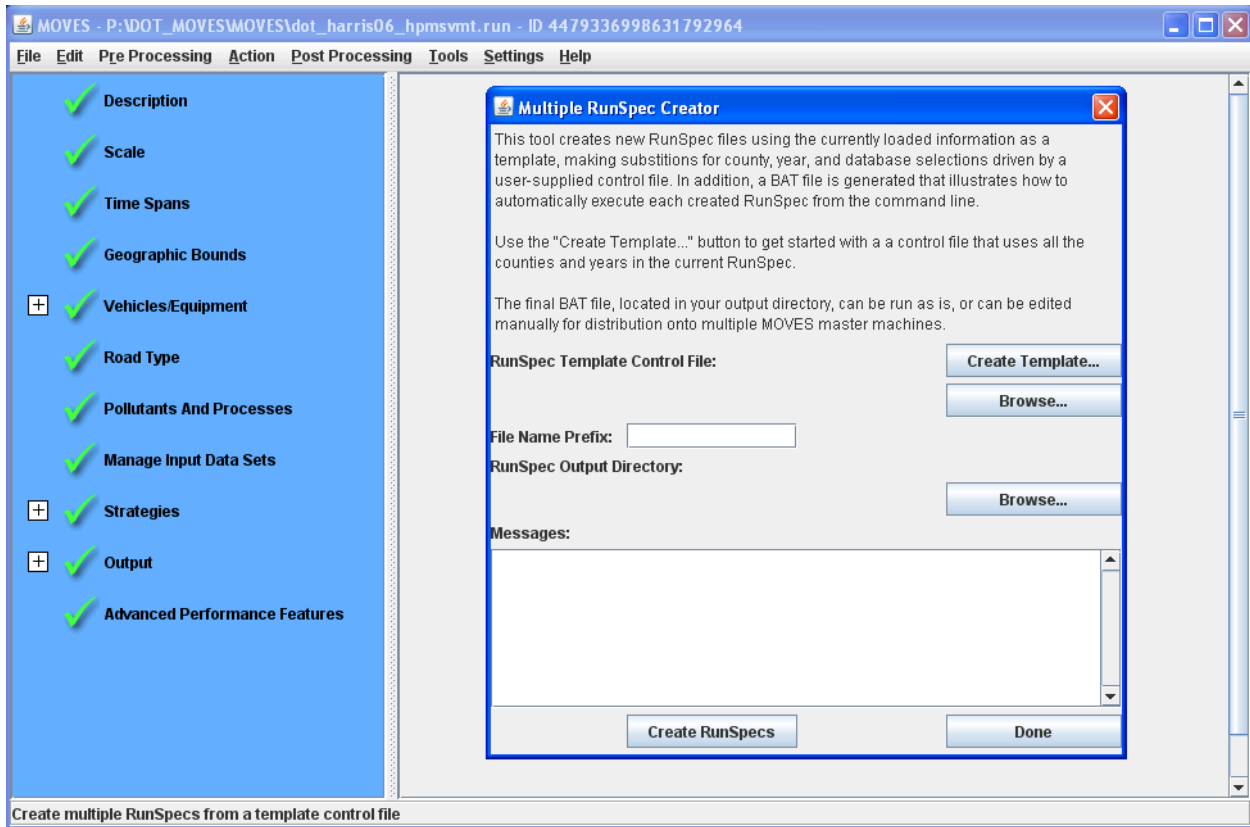
ERG edited the batch file produced by the tool to include all of the different runspecs already prepared, following the guidance presented in section 2.3.6.1 of the MOVES User's Guide. We found it necessary to specify a memory heap size in the batch for execution of the model; without it, MOVES ran out of memory during execution. This may or may not be necessary, depending on the configuration of the computer on which MOVES is executed. A sample batch file used for executing our TDM-based MOVES, which includes the specific memory heap syntax execution used for our runs, is shown in Figure 3-2.

```
@echo off
rem Script generated by the MOVES Multiple RunSpec Creator
rem Based on control file: P:\DOT_MOVES\MOVES\test_template.txt
rem ---------------------------------------------------------
echo Changing to the MOVES folder and compiling code...
C:
cd "C:\Program Files\MOVES20091221"
call setenv.bat
call ant compile
rem ---------------------------------------------------------
echo ******** Running Harris County 2006 TTI VMT Basis ********
java -Xmx300m gov.epa.otaq.moves.master.commandline.MOVESCommandLine -r "P:\DOT_MOVES\MOVES\dot_harris06_ttivmt.run"
echo ******** Running Brazoria County 2006 TTI VMT Basis ********
java -Xmx300m gov.epa.otaq.moves.master.commandline.MOVESCommandLine -r "P:\DOT_MOVES\MOVES\dot_brazoria06_ttivmt.run"
echo ******** Running Chambers County 2006 TTI VMT Basis ********
java -Xmx300m gov.epa.otaq.moves.master.commandline.MOVESCommandLine -r "P:\DOT_MOVES\MOVES\dot_chambers06_ttivmt.run"
echo ******** Running Fort Bend County 2006 TTI VMT Basis ********
java -Xmx300m gov.epa.otaq.moves.master.commandline.MOVESCommandLine -r "P:\DOT_MOVES\MOVES\dot_fortbend06_ttivmt.run"
echo ******** Running Galveston County 2006 TTI VMT Basis ********
java -Xmx300m gov.epa.otaq.moves.master.commandline.MOVESCommandLine -r "P:\DOT_MOVES\MOVES\dot_galveston06_ttivmt.run"
echo ******** Running Liberty County 2006 TTI VMT Basis ********
java -Xmx300m gov.epa.otaq.moves.master.commandline.MOVESCommandLine -r "P:\DOT_MOVES\MOVES\dot_liberty06_ttivmt.run"
echo ******** Running Montgomery County 2006 TTI VMT Basis ********
java -Xmx300m gov.epa.otaq.moves.master.commandline.MOVESCommandLine -r
"P:\DOT_MOVES\MOVES\dot_montgomery06_ttivmt.run"
echo ******** Running Waller County 2006 TTI VMT Basis ********
java -Xmx300m gov.epa.otaq.moves.master.commandline.MOVESCommandLine -r "P:\DOT_MOVES\MOVES\dot_waller06_ttivmt.run"
```

Figure 3-2. Sample Batch File for Execution of Multiple MOVES Runs

We found that modeling each individual county took about 30-35 minutes to execute on our servers, using a single MOVES worker setup. Of course, modeling times will vary based on available computing resources, whether multiple MOVES workers were used (see below), the number and type of pollutants and emissions processes modeled, and other factors. In particular, we have found in other MOVES studies that modeling of evaporative hydrocarbon emissions (as opposed to modeling only exhaust emissions) can increase model execution times by an order of magnitude or more. Model setup time was initially on the order of several hours for many of the County Data Manager inputs described in the previous section, although with repeated model iterations the preparation time for the inputs was significantly decreased.

One way to improve model execution times is to take advantage of MOVES master/worker functionality, which allows users to execute their MOVES runs across multiple computers. When used in this way, the Master computer running MOVES creates calculation "bundles" for other computers to process in a shared network directory. The Worker computers use their local MOVES installation to process the bundle. Multiple Worker computers can process MOVES bundles in parallel, thus reducing the time required for model calculations.

Users wishing to enable this functionality must have multiple machines available with MOVES installed, be willing to make some changes to the Windows operating system on machines running MOVES, and make appropriate edits to some MOVES configuration files. Model execution times can be greatly decreased by implementing a master/worker MOVES run setup, but the users should be aware of both the benefits and drawbacks when using MOVES in this way.

To execute MOVES across multiple computers, it is first necessary to enable File and Printer Sharing For Microsoft Networks within the operating system of the computer to be used. When this feature is enabled, users can then select a directory located on the machine of the Master computer, and enable file sharing for that directory with read/write access. This will allow other Worker computers to see and process the calculation bundles prepared by the Master. Next, configuration files should be edited on both the Master computer (MOVESconfiguration.txt), as well as any Worker computers (Workerconfiguration.txt), to reflect the location of the Shared Work Folder. In both files, the sharedDistributedFolderPath should be set appropriately. For the Master, this will refer to a local folder, while for the Worker, it will refer to a network shared folder (\\MOVESMASTER in the example below). An example of the MOVESconfiguration.txt and Workerconfiguration.txt files, with the sharedDistributedFolderPath highlighted, is presented in Figure 3-3. Having configured MOVES in this way, a model run on the Master computer can either be executed from either the GUI or the command line. Once the Master has been executed, the Worker program can be executed on other machines, which will begin processing "bundles" automatically as they become available.

Figure 3-3. Sample MOVES Configuration Files

Master Configuration

```
defaultServerName = localhost
defaultDatabaseName = MOVESDB20091221
outputServerName = localhost
outputDatabaseName = MOVESOutput
sharedDistributedFolderPath = C:\Program Files\MOVES20091221\SharedWork
GREETWTPApplication = C:\Program Files\MOVES20091221\GREET\GREETGUI.exe
GREETManufactureApplication = C:\Program Files\MOVES20091221\GREET\GREETMfgStub.exe
GREETDirectory = C:\Program Files\MOVES20091221\GREET
computerIDPath =
```

Worker Configuration

```
sharedDistributedFolderPath = \\MOVESMASTER\SharedWork
workFolderPath = C:\Program Files\MOVES20091221\WorkerFolder
workerDatabaseName = MOVESWorker
```

Users should be aware that while a master/worker MOVES setup can be very beneficial in terms of model run time, the system does have some limitations. First of all, only one Worker program can be executed on a single computer at any one time. Per the MOVES User's Guide, "It would be detrimental to performance to operate more than one copy of the MOVES Worker program on a single computer." Secondly, the system does not seems to be fault tolerant – that is, if one of the workers involved in a run fails to process a bundle for some reason (which can occur under a number of conditions: for example, system runs out of disk space, power failure, network disruption) the MOVES run will not complete. In this case, the run will need to be executed again. So long as model runs are closely monitored during execution, however, this drawback can be managed. On the whole, we believe that given appropriate computing resources, most users will benefit from executing MOVES runs across multiple computers using the configuration described here.

3.2   MOVES Output Processing

In reviewing outputs from MOVES, our interest was in comparing modeled CO, NOx, and VOC emissions to those developed by TTI using MOBILE6 in their inventory[22], as well as comparing emissions produced using the default MOVES drive cycles to the cycles developed using the Kansas City data described in Section 4. To produce outputs from the model, the user can either generate summary reports using the MOVES GUI, or evaluate model outputs directly using MySQL. Both approaches are discussed in additional detail below.

MOVES GUI Summary Reports

As mentioned, users can select the Produce Summary Report option, available under the Post Processing menu in the main MOVES interface, to generate emissions summaries for each county of interest. Using this option, users can also select output by total MOVES emissions, or by individual emissions process (for example, crankcase running exhaust, evaporative fuel leaks, refueling spillage loss, etc.) as shown in Figure 3-4.

---

[22] We originally limited our analysis to only CO, NOx, and VOC because those were the pollutants modeled in TTI's original MOBILE6 runs.

Figure 3-4. MOVES Output Emission Process Selection
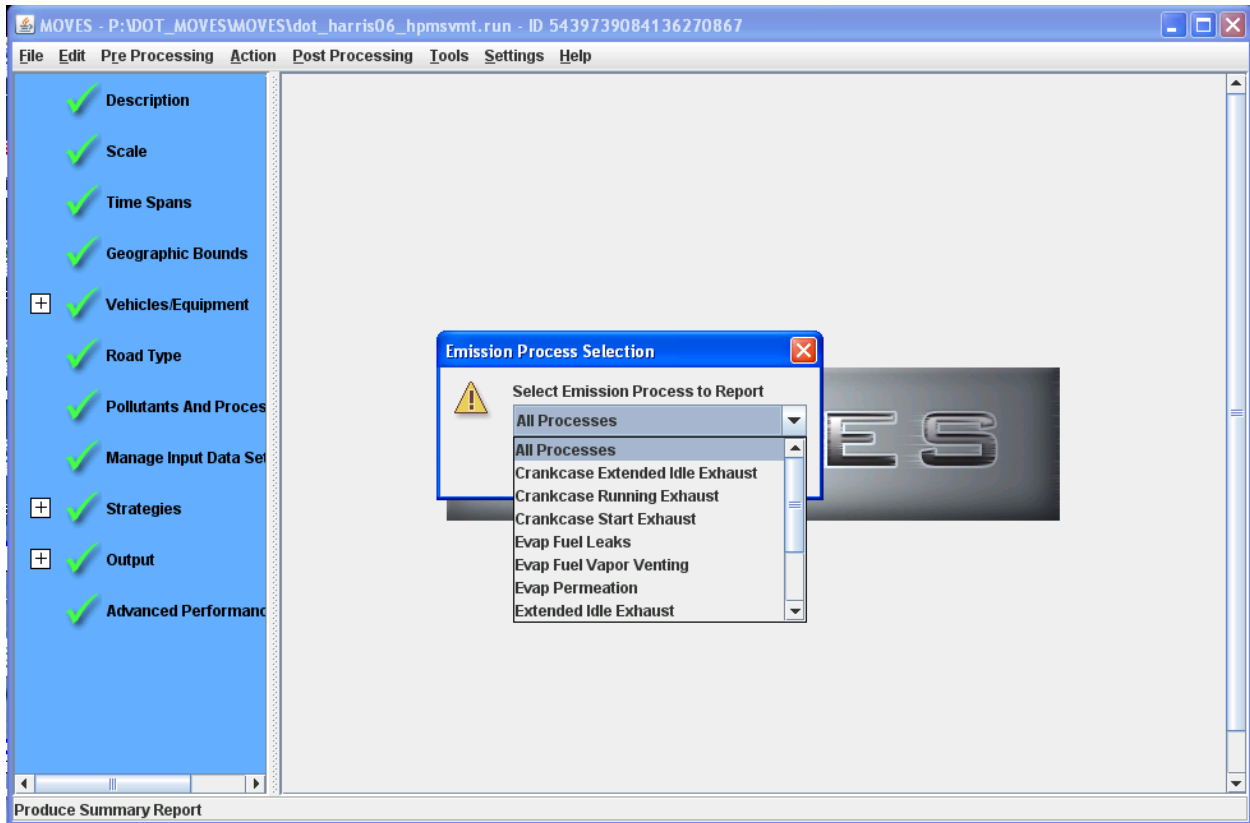
Beyond the modeled emissions processes, the Summary Report input dialog, pictured in Figure 3-5, allows the user to summarize model output with a number of parameters. Of particular interest is the selection of Distance as an output, which serves as an effective check to ensure that VMT input by the users passes through the model appropriately; incorrect pass through of VMT, usually due to setting up model options incorrectly, can lead to very different calculated emissions than what the user might expect. An example of the onscreen summary output that be obtained from MOVES is presented in Figure 3-6. Summaries are easily exported for use in Microsoft Excel or other programs using the Produce Tabbed Output option. Additional information on preparation of Summary Reports is available in Section 2.3.5.2 of the MOVES User's Guide.

Figure 3-5. Summary Report Input Dialog Example
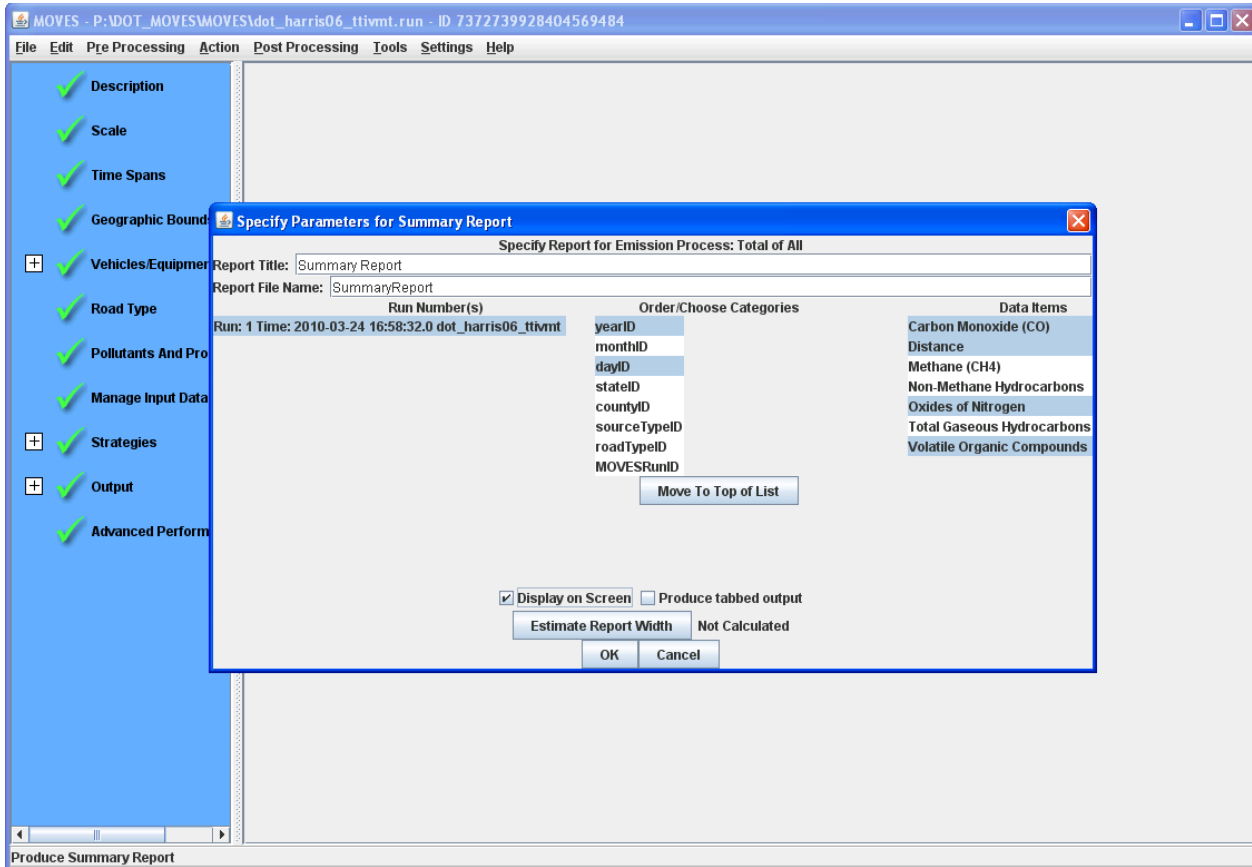
Figure 3-6. Summary Report Output Example

Processing Model Output Using MySQL

Users can also process and analyze output from MOVES outside of the model's interface entirely by using the MySQL Query Browser, which is included with a typical MOVES installation. This requires some familiarity with MySQL query syntax, but has some advantages relative to processing model output via the MOVES GUI.

To create emissions output like the examples below, users can first select the Run MySQL Script on Output Database option (located under the Post Processing menu within the MOVES interface) and choose to execute the DecodeMOVESOutput.sql script23. This script creates additional decoded MOVES emissions and activity tables in the selected output database. Users can then write custom queries and create detailed summaries of their own using the newly decoded tables in the MySQL Query Browser. We have provided two sample queries in Figures 3-7 and 3-8 that recreate the model outputs presented in Figure 3-6. The first script calculates emissions for Harris County for each source type, while the second returns both VMT and vehicle population used in the model calculations. Users are of course free to aggregrate emissions and activity as they choose by using the GROUP BY function on the wide variety of variables available in the output tables, using MySQL scripts that they create.

---

[23] This script can also be executed outside of the MOVES GUI, if desired. In point of fact, running this script is not absolutely necessary at all if a user wishes to perform calculations on the unmodified *movesoutput* and *movesactivityoutput* tables

Figure 3-7. Sample Query and Output using MySQL Query Browser - Emissions

Figure 3-8. Sample Query and Output using MySQL Query Browser - Activity

Although producing query outputs such as those shown above requires some knowledge of MySQL query syntax, there are several advantages to analyzing the outputs of the MOVES model in this way. First of all, the flexibility inherent in MySQL allows for users to investigate output from the model in a more robust way, when compared to the limited options available in the MOVES GUI Summary Reports. Appendix H of the MOVES User's Guide presents a series of script examples that deal with post processing of MOVES outputs in the MySQL environment. Secondly, analysis of model outputs using the GUI cannot be performed when a MOVES model run is in progress; use of MySQL allows users to examine model outputs while concurrently executing other modeling. Finally, depending on the units selected in the model output options, the Summary Reports available in GUI may not provide enough resolution for proper analysis of data, since these reports present data in whole numbers only. The output available by processing MySQL queries, on the other hand, contains a large number of significant figures.

Figure 3-9 presents a list of things to keep in mind while preparing to execute MOVES runs, and process associated model outputs.

Figure 3-9. MOVES Execution and Output Processing Quick Reference

Model Execution

- The MOVES multiple runspec generator can be used to generate batch files to allow for execution of multiple model runs

- Be aware that modeling of evaporative hydrocarbon emissions can increase model execution times significantly, when compared to modeling exhaust emissions only.

- Use of a master/worker MOVES setup, using multiple worker computers, can effectively reduce model run time. However, understand that the failure of any one worker to complete its processing will prohibit the entire run from finishing cleanly.

- Be sure that you don't execute more than one worker program on a given machine, as this is detrimental to model run time.

Model Output Processing

- Processing outputs using the MOVES GUI is quick and effective, but has limited options for aggregation and disaggregation of data.

- Output summaries cannot be produced by the MOVES GUI while another run is in progress; use of the MySQL Query Browser allows for analysis of emissions data from one model run while execution of another run is ongoing.

- Analysis of output in the MOVES GUI can be limited by the selection of units made in model options. The GUI presents output in whole numbers only, while output viewed via MySQL provides many more significant figures.

- If a user is familiar with SQL syntax and query construction, he can investigate outputs from the model in a more robust manner.

- If a user has generated outputs from multiple MOVES runs in a single database, he must be careful to appropriately segregate his output data in MySQL by selecting the appropriate MOVESRunID. Otherwise, he will combine emissions estimates from multiple runs. This same type of error can occur when aggregating across time periods as well (for example, summing emissions over both weekends and weekdays when only output from weekdays is desired). Such an error is much less likely to occur when using the MOVES GUI to summarize outputs.

4.0　Drive Cycle Development

One of the tasks undertaken during this project involved development of drive cycles (also known as drive schedules) based on the Kansas City Emissions Study previously performed by ERG for EPA. MOVES uses real world driving cycle data and the measured emissions from associated drive cycles to model a wide range of possible driving patterns and their resultant emissions. Specifically, MOVES uses the second-by-second speed data included in its drive cycles to calculate the second-by-second vehicle specific power (VSP). It then uses this VSP, along with vehicle type and vehicle age, to weight emission rates before they are applied to the activity data to generate emissions estimates. The resulting estimates are thus based on real world driving.

Drive cycles in MOVES are classified by average speed, vehicle type, and roadway type. In this section, we discuss the methodology we used to develop our own drive cycles based on the collected Kansas City data, and the ways in which we adapted them for inclusion in the MOVES model itself. Having derived alternate drive cycles, we created MOVES modeling runs using those drive cycles for comparison to the modeling runs described in Section 2 above.

In the Kansas City Emissions Study, a large number of light-duty vehicles were equipped with data loggers. The data loggers were in operation while the vehicles were driven on a prescribed "conditioning" run (in which vehicles were driven on a set route in order to prepare them for dynamometer testing at the Kansas City site) and also while the vehicle owners operated the vehicles under regular driving conditions. The data loggers collected and stored second-by-second vehicle operating information, including speed, RPM data, mass air flow, and other data, as well as the time, and the latitude and longitude of the location of each second of driving.

The data stored by the data loggers was used to develop drive cycles that are representative of actual driving in Kansas City. The latitude and longitude coordinates of each second of driving were used to designate each second of driving as having taken place in a rural or an urban area, as well as the type of roadway (freeway, ramp, local road, etc). All of the driving data was then divided into "micro-trips", where each micro-trip is the driving that begins either at vehicle-on or after an idle period, and ends either at vehicle-off or when the vehicle returns to idle. The micro-trips were divided into bins according to road type and average speed, and then for each road type/speed bin, a vector comparison process was used to select a small number of micro-trips that best represented the overall pool of micro-trips for that bin. These selected micro-trips became the driving cycle for each road type/speed combination.

Generally speaking, the development of fleet-specific driving cycles is a resource-intensive process. There are several steps, which include:

1) **Gathering in-use driving data for the fleet of interest**. This typically includes outfitting selected vehicles from the fleet with data-loggers that collect and store, at a minimum, second-by-second vehicle speed. Additionally, storage of information such as engine RPM, mass air flow, and other operational data can be useful in the process of interpreting the driving history for the vehicle. Also, if drive cycles are to be developed for different road types, then the location of the vehicle at each second of driving must be collected and stored. The data that is collected must include enough vehicles, over enough time, to develop a large pool of second-by-second trip information that can be used to represent the general driving characteristics of the fleet. If a small number of vehicles or a short-duration study is used, there is the risk of developing a drive cycle that does not represent the general fleet. This step requires qualified field personnel to install the data loggers (this is critical: improperly installed data loggers result in low quality output data), the data loggers themselves, and a pool of vehicles and drivers. The second-by-second driving should ideally include several million seconds of usable driving.

2) **Interpretation and QA/QC of the driving data**. Once the second-by-second data is collected, it needs to be checked for accuracy and consistency. Among other issues, data loggers can fail intermittently, speeds can get stuck, and random noise sometimes enters the signal. Every data logger and dataset will have its own issues, which need to be discovered and dealt with before the data can be relied upon for the development of drive cycles. This typically involves the manual examination of plots of second-by-second driving data, in addition to analysis of statistics for each trip (for example, mean speed, number of seconds at idle, maximum and minimum acceleration, etc.). This process can be very time consuming, but should be done by an analyst with a good understanding of vehicle operating characteristics, and with skill in extracting pertinent information from a large and variable dataset.

3) **Conversion of the driving data into a pool of micro-trips**. This step prepares the driving data for use in selecting individual micro-trips as components of a drive cycle. It requires software that is capable of matrix manipulation (such as SAS, Matlab, or similar), a system that can store and manipulate datasets on the order of one or more gigabytes, and an analyst with an understanding of matrices as well as vehicle driving characteristics.

4) **Selection of the micro-trips that will comprise the desired driving cycle(s).** This involves comparing each individual micro-trip with the pool of all micro-trips and choosing the micro-trips that best represent the overall pool. Similar to step 3, this step requires software that can work with matrices, a system that can handle the dataset size, and an analyst with an understanding of matrices.

5) **Post-processing of the selected driving cycle(s).** This includes checking statistics for the final cycle, looking at plots of second-by-second speed and acceleration to check for anomalous data, and, depending on the data logger used, possibly smoothing out noise from the stored driving data.

Because of the level of effort required, we expect that many users will elect not to pursue development of vehicle drive cycles specific to their area, and instead choose to use the default drive cycles provided in the MOVES model. Nonetheless, we have documented the methodology used for this study in the hopes that it may be of use to others in the future who wish to undertake similar work. The remainder of the text in this section describes procedures undertaken specifically to convert the collected Kansas City activity data into drive cycles capable of being input into MOVES.

4.1   Preparation of Raw Data

Second-by-second driving data for light duty passenger vehicles was used for this study. The data was collected in Kansas City between July 2004 and April 2005. On-vehicle data loggers recorded the second-by-second driving data, while a GPS unit recorded the latitude and longitude coordinates of the vehicle for each second. A total of 4.6 million seconds of driving data were collected and available for use in this study. The driving data includes a number of repeated trips on a specific vehicle conditioning route, as well as normal, everyday driving by the vehicle owner.

The first step in processing the driving data was to use the latitude/longitude information to determine the location of the vehicle, and therefore the type of road being driven on, for each second of driving. The next step was to identify and attempt to correct any issues with the data that would reduce its quality for use in building drive cycles. These steps are described below.

Assignment of Road Types

A Geographic Information System (GIS) was used to assign road types to the trip data points. Second-by second data points had latitude and longitude coordinates which were mapped in batches with the World Geodetic System 1984 (WGS 84). An initial quality check was done to remove obviously erroneous data, such as (0,0) points, and trips that showed no relation to the area's road network.

Urban/rural designations were obtained by mapping each data point as either within or outside of an urban area from the U.S. Census Bureau's 2000 Urbanized Areas (UAs)

shapefile.[24] The U.S. Census Bureau delineates UAs to provide a better separation of urban and rural territory, population, and housing in the vicinity of large places.

Road type classification was based on ESRI's 2008 U.S. and Canada Detailed Streets layer, which came from the 2003 Tele Atlas Dynamap Transportation version 5.2 product. This detailed road network includes 10 unique road classes indicating various roadway types such as highways, local roads, and so forth. Every road segment in the data set is categorized into one of these 10 road classes (shown in Table 4-1). Spatial analysis tools in the GIS were used to identify the road segment nearest each data point and assign the segment's road class to the point.

Table 4-1. Road classification codes and definitions

Road Class	Road Class Definition
1	Limited Access
2	Highway
3	Major Road
4	Local Road
5	Minor Road
6	Other Road
7	Ramp
8	Ferry
9	Pedestrian Way
0	A special class for high- level routing cross-country and in complicated urban areas.
*00	Off-grid, parking lot, etc.

However, additional revisions were required to address two primary limitations of this automated process. First, data points located close to a cross street were sometimes classified by the software as having the cross street's road class when the trip as a whole clearly continued on the primary road. These points were identified and manually assigned the correct primary road class. Second, this approach ensured that all data points were assigned a road class, even if the trip did not follow road segments. Points that did not closely follow the road network were identified and assigned a separate "00" class indicating points or trip portions that were off-grid such as points off of roads, within parking lots, or with erroneous coordinates. Figure 4-1 presents an example plot of second-by-second observations, along with classifications of nearby roadway types.

---

[24] http://www.census.gov/geo/www/cob/ua_metadata.html

Figure 4-1. Sample Assignment of Road Types and Data Points
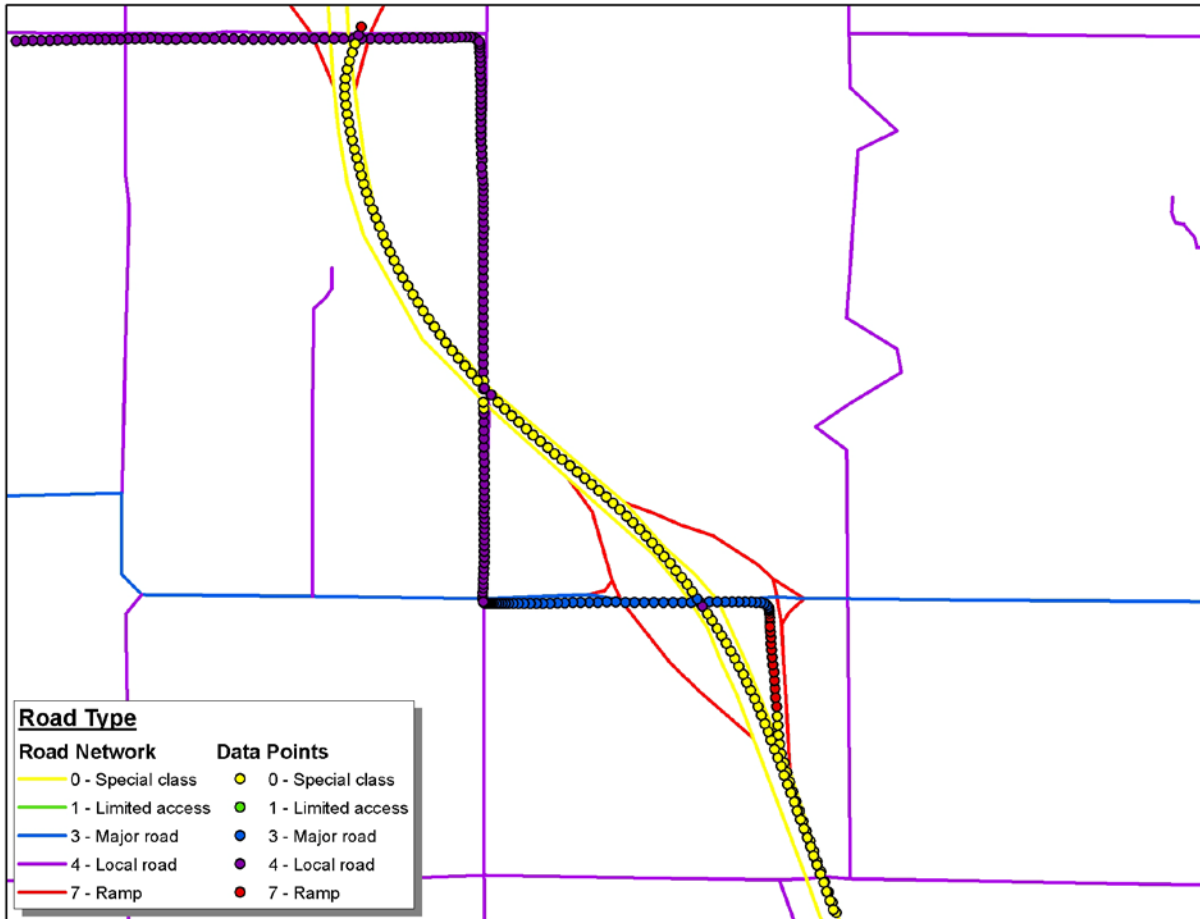
The road class definitions listed in Table 4-1 were next condensed to correlate to the five road types used by MOVES. This mapping was consistent with the correlation presented in Table 4-2.

Table 4-2. Road Classes Mapped to MOVES Road Types

MOVES Road Type	Road Class	Area Classification
1 - Off-Network	00 - Off Grid	All
	9 -Pedestrian Way	All
2 - Rural Restricted Access	0 - Thru Way/Cross Country	Rural
	1 - Limited Access	Rural
	2 - Highway	Rural
	7 - Ramp	Rural
3 - Rural Unrestricted Access	3 - Major Road	Rural
	4 - Local Road	Rural
	5 - Minor Road	Rural
	6 - Other Road	Rural
4 - Urban Restricted Access	0 - Thru Way/Cross Country	Urban
	1 - Limited Access	Urban
	2 - Highway	Urban
	7 - Ramp	Urban
5 - Urban Unrestricted Access	3 - Major Road	Urban
	4 - Local Road	Urban
	5 - Minor Road	Urban
	6 - Other Road	Urban

Removal of Waiting Times at Test Center

Having assigned road types to the raw driving data on a second-by-second basis, the next step in our methodology was to begin the QA/QC process for the driving data. The first issue that was discovered was that many of the vehicle conditioning trips in the dataset included various periods of idle, along with some sporadic low-speed driving in the parking lot. This occurred both prior to driving of the designated test cycle, and also after the driving for the designated test cycle was completed. These instances were identified both by their latitude and longitude, and by their designation as being part of the "off-grid" road class, and thus were removed from the dataset. Any disconnected fragments of driving that were left at the beginning or the end of the trip after the removal of the parking lot periods were also removed from the dataset. This resulted in the deletion of 475,000 seconds from the dataset. An example of the type of data that was removed is shown in Figure 4-2. The data removed is represented by the red line, while the data kept is shown by the black line.

Figure 4-2. Waiting Times at Test Center

/proj1/DOT_MOVES/DriveCycle/readinKC.sas 06JUL10 02:04
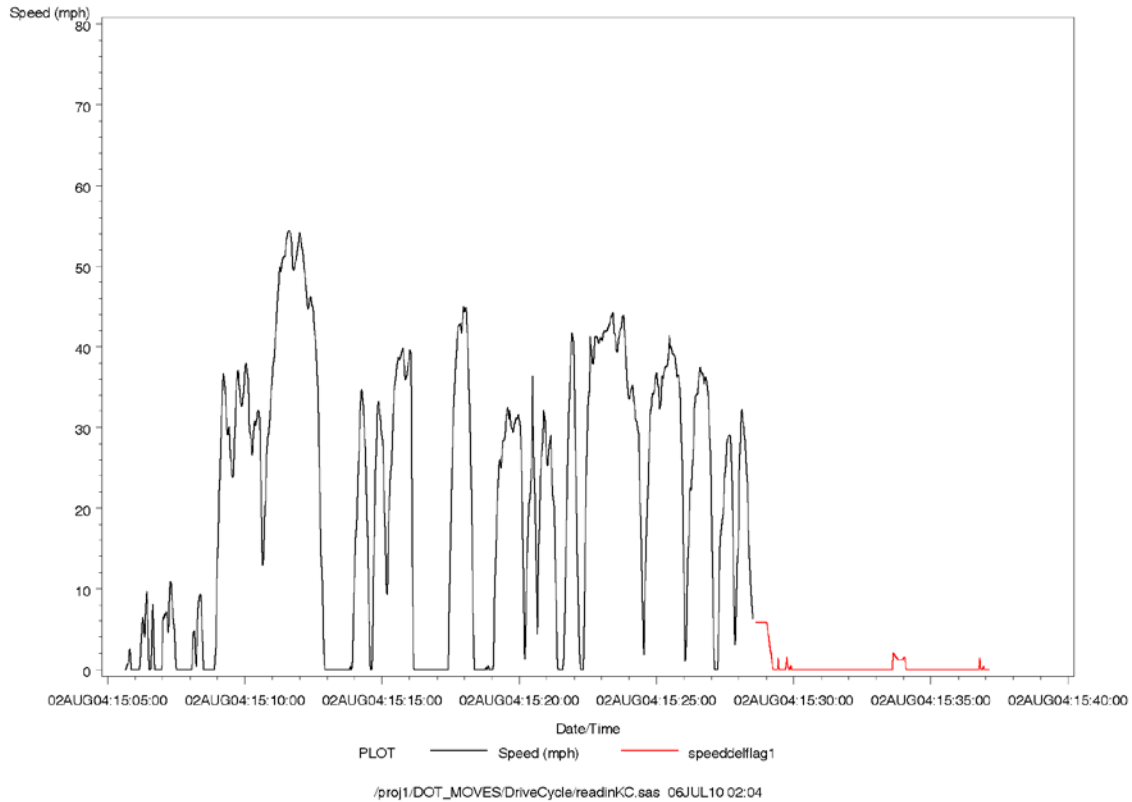
Blocks of Time with Stuck Speeds

Under normal driving conditions, it is rare for a vehicle to travel at the exact same speed for more than a few seconds. Even while driving on the freeway, a vehicle's speed tends to vary. In the Kansas City dataset, some blocks of time were identified for certain vehicles in which vehicle speed stayed constant for extended, improbable stretches. Additionally, a few blocks of time were found where the speed decreased perfectly evenly over several minutes, by a tenth of a mile per hour ever few seconds. Visual examination of the speed versus time history around these stretches indicated that they were not representative of real driving, but rather a stuck voltage or other data logger malfunction. Thus, any blocks of driving where the speed either remained constant for more than 12 seconds, or decreased perfectly steadily, were deleted from the dataset. Ultimately 38,000 seconds were deleted for these reasons. An example of a data logger with a stuck speed is shown in Figure 4-3. The driving with the stuck speed is shown in red.

Figure 4-3. Data Logger with Stuck Speed
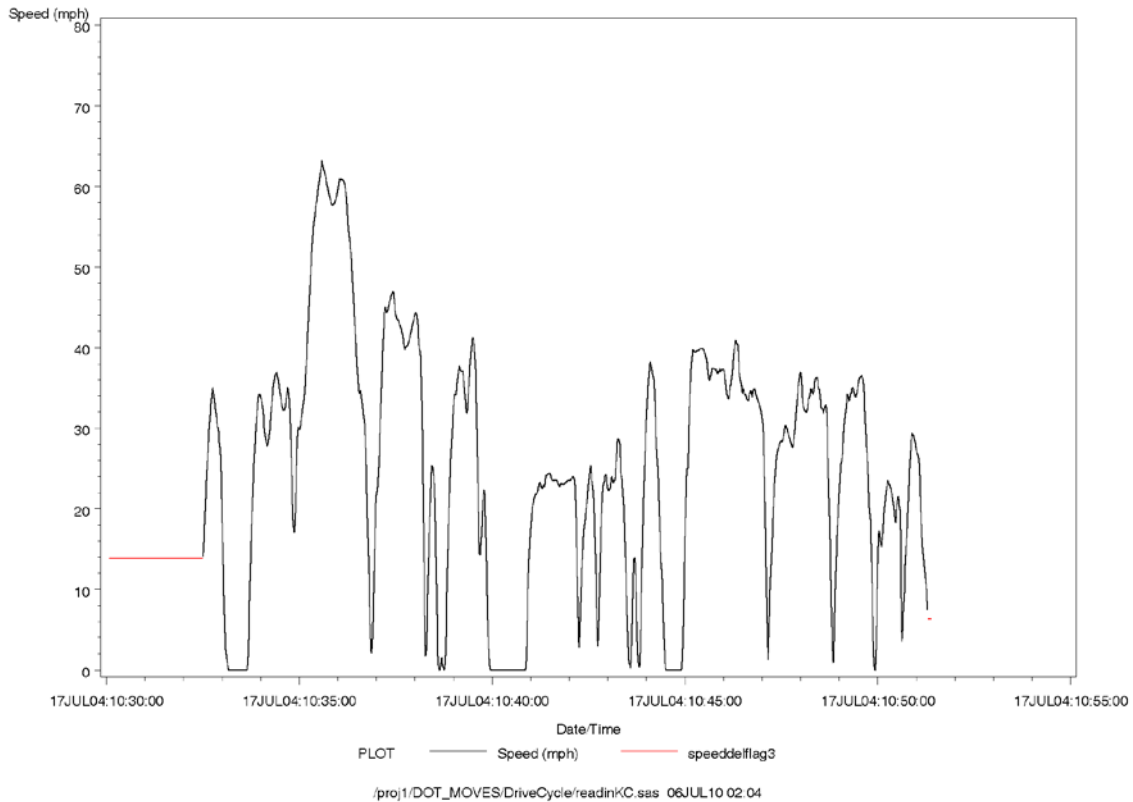
False Trips

In some instances, a number of very short trips were recorded that appeared to be electronic "blips", as opposed to actual driving. For these, the trip duration was very short, speed increased gradually and then decreased only once (as opposed to repeated increases and decreases in speed that are usually seen), and the peak speed was very low – always less than 20 mph, but usually much lower than that. Therefore, a total of 1,211,000 seconds comprising these false trips were deleted. An example of this type of false trip is shown in red in Figure 4-4.

Figure 4-4. False Trips in Drive-Away Driving
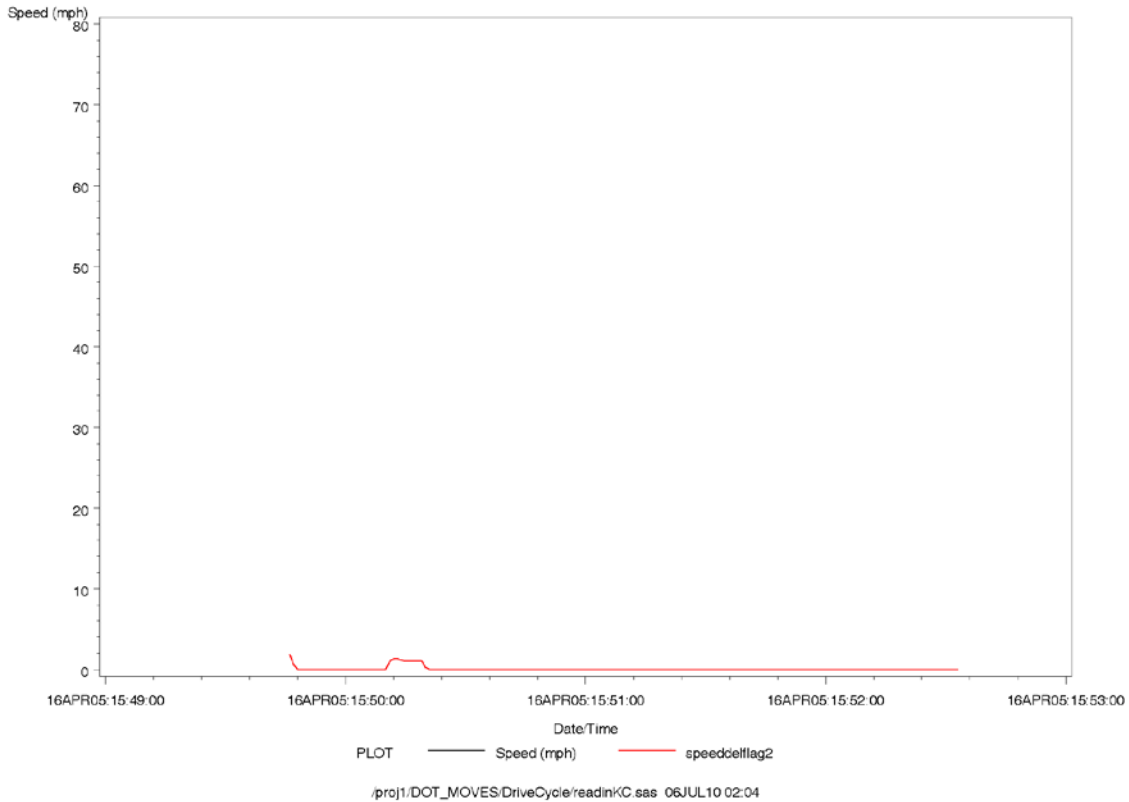
Engine Off Times

During the Kansas City study, the data loggers continued to record vehicle data even when the vehicle was turned off. Usually, engine RPM would be useful in the determination of whether the vehicle is at idle, or turned off; however, engine RPM was not always successfully stored by the data loggers in this study. Mass air flow (MAF) was stored, and was considered for use in differentiating between engine off and idle periods, but it was found that the MAF signal was very noisy, so that even when the vehicle engine was off, the MAF signal fluctuated above zero. Thus the vehicle's speed was the only useful variable in determining whether the vehicle was on or off. Any zero speed traces at the beginning of a trip, before any driving was done, and at the end of a trip, after all driving was done, were deleted from the dataset, which accounted for 1,306,000 seconds of data. This may have resulted in the accidental deletion of some periods where the vehicle was actually idling, rather than turned off, but this should not significantly affect the final cycles. After all of the edits discussed above were made, the remaining dataset contained records for 1,596,000 seconds of data.

Assignment of Trip and Micro-Trip Starts

The driving cycles were created as a set of "micro-trips" with speed and acceleration characteristics that best matched the speed and acceleration of driving in the entire dataset. A micro-trip is defined as a contiguous speed trace of vehicle driving, and is made up of an engine idle, followed by all non-idle driving until the next idle begins. A single vehicle trip may be composed of numerous micro-trips. Each separate trip in the raw Kansas City dataset was labeled with a unique ID by the data logger, and these trip designations were not changed. The next step in the process was to sub-divide the driving in the raw data into micro-trips before the cycles were built. We established the beginning of a new micro-trip:

- When a new trip began; or

- If the current speed was zero and the previous second's speed was non-zero.

Whenever a new micro-trip was detected, the numeric identifier for the micro-trip was incremented. Micro-trip numbers for the entire dataset were unique.

Binning of Continuous Variables

To use the cycle development approach discussed below in Section 4.2, all of the micro-trips in the edited dataset needed to have all of their second-by-second observations binned in terms of speed and acceleration. While the size of the bins is arbitrary, bins in general need to be narrow enough to resolve important emissions effects. In addition, bins need to be sufficiently narrow to distinguish different micro-trips for low speed/low acceleration micro-trips where those variables do not vary over a large range. On the other hand, from a practical perspective, the number of bins needs to be small so that the program that selects micro-trips can run in a reasonable amount of time.

For the cycle development in this project, we used the following binning schemes to bin the data:

- **Speed** – the continuous speed values in miles per hour were rounded to the nearest 2 miles per hour. For example, 5.6 miles per hour was assigned to bin 6, 6.9 miles per hour was assigned to bin 6, and 4.9 miles per hour was assigned to bin 4.

- **Acceleration** – Acceleration values in miles per hour per second were rounded to the nearest one mile per hour per second.

Criteria for Skipping Micro-Trips for a Cycle

Three types of micro-trips were excluded from use in the candidate cycles. First, some micro-trips consisted of entirely idle operation. These micro-trips were not used since a dedicated idle cycle was not needed for use in the MOVES model.

Second, any micro-trips less then 20 seconds in duration were not considered for inclusion in the cycles. The reason for not including these is that many short micro-trips can be produced by common, but non-representative, operation of the vehicle. One example of such operation is when a vehicle starts moving from a standstill, but the engine dies because the clutch is let out too quickly. We have found in this study, as well as in past studies, that micro-trips longer than 20 seconds are adequate to describe the vehicle driving behavior of the entire dataset taken as a whole.

The third type of micro-trip that was excluded from consideration for the cycles was any micro-trip with only one non-zero second of driving.

Assignment of Micro-Trips to Road Type and Speed Bins

For this study, our intent was to create an array of drive cycles for driving on different road types and at different average speeds, similar to the array of cycles that make up the defaults drive cycles in MOVES. Initially, we proposed to developed four cycles: Rural Restricted, Rural Unrestricted, Urban Restricted, and Urban Unrestricted, corresponding to the four on-network road types contained in the MOVES model. However, once the dataset QA/QC process was complete, it was clear that there was far more urban data available than rural, and that there were not enough rural micro-trips to use as a pool from which to derive a driving cycle. Therefore, it was decided to divide the data among only two road types: Restricted (freeways, ramps, etc.), and Unrestricted (local roads).

Even with only two categories for the road types, it was found that the majority of micro-trips included driving on both of the two road types. In large part, this issue was an artifact of using latitude/longitude coordinates to determine what road that the vehicle was on. It was common to see in the dataset stretches of driving that was almost entirely on one road, except that every few seconds, a single record was seen with a different road name. These turned out to be the intersections that the vehicle was driving through, because the coordinates for an intersection are the same for either of the cross streets. Since we needed to assign each micro-trip to a single bin, the assignments were made based on the road type that was found for the majority of the seconds of the micro-trip. A typical mixed-type micro-trip is shown in Figure 4-

5. Restricted (highway) driving is shown by the black line, while unrestricted-road (local) driving is shown in red. The figure shows that the predominant driving in the micro-trip is on the highway, with brief switches to unrestricted roads, which can only be explained by intersections crossing the highway (either over it, or under it). Data processing for work such as this will be greatly facilitated, and the accuracy of road-type assignments will be improved, when software to correlate road names with GPS coordinates is developed to look at the vehicle's trajectory in addition to its latitude and longitude.

Figure 4-5. Highway Driving Interspersed with Unrestricted Cross Streets
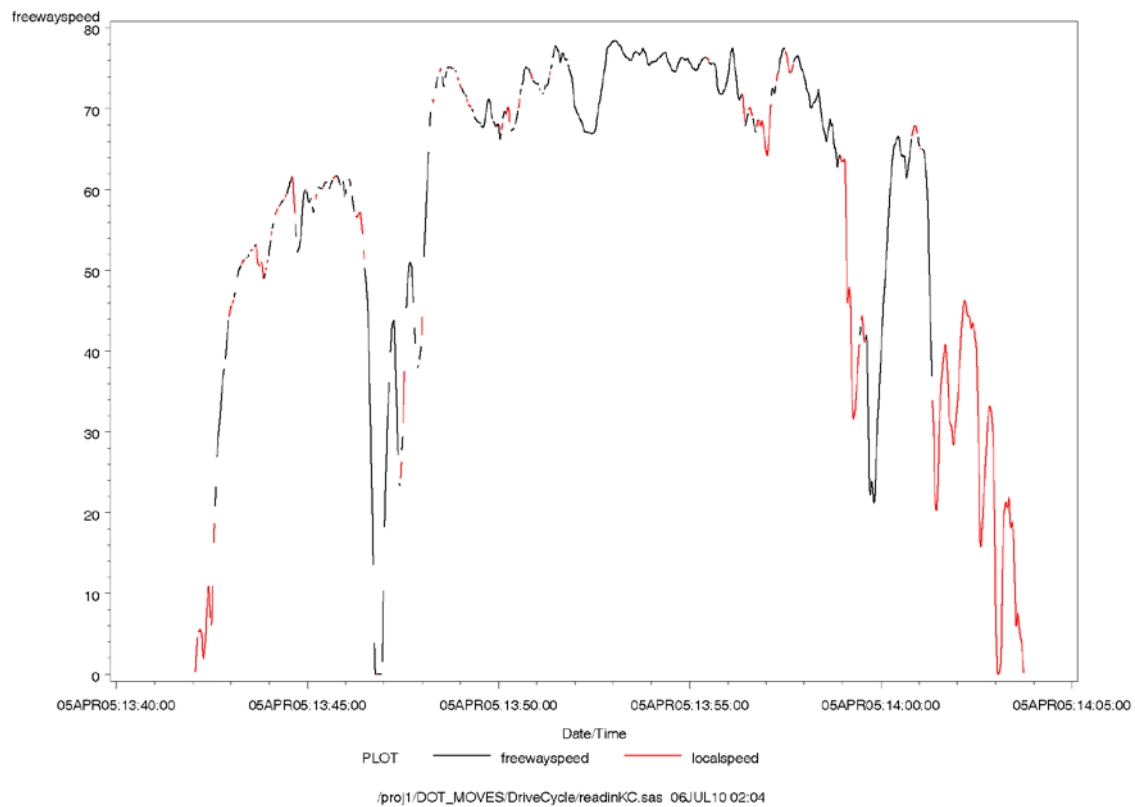
After each micro-trip was assigned to a specific MOVES road type, we then created speed bins. Each micro-trip was assigned to one speed bin, based on the mean speed for the micro-trip. Since an extended idle period at the beginning of a micro-trip could dramatically reduce the mean speed for the micro-trip, all idle periods were separated from the non-idle portion of the dataset.

The speed bins that we initially attempted to use in the binning process were designed to match the mean speeds for the driving cycles for light duty vehicles that currently exist in MOVES. However, Table 4-3 shows that these bins did not match the distribution of the driving

4-12

data for Kansas City, in that the lower bins contain most of the micro-trips, while the higher-speed bins are under populated. Therefore, a second set of speed bins was developed, with the intent of dividing the bins such that each would contain approximately the same amount of data. The second set of speed bins is listed in Table 4-4, and these were the speed bins that were ultimately used. One drive cycle was developed for each of the twelve road type/mean speed bin combinations in Table 4-4. Comparison of the two tables shows that the Kansas City drive cycles will consist of generally lower average speeds than the MOVES drive cycles currently do. It is possible that this discrepancy can be attributed to the use of latitude/longitude to determine an appropriate road type. That is, some of the micro-trips with lower mean speeds that are labeled as Restricted do not have the cruising speed characteristics and long duration that are typically expected of freeway driving, and some of these micro-trips might actually be misclassified Unrestricted-road driving. (In previous analyses, prior to the availability of GPS coordinates associated with driving activity, micro-trips were assigned as restricted/unrestricted based only the mean speed of the trip.)

The idle data that was separated from the non-idle driving data was then evaluated to determine the average idle length for micro-trips from each of the road type/speed bin combinations. It was found that the idle length was close to 20 seconds for all of the road type/speed bin combinations. That amount of idle time was added back to each micro-trip for the drive cycle.

Table 4-3. First Pass Speed Bins, Based on Default MOVES Drive Cycle Speed Ranges

Road Type	Speed Range (mph)	Number of MicroTrips	Total Seconds of Driving
Restricted	00.0-22.5	933	51,516
Restricted	22.5-37.5	1868	217,178
Restricted	37.5-50.0	481	142,060
Restricted	50.0-60.0	258	222,670
Restricted	60.0-70.0	33	32,284
Restricted	70 +	1	109
Unrestricted	00.0-15.0	917	40,154
Unrestricted	15.0-25.0	2136	262,541
Unrestricted	25.0-35.0	1308	343,568
Unrestricted	35.0-47.5	164	42,206
Unrestricted	47.5-60.0	18	12,656

Table 4-4. Final Speed Bins, Based on Uniform Distribution of Trips Across Speed Ranges

Road Type	Speed Range (mph)	Number of MicroTrips	Total Seconds of Driving
Restricted	00-20	543	30,222
Restricted	20-30	1610	123,354
Restricted	30-40	857	160,872
Restricted	40-50	272	96,306
Restricted	50-60	258	222,670
Restricted	60+	34	32,393
Unrestricted	00-15	917	40,154
Unrestricted	15-20	815	88,395
Unrestricted	20-25	1321	174,146
Unrestricted	25-28	600	165,059
Unrestricted	28-32	555	145,134
Unrestricted	32+	335	88,237

4.2   Drive Cycle Development

Representative drive cycles can be built from raw driving data using different methodologies. The methodology we have chosen for this study's cycle is to use pieces of real driving, called micro-trips, derived from available second-by-second driving data, and processed using a series of SAS programs developed by ERG.[25] When these micro-trips are connected together, they can be expected to represent driving behavior from the area of interest. The drive cycles we built were based on parameters of vehicle operation and usage that are known to be closely related to exhaust emissions. By using this approach of matching vehicle operation between measured driving behavior and the candidate cycle, it can be inferred that the emissions behavior of vehicles over the cycle will be similar to the emissions behavior of vehicles on the road.

In the creation of these driving cycles, we have chosen vehicle speed and acceleration as the variables that are important to exhaust emissions. These variables together provide a measure of the load on the engine, which is an important variable associated with exhaust emissions. In this study, we are building a cycle only for warmed up operation of light duty vehicles. That is, we are not building special cycles for cold starts and warm starts[26]. We assume that all data in the datasets represent warmed-up driving.

---

[25] The methodology used in this study was previously developed in ERG's *Roadway-Specific Driving Schedules for Heavy-Duty Vehicles*, prepared for EPA in August 2003.

[26] Drive cycles were only developed for warmed up operations because MOVES exhaust emissions are based on warmed up operations. Emissions associated with vehicle starts are calculated in MOVES using a methodology that is not connected with drive cycles.

General Methodology

   A strategy based on minimizing the difference between a cycle vector C representing the driving in the candidate cycle and a target vector T representing the driving in the activity database for the case was used to select micro-trips from the database for inclusion in the cycle. As micro-trips are used to build-up a candidate cycle, the difference between the two vectors tends to become smaller and smaller. The build-up process ends when the cycle developer decides that the two vectors are substantially the same and the duration of the cycle that has been built up is acceptable. The multi-dimensional space that these vectors are in will be described shortly, but first let us consider how the build-up process works for developing a cycle.

   The goal of building the cycle is to select micro-trips such that when their vectors $\mathbf{M_1}$ are added together, the vector C of the resulting cycle is as similar as possible to the target vector $\mathbf{T}$ of the activity database. Figure 4-6 shows the hypothetical situation of the vectors after two micro-trips have been used to create a cycle. In this hypothetical example, the first micro-trip was selected from the activity database for the case as the one whose vector M1 was closest to the target vector $\mathbf{T}$ for the database. Then, a second micro-trip is searched for such that when its vector $\mathbf{M_2}$ is added to $\mathbf{M_1}$ to create the resultant vector $\mathbf{C}$ shown in Figure 3-1, the distance between the tips of $\mathbf{C}$ and $\mathbf{T}$ is minimized. This distance is the length of the vector $\mathbf{T\text{-}C}$ as denoted in the figure by the dashed vector. As micro-trips are added to create the built-up cycle represented by $\mathbf{C}$, the length of $\mathbf{T\text{-}C}$ is calculated after each additional micro-trip is added to the cycle to follow the progress of the build-up process. It should be noted that the order of the micro-trips in the final cycle is unimportant from the point of view of the selection of the micro-trips. The reason for this is that the resultant $\mathbf{C}$ is independent of the order in which the micro-trip vectors $\mathbf{M_i}$ are added together.

Figure 4-6. Vector Description of Comparing Target and Cycle Activity
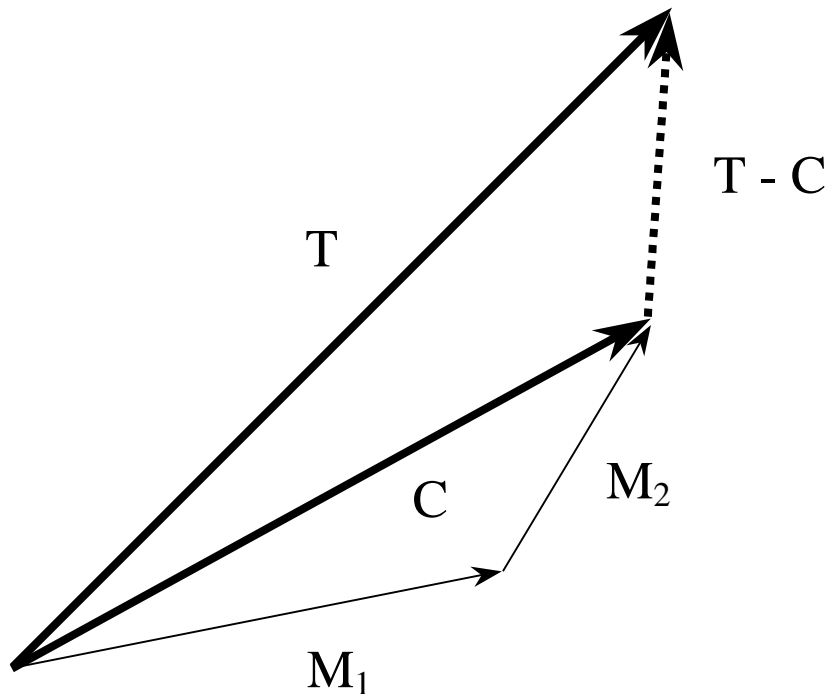
It should also be noted that we are forcing micro-trips to be added to the candidate cycle. This is done even if the addition of the best incremental micro-trip causes the length of **T-C** to increase in some instances. Generally, as the cycle is built up there will be a decrease in the length of **T-C**. After several micro-trips have been added, the length of **T-C** may increase slightly. Later, with the addition of more micro-trips, a "discovery" will be made that will produce a relatively abrupt decrease in the length of **T-C** so that the accumulated cycle will be substantially better than the cycle was much earlier in the build-up process.

All of the vectors used above to describe the build-up process are based on representations of the frequency distributions of observations in cumulative speed, acceleration, vehicle specific power space. This statement requires some explanation. A segment of driving, whether it is a micro-trip, a piece of a driving cycle, or the entire activity database, can be described as a frequency distribution. The distribution consists of combinations of two variables: speed and acceleration. The continuous values for these variables were converted into frequency distributions through the use of bins. Each observation in the database was placed in a particular speed/acceleration bin. The <u>cumulative</u> frequency distribution is made up of the number of observations that fall "below" the current bin for each of the two-binned variables. The binning criteria for the two variables is described above in Section 4.1. To help the reader understand the

process, we will present a numerical example in one dimension and another example in two dimensions to demonstrate how the comparison of the vectors **T** and **C** works.

Suppose we wanted to compare a candidate cycle with the database using a single vehicle operation variable that was monitored second-by-second in the collection of data for the activity database. The single variable might be engine load. In this hypothetical example, we have 35,900 one-second observations of engine load in the target activity database and 68 one-second observations in the cycle. The first step in comparing **T** and **C** is to bin the observations of load in the target data and in the cycle data. Table 4-5 shows the binning of the hypothetical data in Columns 2 and 3. Note that the target counts in Column 2 are much larger than the cycle counts in Column 3. This is a consequence of the activity database containing all of the observations for all micro-trips, while the cycle has just one micro-trip. The counts in Columns 2 and 3 were converted to cumulative counts in Columns 4 and 5. This is done to provide proximity information for the micro-trip searching algorithm. In other words, we wanted the algorithm to be able to select a micro-trip even if the observations for a given micro-trip were not in exactly the same bins as the target, but did have observations at least in a nearby bin. The use of the cumulative distributions helps ensure that proximity information is available.

Table 4-5. Comparison of Cycle and Target Vectors for a Hypothetical One-Dimensional Example

Bin	Counts		Cumulative Counts		Vector (Normalized Cumulative Counts)		Vector Length		
	Target	Cycle	Target	Cycle	Target	Cycle	T	C	T-C
1	1000	0	1000	0	0.028	0.000	1.246	1.266	0.138
2	11000	30	12000	30	0.334	0.441			
3	7000	10	19000	40	0.529	0.588			
4	6000	7	25000	47	0.696	0.691			
5	4500	5	29500	52	0.822	0.765			
6	2800	1	32300	53	0.900	0.779			
7	1500	4	33800	57	0.942	0.838			
8	800	6	34600	63	0.964	0.926			
9	600	1	35200	64	0.981	0.941			
10	700	4	35900	68	1.000	1.000			

A comparison of the cumulative counts for the target and cycle information in Columns 4 and 5 shows that if we used these counts to create the **T** and **C** vectors, the lengths of the vectors would be greatly different simply because the target vector, which is made up of the 10 elements in Column 4, would be a much longer vector then the cycle vector, which is made up of the 10 elements in Column 5. Accordingly, we normalize the target and cycle cumulative counts in 4

and 5 to produce the target vector elements and the cycle vector elements as the fractional values between 0 and 1 shown in Columns 6 and 7.

The values in Columns 6 and 7 become the elements of the **T** and **C** vectors, which are in 10-dimensional space. A visualization of the elements of these vectors is provided in Figure 4-7. This figure shows the normalized cumulative counts of the target and cycle from Columns 6 and 7 as a function of the bin number. What we want to do in developing the cycle is select micro-trips so that the curve for the cycle is as close as possible to the curve for the target in this figure. The way we do this is to minimize the sums of the squares of the differences between the value for the corresponding elements of the target and cycle vectors. This corresponds to the square of the length of **T-**C. Table 3-1 shows the calculated length of **T**, **C**, and **T-C**. These lengths can be determined from the values of the elements for **T** and **C** in Columns 6 and 7 using the standard relationship for determining the length of a vector if its elements are known.

Figure 4-7. Visual Comparison of Vector Elements
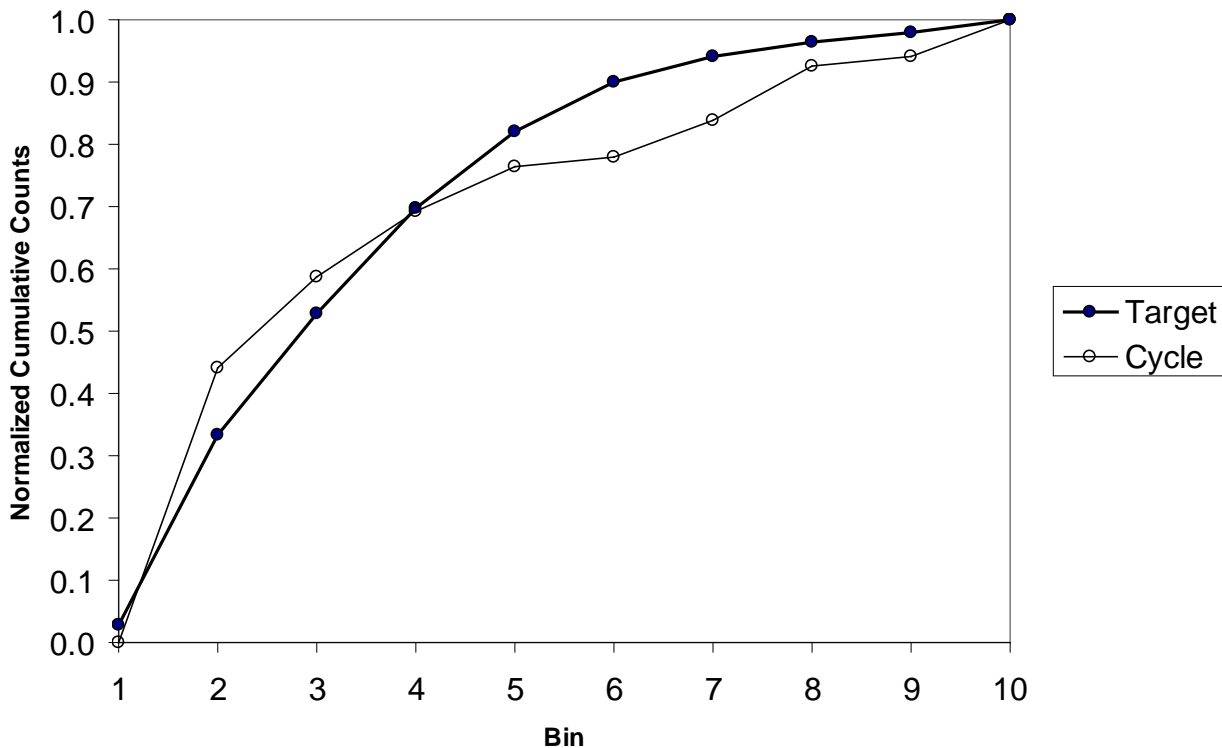
Extension of the one-dimensional example shown in Table 4-5 and Figure 4-7 to multiple dimensions is demonstrated by the spreadsheet calculations shown in Table 4-6. In this example, 100 matrix elements are used. The table shows 10 rows which might be accelerations and 10 columns which might be speeds. The left side of Table 3-2 shows the calculations for the target

4-18

matrix and the right side shows the calculations for the cycle matrix. In Tables a) and b), the second-by-second observations of the target and cycle data are binned. The numbers in each bin represent the frequency of observations that meet the criteria for those bins. In Tables c) and d), the counts in the Tables a) and b) are accumulated across each row. Then, in Tables e) and f), the accumulated frequencies in Tables c) and d) are accumulated down each column. This produces a field of frequencies on a cumulative basis that run from a low value in the upper left corner of each matrix to a high number in the lower right corner of each matrix. The value in the lower right hand corner of Tables e) and f) is equal to the total number of observations in the target or cycle matrix. These total observation numbers in the lower right hand corner of e) and f) are used to normalize all of the frequencies in Tables e) and f) to arrive at the normalized cumulative matrices in g) and h). The values in g) and h) are then used to calculate the square of the differences in each corresponding matrix element to produce the values in Table i). The value in Table j) is just the summation of all of the elements of Table i) and represents the square of the length of the **T-C** vector. This is the value that we attempt to minimize when selecting micro-trips for the cycle.

Note that the counts in a) and b) did not need to be in corresponding bins for this comparison process to work. The use of cumulative distributions permitted the two matrices to be compared successfully.

Extension of the technique to the third dimension for vehicle specific power or any number of higher dimensions could be made by analogy.

Table 4-6. Comparison of Cycle and Target Matrices for a Hypothetical Two-Dimensional Example

Target Activity Matrix

a) Count the second-by-second observations in each bin.

	A	B	C	D	E	F	G	H	I	J
1	2									
2		1								
3		2		5						
4			5		3		2	1		
5		5		9	1			2	9	3
6			2			4	1			
7										
8			6			1				
9		1								
10										

c) Accumulate the above frequencies across each row

	A	B	C	D	E	F	G	H	I	J
1	2	2	2	2	2	2	2	2	2	2
2	0	1	1	1	1	1	1	1	1	1
3	0	2	2	7	7	7	7	7	7	7
4	0	0	5	5	8	8	10	11	11	11
5	0	5	5	14	15	15	15	17	26	29
6	0	0	2	2	2	6	7	7	7	7
7	0	0	0	0	0	0	0	0	0	0
8	0	0	6	6	6	7	7	7	7	7
9	0	1	1	1	1	1	1	1	1	1
10	0	0	0	0	0	0	0	0	0	0

e) Accumulate the above frequencies down each column.

	A	B	C	D	E	F	G	H	I	J
1	2	2	2	2	2	2	2	2	2	2
2	2	3	3	3	3	3	3	3	3	3
3	2	5	5	10	10	10	10	10	10	10
4	2	5	10	15	18	18	20	21	21	21
5	2	10	15	29	33	33	35	38	47	50
6	2	10	17	31	35	39	42	45	54	57
7	2	10	17	31	35	39	42	45	54	57
8	2	10	23	37	41	46	49	52	61	64
9	2	11	24	38	42	47	50	53	62	65
10	2	11	24	38	42	47	50	53	62	65

g) Normalize the elements in the above matrix.

	A	B	C	D	E	F	G	H	I	J
1	0.031	0.031	0.031	0.031	0.031	0.031	0.031	0.031	0.031	0.031
2	0.031	0.046	0.046	0.046	0.046	0.046	0.046	0.046	0.046	0.046
3	0.031	0.077	0.077	0.154	0.154	0.154	0.154	0.154	0.154	0.154
4	0.031	0.077	0.154	0.231	0.277	0.277	0.308	0.323	0.323	0.323
5	0.031	0.154	0.231	0.446	0.508	0.508	0.538	0.585	0.723	0.769
6	0.031	0.154	0.262	0.477	0.538	0.600	0.646	0.692	0.831	0.877
7	0.031	0.154	0.262	0.477	0.538	0.600	0.646	0.692	0.831	0.877
8	0.031	0.154	0.354	0.569	0.631	0.708	0.754	0.800	0.938	0.985
9	0.031	0.169	0.369	0.585	0.646	0.723	0.769	0.815	0.954	1.000
10	0.031	0.169	0.369	0.585	0.646	0.723	0.769	0.815	0.954	1.000

Cycle Activity Matrix

b) Count the second-by-second observations in each bin.

	A	B	C	D	E	F	G	H	I	J
1										
2		1								
3				4						
4		4				3		1		
5								4	1	
6				8						2
7					3					
8							1			
9	1	5								
10										

d) Accumulate the above frequencies across each row

	A	B	C	D	E	F	G	H	I	J
1	0	0	0	0	0	0	0	0	0	0
2	0	1	1	1	1	1	1	1	1	1
3	0	0	0	4	4	4	4	4	4	4
4	0	4	4	4	4	7	7	8	8	8
5	0	0	0	0	0	0	0	4	5	5
6	0	0	0	8	8	8	8	8	8	10
7	0	0	0	0	3	3	3	3	3	3
8	0	0	0	0	0	0	1	1	1	1
9	1	6	6	6	6	6	6	6	6	6
10	0	0	0	0	0	0	0	0	0	0

f) Accumulate the above frequencies down each column.

	A	B	C	D	E	F	G	H	I	J
1	0	0	0	0	0	0	0	0	0	0
2	0	1	1	1	1	1	1	1	1	1
3	0	1	1	5	5	5	5	5	5	5
4	0	5	5	9	9	12	12	13	13	13
5	0	5	5	9	9	12	12	17	18	18
6	0	5	5	17	17	20	20	25	26	28
7	0	5	5	17	20	23	23	28	29	31
8	0	5	5	17	20	23	24	29	30	32
9	1	11	11	23	26	29	30	35	36	38
10	1	11	11	23	26	29	30	35	36	38

h) Normalize the elements in the above matrix.

	A	B	C	D	E	F	G	H	I	J
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026
3	0.000	0.026	0.026	0.132	0.132	0.132	0.132	0.132	0.132	0.132
4	0.000	0.132	0.132	0.237	0.237	0.316	0.316	0.342	0.342	0.342
5	0.000	0.132	0.132	0.237	0.237	0.316	0.316	0.447	0.474	0.474
6	0.000	0.132	0.132	0.447	0.447	0.526	0.526	0.658	0.684	0.737
7	0.000	0.132	0.132	0.447	0.526	0.605	0.605	0.737	0.763	0.816
8	0.000	0.132	0.132	0.447	0.526	0.605	0.632	0.763	0.789	0.842
9	0.026	0.289	0.289	0.605	0.684	0.763	0.789	0.921	0.947	1.000
10	0.026	0.289	0.289	0.605	0.684	0.763	0.789	0.921	0.947	1.000

i) Calculate the squares of the differences in corresponding elements of the above two matrices.

	A	B	C	D	E	F	G	H	I	J
1	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
2	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.001	0.003	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	0.001	0.003	0.000	0.000	0.002	0.002	0.000	0.000	0.000	0.000
5	0.001	0.000	0.010	0.044	0.073	0.037	0.050	0.019	0.062	0.087
6	0.001	0.000	0.017	0.001	0.008	0.005	0.014	0.001	0.021	0.020
7	0.001	0.000	0.017	0.001	0.000	0.000	0.002	0.002	0.005	0.004
8	0.001	0.000	0.049	0.015	0.011	0.010	0.015	0.001	0.022	0.020
9	0.000	0.014	0.006	0.000	0.001	0.002	0.000	0.011	0.000	0.000
10	0.000	0.014	0.006	0.000	0.001	0.002	0.000	0.011	0.000	0.000

j) Sum the squares of the differences.

0.754

4.3 Development of the Twelve Drive Cycles for Inclusion in MOVES

The dataset containing all of the second-by-second driving activity for the Kansas City vehicles was edited and prepared as described in Section 4.1, resulting in a set of micro-trips distributed into twelve bins according to mean speed and dominant road type. The approach described in Section 4.2 was then used to select the micro-trips that would comprise a drive cycle for each of the twelve bins. The micro-trip speed and acceleration information was used to find those micro-trips which, when concatenated, best described the driving activity for each bin. For each bin, ERG's SAS-based cycle development program converted each micro-trip into a speed/acceleration vector and the entire set of micro-trips for that bin into the speed/acceleration target vector **T**. The program then found the micro-trip whose sum of the squares difference between the cumulative normalized elements of the micro-trip with the corresponding elements of the target was the smallest. This corresponded to finding the micro-trip such that the **T-C** vector was the smallest. This became the first micro-trip in the cycle, for that bin. Then, the program looked through all remaining micro-trips from the bin to find the best second micro-trip such that when it was added to the first micro-trip the new vector **T-C** had a minimum length. This process may be repeated until the developer wants to stop searching. In this study, we stopped searching after 25 micro-trips were added to a cycle.

Finally, a 20-second idle period was added to the beginning of each micro-trip, representing the mean idle time that was found for micro-trips from each road type/ speed bin combination.

Plots of the square of the length of the **T-C** vector as micro-trips were added to the cycle for each bin are presented in Appendix B in Figures B-1a through B-1l. These are the sum of the squared differences between the cycle vector **C** and the target vector **T**. The left vertical axis of each figure shows that the square of the length of the **T-C** vector drops continuously as additional micro-trips are added to the cycle (red line). The right vertical axis of each figure shows the cumulative duration of the cycle as additional micro-trips are added to the cycle (dashed black line). What these figures reveal is that none of the cycles really need all 25 micro-trips to adequately represent the driving for the bin. For each of the twelve cycles, little additional benefit was achieved by adding micro-trips after about the 15th trip; the sum of the squared differences between the vectors did not continue to decrease after that point. Additionally, the use of 25 micro-trips tended to make the final cycle much longer than the cycles of 1000 seconds or so that are currently used in MOVES. This was especially true for the higher-speed bins, whose micro-trips tended to be much longer than the micro-trips for the lower-speed bins. Therefore, we elected to use only the first 15 micro-trips for each of the cycles,

with the exception of the two highest speed cycles. These are the Restricted 50-60 cycle, for which only the first five micro-trips would be used, and the Restricted 60+ cycle, for which only the first seven micro-trips would be used. These two cycles contain micro-trips that are very long, but the figures show slight minimums in the values of the sum of the squared differences, at the fifth and the seventh micro-trip, respectively.

Speed versus Time Traces for Candidate Cycles

After the micro-trips for the twelve cycles were selected, several evaluations of the cycles were made, including comparisons of the cycle to the target dataset. First, we examined a speed versus time plot of the micro-trips that make up each cycle, shown in Appendix B in Figures B-2a through B-2l. The small circles on the plot indicate the beginning of each micro-trip. These candidate cycle plots were used to examine the overall appearance of the cycle and to show the duration of the cycle. These figures show that the micro-trips for the different road types and speed bins clearly represent different types of driving.

Comparison of Statistics for Cycle and Target

We also made comparisons of the chosen cycles to the corresponding target datasets. In particular, we wished to confirm that the speed and acceleration characteristics of the two were similar. We first examined scatter plots of acceleration versus speed for the candidate cycle and for the target dataset. Rather than show all 24 plots here, only a sample are shown in Appendix B in Figures B-3a through B-3d: a comparison of speed versus acceleration for the cycle and the target for the Restricted 50-60 MPH bin, and for the Unrestricted 25-28 MPH bin. All of the data points for the cycle are shown on the two cycle plots, but due to the large size of the target databases, only a random subset of the target data is shown on either of the two target plots. Comparison of Figures B-3a and B-3b shows that the Cycle and the Target contain a very similar speed versus acceleration profile, with the highest density of data between speeds of 60 and 70 MPH and accelerations between -2 to 2 MPH/S, somewhat lighter concentration between 30 and 40 MPH. However, while the target database contains quite a few observations with speeds near 80 MPH, there are fewer of these high speeds in the cycle. For Figures B-3c and B-3d, the cycle and the target again look similar. Here the densest area of points is at about 30 MPH, and again, we see more observations at the highest speeds (above 45 MPH) in the target than in the cycle. It appears that a few trips in the target bins contain outliers: short amounts of especially high speed (for that bin) driving. Because they are few in number, they do not comprise the majority of the driving for the bin, and do not get selected for inclusion in the cycle. This was observed for most of the 12 bins, and has been seen in ERG's previous cycle building projects as well.

The frequency distributions of speeds found in the new cycles and the corresponding target databases are shown in Figures B-4a through B-4d. Again, this is only a sample, for the Restricted 50-60 MPH bin and the Unrestricted 25-28 MPH bin. From these figures we see once more that while the speed distributions for the cycles and the targets are similar, the target does contain a few slightly higher speeds than the cycle.

A number of statistics were calculated so that the characteristics of the cycles could be compared to the characteristics of the respective target datasets. These statistics are listed in Table 4-7. It is important to remember when comparing any of these statistics that the micro-trips in the each cycle were selected only because their non-idle speed and acceleration characteristics match those of the target. Any other statistics that are calculated and compared were not the basis, or at least not the direct basis, for choosing the micro-trips for the cycles.

Table 4-7 shows that the average second-by-second speeds of the selected cycle and target database are similar, although the cycle speeds are slightly lower than those in the target dataset. This corresponds to what was observed in Figures B-3 and B-4 in Appendix B: there are a few especially high speeds in each target bin that are not representative of the target bin as a whole, and thus were not selected for inclusion in the cycles. They do, however, raise the mean speed of the target slightly above that of the cycle. Finally, Table-4-7 shows that the square of the length of **T-C** was very low for each of the cycles. This indicates a good fit of the driving conditions (speed and acceleration) in the selected cycle, compared to the target database.

Table 4-7. Comparison of Dataset and Cycle Operation Characteristics

	Target			Cycle						
	Average Speed (mph)	Average Micro-Trip Time (s)	Average Micro-Trip Distance (miles)	Average Speed (mph)	Average Micro-Trip Time (s)	Average Micro-Trip Distance (miles)	Total Time (s)	Total Distance (miles)	Micro-Trips (count)	Final Square of L. of T-C
Rest. 0-20	16	56	0.25	12	75	0.26	1127	3.91	15	0.012
Rest. 20-30	26	77	0.55	20	87	0.49	1300	7.38	15	0.013
Rest. 30-40	35	188	1.85	32	189	1.70	2842	25.45	15	0.013
Rest. 40-50	45	354	4.43	42	320	3.77	4801	56.54	15	0.015
Rest. 50-60	55	863	13.12	54	834	12.43	4170	62.15	5	0.022
Rest. 60+	63	953	16.57	61	713	12.05	4989	84.37	7	0.054
Unrst. 0-15	9	44	0.11	6	62	0.11	924	1.66	15	0.028
Unrst. 15-20	18	108	0.55	15	94	0.38	1405	5.76	15	0.013
Unrst. 20-25	23	132	0.83	19	109	0.57	1634	8.60	15	0.011
Unrst. 25-28	27	275	2.03	24	181	1.20	2713	18.00	15	0.011
Unrst. 28-32	30	262	2.15	27	186	1.38	2783	20.72	15	0.01
Unrst. 32+	39	263	2.86	35	190	1.85	2845	27.73	15	0.037

MOVES Modeling of Kansas City Drive Cycles

Drive cycles within the MOVES database are stored in three tables of interest: *driveschedule*, *drivescheduleassoc*, and *driveschedulesecond*. Per discussions with EPA staff[27], it was necessary to completely replace the existing MOVES light duty drive cycles, in order to avoid potential interpolation issues during model calculations. This was done by removing selected light duty drive cycles from the default database, and importing the new drive cycle data described above using completely new *drivescheduleIDs*.

First, using the default *driveschedule* table as a basis, we eliminated all existing light duty drive cycles with the exception of *drivescheduleIDs* 101 and 199. In this case, the former was intended to simulate very low speeds (average 2.5 mph), and the latter was intended to simulate activity on ramps. Both instances were such that we did not feel they could be accurately simulated with the Kansas City drive cycles we developed dataset, so they were left intact. The *driveschedule* table was then populated with the proper data.

Similarly, the default *drivescheduleassoc* and *driveschedulesecond* tables were cleansed of existing light duty cycles, and new information was imported. In the case of *drivescheduleassoc*, Restricted cycles were associated with MOVES roadtypeIDs 2 and 4, while Unrestricted cycles were associated with MOVES roadtypeIDs 3 and 5. Associations were made for MOVES sourcetypeIDs 11, 21, 31, and 32, which correspond to motorcycles, passenger cars, passenger trucks, and light commercial trucks, respectively. These associations are consistent with the default drive cycles removed from the MOVES database. In the case of *driveschedulesecond*, new data was added for each cycle on a second-by-second basis, using speeds calculated during the development process described above.

Having populated updated tables containing new Kansas City study-based cycles for light duty vehicles, it was necessary to then replace each of the existing MOVES default drive cycle tables and perform additional modeling. Before doing so, ERG used the MySQL Administrator program, which is typically included with during a normal MOVES installation, to make a SQL backup of the existing three tables of interest. This is important for two reasons: first, to have a backup of the affected tables in case there is an error during the import of the new drive cycles tables; and secondly, to be able to easily switch back to the default MOVES drive cycles for future modeling. Users should be aware that alteration of drive cycle tables in this way will

---

[27] Note that we are not aware of any formal guidance from EPA on replacement of drive cycles in the model. We consulted directly David Brzezinski and Sean Hillson at EPA during this stage of our study.

affect all future modeling runs done on an individual computer system, so it's crucial to backup the default drive cycles in such a way that they can be easily restored.[28]

Loading of the updated drive cycle tables was done with a simple SQL script, in which the existing default tables were truncated, and the new tables were imported in the form of .tab-delimited text files. Having successfully loaded the new tables, a backup was made again, using MySQL administrator, for ease of switching between drive cycle sets. Finally, the TDM-based and HPMS-based MOVES runs described in Sections 2 and 3 were re-created and re-run using the updated drive cycles to provide additional model outputs for further analysis. Comparisons of model outputs are discussed in Sections 5 and 6 below.

For reference, all of the MOVES database tables and SQL scripts described in this section are provided as electronic files in Appendix A.

---

[28] Note that the most current version of MOVES at this writing, MOVES2010a, allows users to alter the drive cycle tables discussed in this section via the County Data Manager GUI, rendering obsolete the procedure described here for importing alternate drive cycles into MOVES.

5.0   Model Outputs

MOVES outputs from the model runs performed during this study are summarized in this section. These include the TDM-based and HPMS-based model runs originally performed for each of the eight counties in the Houston-Galveston non-attainment area, as well as the model runs conducted for the same region and VMT bases using the alternate Kansas City drive cycles described above in Section 4. These outputs are described in additional detail below.

5.1   Comparison of MOVES Outputs to MOBILE-6 Based Inventory

First, ERG compared the two sets of MOVES runs conducted using the default drive cycles to the original TTI MOBILE6 based inventory[29], which was previously described in detail in Section 2.0. Table 5-1 presents ERG's modeled TDM-based and HPMS-based MOVES model outputs, along with TTI's MOBILE6 inventory emissions in tons, by county, for a Non-School ozone season weekday. Table 5-2 presents the percent difference between both sets of MOVES runs and the original TTI inventory. Note that according to these tables, MOVES is calculating significantly higher NOx emissions across the board, regardless of the VMT basis used for model inputs. There is a wider variance in emissions for CO and VOC; counties with relatively few VMT return a decrease in CO and VOC, while counties with more VMT are associated with a significant increase in CO and VOC. The largest county (in terms of both VMT and population), Harris county, demonstrates much higher emissions of all three pollutants of interest relative to the MOBILE6 inventory.

Note that generally, as shown in Table 5-1, the HPMS VMT being modeled is somewhat higher (for most counties) relative to the modeled MOBILE6 inventory, which in turn leads to higher emissions relative to that same inventory. We conclude here that it's very important, regardless of the VMT basis chosen for calculation, for users to be sure that their activity data is representative, since the model outputs are highly sensitive to it.

---

[29] All references to the MOBILE6-based TTI inventory in this section are based on TTI's *Production of MOVES On-Road Mobile, Link-Based Emissions Estimates and Document Preparation Technical Note*, submitted to TCEQ in July 2010, in which TTI assessed the 2005 and 2006 years using both MOVES and MOBILE6. This is the most recently available document from TTI that discusses summaries of their MOBILE6 –based inventory for 2005 and 2006. This document was produced subsequent to TTI's *2005/2006 On-Road Mobile Source, Ozone Episode Modeling Emission Inventories for the HGB Eight-Hour Nonattainment Area,* which was the source of the data we based this analysis upon originally.

Table 5-1. Summary of VMT and Tons of Emissions, by County, Non-school Ozone Season Weekday

County	*ERG MOVES, TDM-based VMT*				*ERG MOVES, HPMS-based VMT*				*TTI MOBILE6 Inventory*			
	VMT	VOC	CO	NOx	VMT	VOC	CO	NOx	VMT	VOC	CO	NOx
Brazoria	5,708,326	4.233	46.032	11.457	5,810,915	4.427	47.051	12.852	5,728,510	3.659	47.044	8.704
Chambers	2,508,488	0.950	17.900	4.638	2,565,380	1.043	21.123	5.389	2,519,042	1.815	29.611	5.072
Fort Bend	7,948,707	5.493	61.357	15.746	8,382,481	5.834	64.235	18.171	7,994,337	4.662	62.013	11.965
Galveston	5,816,257	4.117	48.031	11.762	5,941,357	4.218	51.160	13.253	5,850,101	3.89	47.37	8.696
Harris	96,544,278	78.159	870.197	204.910	104,118,424	81.555	936.121	241.000	97,155,857	66.433	792.438	147.484
Liberty	2,524,022	1.539	20.206	5.223	2,369,354	1.532	19.093	5.334	2,529,328	2.291	29.303	4.502
Montgomery	10,096,087	6.379	76.684	20.107	10,079,347	6.510	77.152	21.582	10,141,862	6.173	82.879	16.415
Waller	1,946,771	1.490	19.040	4.188	1,880,278	1.526	19.375	4.433	1,949,623	1.785	24.601	3.901
TOTAL	**133,092,936**	**102.359**	**1159.447**	**278.032**	**141,147,536**	**106.645**	**1235.310**	**322.014**	**133,868,660**	**90.708**	**1115.259**	**206.739**

Table 5-2. Percent Difference between MOVES and MOBILE6 Inventory Results

County	*% Difference, TDM VMT basis, relative to MOBILE6*				*% Difference, HPMS VMT basis, relative to MOBILE6*			
	VMT	VOC	CO	NOx	VMT	VOC	CO	NOx
Brazoria	-0.4%	15.7%	-2.2%	31.6%	1.4%	21.0%	0.0%	47.7%
Chambers	-0.4%	-47.7%	-39.6%	-8.6%	1.8%	-42.5%	-28.7%	6.3%
Fort Bend	-0.6%	17.8%	-1.1%	31.6%	4.9%	25.1%	3.6%	51.9%
Galveston	-0.6%	5.8%	1.4%	35.3%	1.6%	8.4%	8.0%	52.4%
Harris	-0.6%	17.7%	9.8%	38.9%	7.2%	22.8%	18.1%	63.4%
Liberty	-0.2%	-32.8%	-31.0%	16.0%	-6.3%	-33.1%	-34.8%	18.5%
Montgomery	-0.5%	3.3%	-7.5%	22.5%	-0.6%	5.5%	-6.9%	31.5%
Waller	-0.1%	-16.5%	-22.6%	7.4%	-3.6%	-14.5%	-21.2%	13.6%
TOTAL	**-0.6%**	**12.8%**	**4.0%**	**34.5%**	**5.4%**	**17.6%**	**10.8%**	**55.8%**

Tables 5-3 and 5-4 present emissions and VMT split by light-duty and heavy duty vehicle classification. Consistent with summaries found in TTI's July 2010 *Technical Note*, vehicle classes in the MOBILE6 summaries are defined as follows:

- *Light duty vehicles* include those classified as LDGV, LDGT1, LDGT2, LDGT3, LDGT4, LDDV, LDDT12, MC, and LDDT34.

- *Heavy duty vehicles* include all other classes.

Similarly, TTI defined light- and heavy-duty vehicles in MOVES output summaries as follows:

- *Light duty vehicles* include the source types Motorcycles, Passenger Cars, Passenger Trucks, and Light commercial Trucks.

- *Heavy duty vehicles* include all MOVES source type.

For consistency we have duplicated this classification scheme in the following tables. The reader will immediately note in Table 5-4 the large discrepancy in VMT when comparing heavy duty VMT modeled in MOVES to the heavy duty VMT originally modeled by TTI in MOBILE6. For example, the HD VMT modeled under the TDM-based VMT scenario is between 35 and 40 percent less than that modeled in MOBILE6. This is a result of attempting to directly compare emissions aggregated under MOBILE6 *vehicle* types (which classify vehicles on a GVWR weight basis) versus emissions aggregated under MOVES source types (which classify vehicles based on their expected mode of operation). In particular, vehicle types HDDV2B through HDDV5 inclusive (along with their gasoline counterparts), considered heavy-duty in MOBILE6, map to MOVES source types 31 (passenger trucks) and 32 (light commercial trucks). These are considered "light-duty" under the TTI classification. Because of this, it is difficult to draw conclusions about emissions when differentiating by vehicle classification, but we include the tables below for reference.

Table 5-3. Summary of VMT and Tons of Emissions, by County and Vehicle Classification, Non-school Ozone Season Weekday

ERG MOVES, TDM-based VMT								
	VMT		VOC		CO		NOx	
County	Light Duty	Heavy Duty	Light Duty	Heavy Duty	Light Duty	Heavy Duty	Light Duty	Heavy Duty
Brazoria	5,384,087	324,239	3.873	0.360	43.806	2.226	6.511	4.945
Chambers	2,368,948	139,540	0.836	0.113	17.175	0.725	2.636	2.002
Fort Bend	7,486,363	462,344	4.929	0.563	57.827	3.530	8.396	7.349
Galveston	5,490,933	325,324	3.725	0.392	45.727	2.304	6.566	5.196
Harris	91,292,024	5,252,254	71.501	6.658	828.843	41.353	117.144	87.766
Liberty	2,379,054	144,968	1.394	0.145	19.192	1.014	3.119	2.104
Montgomery	9,487,457	608,630	5.720	0.659	72.649	4.035	10.620	9.487
Waller	1,834,222	112,549	1.393	0.096	18.427	0.614	2.593	1.596
TOTAL	**125,723,088**	**7,369,848**	**93.373**	**8.986**	**1103.645**	**55.802**	**157.586**	**120.446**

ERG MOVES, HPMS-based VMT								
	VMT		VOC		CO		NOx	
County	Light Duty	Heavy Duty	Light Duty	Heavy Duty	Light Duty	Heavy Duty	Light Duty	Heavy Duty
Brazoria	5,350,307	460,608	3.922	0.505	43.709	3.342	6.399	6.453
Chambers	2,361,987	203,393	0.881	0.162	19.870	1.253	2.749	2.640
Fort Bend	7,718,045	664,436	5.048	0.786	58.974	5.261	8.483	9.689
Galveston	5,470,464	470,893	3.692	0.525	47.672	3.488	6.474	6.779
Harris	95,863,528	8,254,896	72.015	9.540	870.072	66.049	119.679	121.321
Liberty	2,181,503	187,851	1.347	0.185	17.744	1.349	2.860	2.474
Montgomery	9,280,219	799,128	5.682	0.828	71.543	5.609	10.299	11.283
Waller	1,731,203	149,075	1.397	0.129	18.457	0.918	2.521	1.912
TOTAL	**129,957,256**	**11,190,280**	**93.986**	**12.659**	**1148.040**	**87.270**	**159.462**	**162.552**

TTI MOBILE6 Inventory								
	VMT		VOC		CO		NOx	
County	Light Duty	Heavy Duty	Light Duty	Heavy Duty	Light Duty	Heavy Duty	Light Duty	Heavy Duty
Brazoria	5,188,114	540,396	3.424	0.235	45.400	1.644	3.726	4.978
Chambers	2,302,383	216,659	1.735	0.080	28.854	0.757	2.050	3.022
Fort Bend	7,227,617	766,720	4.303	0.359	59.541	2.472	4.859	7.106
Galveston	5,310,752	539,349	3.633	0.257	45.601	1.770	3.713	4.984
Harris	88,607,611	8,548,246	62.147	4.285	762.826	29.612	64.143	83.341
Liberty	2,289,017	240,311	2.194	0.096	28.581	0.722	2.228	2.275
Montgomery	9,135,188	1,006,674	5.734	0.439	79.721	3.157	6.491	9.924
Waller	1,769,669	179,955	1.718	0.069	24.013	0.589	1.752	2.149
TOTAL	**121,830,351**	**12,038,310**	**84.888**	**5.820**	**1074.537**	**40.723**	**88.962**	**117.779**

Table 5-4. Percent Difference between MOVES and MOBILE6 Inventory Results, by Vehicle Classification

% Difference, TDM VMT basis, relative to MOBILE6								
	VMT		**VOC**		**CO**		**NOx**	
County	**Light Duty**	**Heavy Duty**	**Light Duty**	**Heavy Duty**	**Light Duty**	**Heavy Duty**	**Light Duty**	**Heavy Duty**
Brazoria	3.8%	-40.0%	13.1%	53.1%	-3.5%	35.4%	74.8%	-0.7%
Chambers	2.9%	-35.6%	-51.8%	41.6%	-40.5%	-4.2%	28.6%	-33.7%
Fort Bend	3.6%	-39.7%	14.6%	56.9%	-2.9%	42.8%	72.8%	3.4%
Galveston	3.4%	-39.7%	2.5%	52.5%	0.3%	30.2%	76.8%	4.3%
Harris	3.0%	-38.6%	15.1%	55.4%	8.7%	39.7%	82.6%	5.3%
Liberty	3.9%	-39.7%	-36.5%	51.3%	-32.9%	40.5%	40.0%	-7.5%
Montgomery	3.9%	-39.5%	-0.2%	50.0%	-8.9%	27.8%	63.6%	-4.4%
Waller	3.6%	-37.5%	-18.9%	39.8%	-23.3%	4.2%	48.0%	-25.7%
TOTAL	**3.2%**	**-38.8%**	**10.0%**	**54.4%**	**2.7%**	**37.0%**	**77.1%**	**2.3%**

% Difference, HPMS VMT basis, relative to MOBILE6								
	VMT		**VOC**		**CO**		**NOx**	
County	**Light Duty**	**Heavy Duty**	**Light Duty**	**Heavy Duty**	**Light Duty**	**Heavy Duty**	**Light Duty**	**Heavy Duty**
Brazoria	3.1%	-14.8%	14.6%	114.8%	-3.7%	103.3%	71.7%	29.6%
Chambers	2.6%	-6.1%	-49.2%	102.2%	-31.1%	65.6%	34.1%	-12.6%
Fort Bend	6.8%	-13.3%	17.3%	118.9%	-1.0%	112.8%	74.6%	36.3%
Galveston	3.0%	-12.7%	1.6%	104.4%	4.5%	97.1%	74.4%	36.0%
Harris	8.2%	-3.4%	15.9%	122.6%	14.1%	123.0%	86.6%	45.6%
Liberty	-4.7%	-21.8%	-38.6%	92.3%	-37.9%	86.8%	28.3%	8.8%
Montgomery	1.6%	-20.6%	-0.9%	88.7%	-10.3%	77.7%	58.7%	13.7%
Waller	-2.2%	-17.2%	-18.7%	86.5%	-23.1%	55.9%	43.9%	-11.0%
TOTAL	**6.7%**	**-7.0%**	**10.7%**	**117.5%**	**6.8%**	**114.3%**	**79.2%**	**38.0%**

Tables 5-5 and 5-6 present emissions and VMT split by on-network and off-network process classification. Processes associated with running exhaust and evaporative emissions were classified as on-network, while off-network processes included emissions due to vehicle starts and extended idling. This classification is consistent with that proposed in TTI's July 2010 *Technical Note*.

Table 5-5. Summary of VMT and Tons of Emissions, by County and Emissions Process Classification, Non-school Ozone Season Weekday

ERG MOVES, TDM-based VMT								
	VMT		VOC		CO		NOx	
County	On-Network	Off-Network	On-Network	Off-Network	On-Network	Off-Network	On-Network	Off-Network
Brazoria	5,708,326	0	2.655	1.578	33.711	12.321	9.896	1.561
Chambers	2,508,488	0	0.776	0.173	16.787	1.112	4.471	0.167
Fort Bend	7,948,707	0	3.465	2.027	45.518	15.839	13.631	2.114
Galveston	5,816,257	0	2.747	1.370	37.998	10.033	10.371	1.391
Harris	96,544,278	0	51.308	26.850	672.636	197.561	177.997	26.914
Liberty	2,524,022	0	1.153	0.386	17.275	2.931	4.858	0.364
Montgomery	10,096,087	0	4.254	2.125	59.757	16.927	17.963	2.144
Waller	1,946,771	0	0.945	0.544	14.301	4.740	3.720	0.469
TOTAL	**133,092,936**	**0**	**67.304**	**35.055**	**897.982**	**261.464**	**242.907**	**35.124**

ERG MOVES, HPMS-based VMT								
	VMT		VOC		CO		NOx	
County	On-Network	Off-Network	On-Network	Off-Network	On-Network	Off-Network	On-Network	Off-Network
Brazoria	5,810,915	0	2.825	1.602	34.699	12.351	11.231	1.621
Chambers	2,565,380	0	0.863	0.180	20.005	1.118	5.211	0.179
Fort Bend	8,382,481	0	3.765	2.068	48.346	15.889	15.958	2.213
Galveston	5,941,357	0	2.829	1.389	41.105	10.056	11.818	1.435
Harris	104,118,424	0	54.261	27.294	738.071	198.050	213.114	27.886
Liberty	2,369,354	0	1.143	0.389	16.157	2.936	4.959	0.375
Montgomery	10,079,347	0	4.367	2.144	60.201	16.951	19.389	2.193
Waller	1,880,278	0	0.979	0.547	14.631	4.744	3.956	0.477
TOTAL	**141,147,536**	**0**	**71.032**	**35.613**	**973.216**	**262.094**	**285.635**	**36.379**

TTI MOBILE6 Inventory								
	VMT		VOC		CO		NOx	
County	On-Network	Off-Network	On-Network	Off-Network	On-Network	Off-Network	On-Network	Off-Network
Brazoria	5,728,510	0	1.540	2.119	35.645	11.399	8.074	0.630
Chambers	2,519,042	0	0.680	1.135	23.798	5.813	4.752	0.320
Fort Bend	7,994,337	0	2.105	2.557	47.126	14.887	11.194	0.771
Galveston	5,850,101	0	1.705	2.185	35.768	11.602	8.077	0.619
Harris	97,155,857	0	29.493	36.940	601.275	191.163	136.906	10.578
Liberty	2,529,328	0	0.883	1.408	22.528	6.775	4.128	0.374
Montgomery	10,141,862	0	2.615	3.558	62.455	20.424	15.355	1.060
Waller	1,949,623	0	0.637	1.148	19.379	5.222	3.613	0.288
TOTAL	**133,868,660**	**0**	**39.658**	**51.050**	**847.974**	**267.285**	**192.099**	**14.640**

Table 5-6. Percent Difference between MOVES and MOBILE6 Inventory Results, by Emissions Process Classification

% Difference, TDM VMT basis, relative to MOBILE6							
	VMT		VOC		CO		NOx	
County	On-Network	Off-Network	On-Network	Off-Network	On-Network	Off-Network	On-Network	Off-Network
Brazoria	-0.4%	0.0%	72.4%	-25.5%	-5.4%	8.1%	22.6%	147.7%
Chambers	-0.4%	0.0%	14.1%	-84.7%	-29.5%	-80.9%	-5.9%	-47.8%
Fort Bend	-0.6%	0.0%	64.6%	-20.7%	-3.4%	6.4%	21.8%	174.2%
Galveston	-0.6%	0.0%	61.1%	-37.3%	6.2%	-13.5%	28.4%	124.7%
Harris	-0.6%	0.0%	74.0%	-27.3%	11.9%	3.3%	30.0%	154.4%
Liberty	-0.2%	0.0%	30.6%	-72.6%	-23.3%	-56.7%	17.7%	-2.6%
Montgomery	-0.5%	0.0%	62.7%	-40.3%	-4.3%	-17.1%	17.0%	102.2%
Waller	-0.1%	0.0%	48.4%	-52.6%	-26.2%	-9.2%	2.9%	62.8%
TOTAL	**-0.6%**	**0.0%**	**69.7%**	**-31.3%**	**5.9%**	**-2.2%**	**26.4%**	**139.9%**

% Difference, HPMS VMT basis, relative to MOBILE6							
	VMT		VOC		CO		NOx	
County	On-Network	Off-Network	On-Network	Off-Network	On-Network	Off-Network	On-Network	Off-Network
Brazoria	1.4%	0.0%	83.4%	-24.4%	-2.7%	8.4%	39.1%	157.4%
Chambers	1.8%	0.0%	27.0%	-84.2%	-15.9%	-80.8%	9.7%	-44.2%
Fort Bend	4.9%	0.0%	78.9%	-19.1%	2.6%	6.7%	42.6%	187.0%
Galveston	1.6%	0.0%	65.9%	-36.4%	14.9%	-13.3%	46.3%	131.9%
Harris	7.2%	0.0%	84.0%	-26.1%	22.8%	3.6%	55.7%	163.6%
Liberty	-6.3%	0.0%	29.4%	-72.4%	-28.3%	-56.7%	20.1%	0.2%
Montgomery	-0.6%	0.0%	67.0%	-39.8%	-3.6%	-17.0%	26.3%	106.9%
Waller	-3.6%	0.0%	53.7%	-52.3%	-24.5%	-9.2%	9.5%	65.7%
TOTAL	**5.4%**	**0.0%**	**79.1%**	**-30.2%**	**14.8%**	**-1.9%**	**48.7%**	**148.5%**

Complete summaries of emissions and VMT output by county for each model run performed, by VMT-basis (TDM or HPMS), drive cycle (default or Kansas City derived), vehicle classification (light- or heavy-duty), emissions process, and process classification (on- or off-network) are presented in Appendix C. Raw model outputs are also available in Appendix A.

5.2   Comparison of MOVES Default Drive Cycle Output to Alternate Drive Cycle Outputs

In performing comparisons of the modeling performed using MOVES default drive cycles to the modeling conducted using the Kansas City drive cycles developed by ERG, we limited our summaries to only those source types affected by the drive cycle changes. These included motorcycles, passenger cars, passenger trucks, and light commercial trucks (source types IDs 11, 21, 31, and 32, respectively).

Table 5-7 presents a comparison of modeled light duty emissions for the combined 8-county Houston-Galveston area. Note that the Kansas City drive cycles produce slightly lower emissions for both VMT bases, in the range of approximately 1-3%. This difference is dominated by emissions from Harris county; in Appendix D, it can be seen that percent differences range from 1-17% on a by-county basis, with emissions from counties with fewer VMT showing the widest variance. It can also be seen that VOC tends to vary less than the other two pollutants modeled.

Table 5-7. Comparison of Eight County Light Duty Emissions for Both Drive Cycles and VMT Bases (Overall tons)

	Default Cycle			KC Cycle			% Difference		
VMT Basis	**CO**	**NOx**	**VOC**	**CO**	**NOx**	**VOC**	**CO**	**NOx**	**VOC**
HPMS	1148.04	159.46	93.99	1129.63	157.46	92.83	-1.60%	-1.25%	-1.23%
TDM	1103.64	157.59	93.37	1073.96	153.37	91.97	-2.69%	-2.67%	-1.50%

Table 5-8 presents a comparison of combined light duty emissions by MOVES source type. The largest differences between the modeled drive cycles are seen in motorcycles, which are a relatively small portion of the light-duty fleet. Overall differences across source types and counties range from approximately +1% to -3%. As before, these emissions are dominated by results from Harris county, and other more rural counties exhibit wider variance in emissions between modeled drive cycles, as shown Appendix D.

Table 5-8. Comparison of Eight County Light Duty Emissions for Both Drive Cycles and VMT Bases (Tons, by Source Type)

VMT Basis	Source	Default Cycle			KC Cycle			% Difference		
		CO	NOx	VOC	CO	NOx	VOC	CO	NOx	VOC
HPMS	Motorcycles	12.62	0.49	2.66	11.74	0.45	2.63	-6.97%	-9.07%	-1.26%
	Passenger Cars	597.44	78.59	57.21	590.05	77.23	56.69	-1.24%	-1.73%	-0.90%
	Passenger Trucks	364.98	48.49	22.60	358.27	47.79	22.19	-1.84%	-1.45%	-1.81%
	Light Commercial Trucks	173.00	31.90	11.52	169.56	32.00	11.32	-1.99%	0.33%	-1.73%
TDM	Motorcycles	2.75	0.11	1.70	2.56	0.10	1.70	-6.92%	-8.91%	-0.44%
	Passenger Cars	588.28	78.81	57.55	575.52	76.43	56.91	-2.17%	-3.02%	-1.11%
	Passenger Trucks	346.57	47.11	22.56	335.32	45.64	22.04	-3.25%	-3.13%	-2.27%
	Light Commercial Trucks	166.04	31.55	11.56	160.56	31.20	11.32	-3.30%	-1.11%	-2.09%

Table 5-9 presents combined light-duty emissions by MOVES road type. Variations in this table are fairly pronounced. Although there is virtually no change in off-network emissions (which is expected, as no drive cycles for that road type was modified), the Kansas City restricted road types (freeways) have approximately 3-8% higher emissions, while unrestricted road types (local roads) have emissions that are about 5-15% lower. These differences are independent from the rural/urban distinction, which is expected, since as described above we did not have enough microtrips in the Kansas City Data to make that distinction in the development of our final drive cycles. The variation also seems to be independent of VMT calculation basis.

Table 5-9. Comparison of Eight County Light Duty Emissions for Both Drive Cycles and VMT Bases (Tons, by Road Type)

VMT Basis	Road Type	Default Cycle			KC Cycle			% Difference		
		CO	NOx	VOC	CO	NOx	VOC	CO	NOx	VOC
HPMS	Off-Network	254.27	29.35	52.59	254.27	29.35	52.53	0.00%	0.00%	-0.13%
	Rural Restricted	36.66	4.80	1.18	37.41	4.81	1.17	2.04%	0.32%	-0.13%
	Rural Unrestricted	99.23	16.50	5.29	84.53	14.18	4.92	-14.82%	-14.06%	-7.06%
	Urban Restricted	378.54	48.43	12.93	406.51	52.47	13.28	7.39%	8.32%	2.67%
	Urban Unrestricted	379.33	60.39	22.00	346.90	56.66	20.94	-8.55%	-6.17%	-4.81%
TDM	Off-Network	254.27	29.35	52.53	254.27	29.35	52.46	0.00%	0.00%	-0.15%
	Rural Restricted	197.29	25.91	6.67	209.24	27.48	6.80	6.06%	6.06%	1.95%
	Rural Unrestricted	404.17	66.26	22.81	364.38	60.38	21.59	-9.85%	-8.87%	-5.34%
	Urban Restricted	121.79	15.48	4.15	131.25	16.95	4.27	7.77%	9.46%	2.95%
	Urban Unrestricted	126.13	20.58	7.22	114.82	19.21	6.86	-8.97%	-6.67%	-4.99%

Table 5-10 presents combined light duty emissions by MOVES emission process ID. For CO and NOx, the only difference between modeled drive cycles is attributable to running exhaust and crankcase running exhaust, and is on the order of about 1-4% less, depending on

VMT basis. There is no difference at all in start emissions. For VOC, there is also no difference in start emissions, and further, no difference in evaporative emissions either. This is expected, since start emissions and evaporative emissions are both off-network processes, and as such, should be unaffected by changes to on-network drive cycles. However, there is larger variance in both exhaust and refueling emissions, approximately 4-6% less, depending on the VMT basis modeled.

Table 5-10. Comparison of Eight County Light Duty Emissions for Both Drive Cycles and VMT Bases (Tons, by Emission Process ID)

VMT Basis	Process ID	Default Cycle			KC Cycle			% Difference		
		CO	NOx	VOC	CO	NOx	VOC	CO	NOx	VOC
HPMS	Running Exh	893.63	130.10	29.42	875.22	128.11	28.35	-2.06%	-1.54%	-3.65%
	Start Exh	254.24	29.35	31.09	254.24	29.35	31.09	0.00%	0.00%	0.00%
	Evap Permeation			8.04			8.04			0.00%
	Evap Fuel Vent			19.83			19.83			0.00%
	Evap Fuel Leak			3.55			3.55			0.00%
	Crank Run Exh	0.14	0.01	0.39	0.13	0.01	0.37	-1.42%	2.87%	-3.59%
	Crank Start Exh	0.03	0.00	0.41	0.03	0.00	0.41	0.00%	0.00%	0.00%
	Refuel Disp Vap			0.52			0.49			-5.39%
	Refuel Spillage			0.73			0.69			-5.42%
TDM	Running Exh	849.24	128.23	28.73	819.56	124.01	27.42	-3.50%	-3.29%	-4.56%
	Start Exh	254.24	29.35	31.09	254.24	29.35	31.09	0.00%	0.00%	0.00%
	Evap Permeation			8.04			8.04			0.00%
	Evap Fuel Vent			19.90			19.90			0.00%
	Evap Fuel Leak			3.59			3.59			0.00%
	Crank Run Exh	0.13	0.01	0.39	0.13	0.01	0.37	-2.30%	1.55%	-4.46%
	Crank Start Exh	0.03	0.00	0.41	0.03	0.00	0.41	0.00%	0.00%	0.00%
	Refuel Disp Vap			0.51			0.47			-6.25%
	Refuel Spillage			0.72			0.67			-6.21%

The information presented in Tables 7 through 10 above is also available on a by-county basis in Appendix D. As shown in these tables, the effects of implementing the drive cycles developed by ERG in MOVES varied considerably depending on which county was being modeled. The drive cycles we developed were based on largely urban driving data as discussed in Section 4, and applying these drive cycles to more rural counties in the Houston-Galveston area (as exhibited by fewer VMT driven in those counties) produced emissions that may not be representative. On the other hand, emissions from Harris County, the most urban of the counties modeled (with the corresponding largest daily VMT modeled), did not vary much at all. We believe that any alternate drive cycles developed and used in MOVES should closely match the type of driving activity for the area being modeled for emissions to be considered representative.

6.0 Conclusions

Although the technical analyses ERG performed during the course of this work are sometimes specific to the Houston and Kansas City metropolitan areas, we believe that the broadest benefit of this study will be in applying lessons learned in transitioning from traditional MOBILE6 modeling and inventory planning inputs to using MOVES effectively. With that in mind, we have included in this report not only analysis of model inputs and outputs, but also documentation of the process of running MOVES, example model runs that users can execute on their own computer systems, inventories they can compare their results against, and analysis of the ramifications of transitioning to the new MOVES system.

In the section that follows, we present important issues dealt with during the course of this project, with an emphasis on how users can deal with these issues themselves. We also discuss the study questions originally proposed in the project Work Plan.

6.1 Important Transition Issues

In order to effectively transition from use of MOBILE6 to MOVES in developing regional emissions analyses, users must not only understand how to use MOVES to develop inventories, but also keep in mind several issues arising from characterization of data that is input to both models. We list some of these significant concerns below.

Proper VMT Derivation and Inputs

Because VMT is perhaps the single most important variable in calculation of emissions, it is crucial to understand both the process of converting VMT into a format usable by MOVES, and the pros and cons of using various VMT bases. In this study, ERG explored the usage of two separate VMT types. The first was based on VMT derived from a travel demand model (TDM), and the second on VMT summarized in Highway Performance Monitoring System (HPMS) format. Users may have sources of VMT available to them in other formats as well, which may require the development of spreadsheets or other software tools in order to process the VMT for use in modeling.

Difficulties and Sensitivity to Emissions

Running exhaust emissions and evaporative hydrocarbon emissions calculated in MOVES are directly proportional to the VMT input to the model. In addition to translation of raw VMT from MOBILE6 to MOVES format, users must also be careful to properly allocate

VMT across various speed, age, and road type profiles to ensure that calculated emissions are truly representative.

In the case of VMT formats used in this study, the travel model-based VMT process generally provides a more detailed VMT basis, but it also requires a greater degree of data manipulation and processing on the part of the user. Use of HPMS-based VMT, meanwhile, requires less upfront data processing, but does require a separate MOVES model run, as described in Section 2.1, to arrive at a default source type distribution required for use in EPA's VMT converter tool. The use of HPMS based VMT also necessitates the independent development of road type distributions appropriate to the modeled area of interest.

Further, the VMT converter tools provided by EPA, while effective in calculating inputs for the MOVES model, are prone to potential user error due to their complexity. Such user errors, if unchecked, can cascade throughout the created model inputs files.

<u>Solution</u>

To ensure that VMT is appropriately developed for input to MOVES, we recommend that users take advantage, to the extent possible, of any existing MOBILE6 files (such as age distributions, speed distributions, VMT by hour, and so forth) during VMT pre-processing. Any software developed by users to process VMT and its associated distributions should be extensively quality checked to ensure that VMT transformations are being applied correctly.

Users should also make themselves familiar with the wide variety of converter tools that EPA has developed and made available. Our methodology for using these tools on both the TDM and HPMS VMT basis is discussed above in Section 2.2.

Vehicle / Source Type Characterization

The MOVES model has been built from the ground up to utilize emission factors in units of grams per hour of operation. This is a distinct difference from the grams per mile basis used in MOBILE6. In addition, the MOVES model produces emission estimates for emissions processes that could not be accurately produced by the MOBILE6 model, such as extended idling emissions. In the development of MOVES, EPA transitioned away from the standard MOBILE6 vehicle classes (based on gross vehicle weight rating, or GVWR) and instead characterizes vehicles by source type based on their expected mode of operation. However, the emission factors contained within the model are still primarily aggregated on a vehicle class/weight basis.

This requires the MOVES model to map the source types to vehicle classes internally in what is called the Source Bin Generator.

Difficulties and Sensitivity to Emissions

This change in calculation basis presents challenges to users on a number of fronts. The first is proper conversion of VMT and associated distributions. Fortunately, EPA has provided tools and documentation to assist users in transition from MOBILE6 vehicle types to MOVES source types for total VMT, speed distribution, age distribution, and vehicle populations. The second challenge this transition presents to users comes in the form of development of appropriate I/M program inputs. This is discussed in further detail in the next subsection.

The third challenge manifests itself in comparison of model outputs between MOBILE6 and MOVES. In particular, vehicle types HDDV2B through HDDV5 inclusive (along with their gasoline counterparts), considered heavy-duty in MOBILE6, map to MOVES source types 31 (passenger trucks) and 32 (light commercial trucks) in EPA's converter tools and associated guidance. These source types are often considered "light-duty" when analyzing aggregated MOVES outputs. Because of this, difficulties arise when attempting to directly compare MOBILE6 outputs by vehicle type with MOVES outputs by source type.

An example of such a difficulty in presented in Section 5.2, where a first glance at VMT output from the models indicates a disparity between light-duty and heavy duty vehicle activity. This disparity turns out to be a result of the vehicle classification scheme, and not necessarily indicative of any calculation errors in the model.

Solution

Because the vehicle/source type characterization issue is present on a number of fronts when transitioning from MOBILE6 to MOVES, users will do well to keep it in mind at all points of the process: in preparing and converting model inputs, in model execution, and in comparison and analysis of model outputs. Proper use of EPA's provided converter tools, as well as the methodology for model inputs development presented in EPA's *Technical Guidance*, will be helpful in developing VMT by source type.

Inspection/Maintenance Program Inputs

If users elect to model areas affected by local Inspection and Maintenance (I/M) programs, they need to take care to ensure that not only is the I/M representative of the area being modeled, but also that inputs are formatted properly, and conversions from MOBILE6

vehicle to MOVES source types, as alluded to above, are applied in the correct manner (if appropriate).

Difficulties and Sensitivity to Emissions

I/M program information input to the model will reduce calculated model emissions by varying degrees, depending on the type of program implemented, model year coverage, and other factors. If users elect to utilize MOVES I/M defaults for the area of interest, they should first be aware that not all such defaults are accurate, or even present at all, in the model. At the very least, users should carefully examine the defaults and make changes to the I/M program as appropriate.

Secondly, as mentioned above, one of the challenges of applying I/M programs in MOVES is that MOBILE6 vehicle types do not map precisely with MOVES source types. In particular, MOVES source types 31 and 32 can encompass MOBILE6 source types LDGT1 all the way up to HDGV5. As stated in Section 3.10.6 of the *Technical Guidance*, "Inspection and Maintenance (I/M) programs entered in MOVES can only be applied to source types. However, as discussed in section 3.10.2, this association of I/M programs and source type may be inconsistent with state I/M program regulations that define I/M programs by the vehicle weight classes. MOVES source types are composed of several vehicle weight classes and, therefore, applying I/M benefits to the entire MOVES source type may be inappropriate." Users should be aware that they may run into similar difficulties when trying to apply existing I/M programs information across source types in MOVES.

Solution

When making changes to an existing I/M program, users need to be aware that this is a tricky process in which errors can easily be introduced. In particular, it is critical to change the useIMyn field to N for existing records in the IM table, and alter the IMProgramID field for any new or updated records. ERG's methodology for preparation of I/M inputs was discussed previously in Section 2.3.

With respect to proper application of I/M programs across MOVES source types, EPA recommends in section 3.10.6 of the *Technical Guidance* the calculation of a compliance factor based not only on compliance rate and waiver rate (which both provided in MOBILE6 inputs), but also on regulatory class coverage. Calculating the regulatory class coverage is now fairly straightforward, as the information in Table A.3 of the April 2010 *Technical Guidance* Appendix (Gasoline I/M Regulatory Coverage Adjustments) provides the proper conversion factors, along with example calculations.

Vehicle Population Data Selection

There are a couple of methods for obtaining vehicle population data for input to MOVES, which is used in calculation of vehicle start and extended idle emission components in the model. We ultimately used by-county registration data provided by the Texas Department of Transportation, but if such data is not available, users can go through the process of developing populations surrogates based on VMT, as described in detail in section 2.3 of this report, as well as the EPA *Technical Guidance*.

Difficulties and Sensitivity to Emissions

Vehicle start emissions and extended idle emissions calculated in MOVES are directly proportional to the vehicle population input to the model. Although usage of registration data to develop population is suggested by EPA in the *Technical Guidance*, users should understand that the vehicles registered in a county for a given point in time do not necessarily correspond to the VMT driven over that same period in the county. Using the counties we modeled as an example, commuters living in Fort Bend county may drive a significant portion of their vehicle miles in Harris county. In this case, using registration data as a population surrogate may lead to overestimation of start and evaporative emissions in Fort Bend county, while underestimating those same emissions in Harris County.

Solution

While a better alternative for determining representative source populations does not currently exist, this is an important issue for users to be aware of. Ultimately, users must be confident that whatever population data they choose to input to the model is accurate.

MOVES Bugs and Errata

For this study, ERG used the first non-draft version of the MOVES model (MOVES2010) released by EPA in late December 2009. The EPA has released subsequent versions of the model since that time, and will continue to update and improve the model going forward into the future. However, in the course of their modeling efforts, users may, like we did, discover discrepancies in model outputs that only be explained by problems in MOVES.

While using MOVES2010, ERG found two bugs in model calculations during the course of this study, both mentioned previously in Section 2.3. One of these bugs was related to fuel type IDs and associated oxygenate volume percentages. The other bug we found has been documented in the latest version of the MOVES Errata, and involved user creation of new fuel

formulation IDs for input to MOVES. When a user creates and imports new fuel formulation IDs, MOVES does not currently "zero out" the market share of existing fuel formulation IDs.

<u>Difficulties and Sensitivity to Emissions</u>

It is difficult to speculate on how general calculation errors in the model might effect emissions output. In fact, some bugs that users come across may not have an effect on emissions at all, but rather relate to user interface difficulties, Java execution errors, or other issues. In our case, the first bug we found (related to fuel type IDs) led to underestimation of VOC emissions, underestimating them by approximately an order of magnitude[30] The second bug we encountered, involving user creation of fuel formulation IDs, led to a pollutant overestimation of approximately two to four times in model calculations.

<u>Solution</u>

Careful examination of outputs in each case led us to contact EPA staff, who provided us with temporary workarounds for each bug while adjustments were made to the model itself. Other potential issues are explained in the currently available version of the MOVES Errata[31]. Users should carefully examine these Errata before performing a MOVES analysis, and be ready to report to EPA any potential issues identified in their own outputs for correction in future versions of the model.

6.2 Study Questions

In this section we attempt to answer the questions originally proposed in the Work Plan submitted at the beginning of this study. These questions included the following:

What is the best approach for creating an emissions inventory with MOVES based on activity data obtained from a travel demand model (TDM)?

We believe the best approach for developing a MOVES emissions inventory using TDM activity data, as described in detail in Sections 2 and 3, involves a number of steps. First, methodical preparation of TDM VMT data, consistent with EPA's *Technical Guidance*, must be undertaken, in which VMT, VHT, and associate activity distributions are derived with an eye on inputs to EPA's multiple Converter Tools. Second, these Converter Tools must be properly utilized to arrive at inputs that can directly input into the MOVES County Data Importer itself.

---

[30] MOVES tables of interest related to this bug include etohbin and fuelsubtypeID.
[31] MOVES 2010 Errata/Information Sheet, May 2010, EPA-420-B-10-026. Available at http://www.epa.gov/otaq/models/moves/index.htm

Third, analysis and processing of other MOVES inputs, including vehicle populations, I/M program data, fuel characterization, and meteorological data should be appropriately developed. Fourth, execution of MOVES with appropriate model options is required. Fifth, care must be taken in processing of model output, whether using the MOVES GUI or external SQL queries, to ensure emissions summaries are representative. Finally, at all stages of the modeling process, effective QC and QA procedures should be developed to minimize errors and ensure accurate production of model outputs.

> *What is the best approach for creating an emissions inventory with MOVES based on activity data obtained from the Highway Performance Monitoring System (HPMS)?*

The best approach for developing a MOVES emission inventory based on HPMS activity data is fairly similar to that described above for a TDM activity basis. It differs in that the HPMS VMT data requires a different type of pre-processing prior to integration with EPA's Converter Tools; all other modeling steps are the same. Effective pre-processing of HPMS data involves application of default MOVES source type distribution to the data, as described above in Section 2.1. It also necessitates the development of road type distributions outside of the EPA Converter Tools.

> *What will be the likely impacts on an emission inventory from developing and implementing different drive cycles derived from real world testing data within MOVES?*

In this study, the effects of implementing drive cycles in MOVES, based on real world testing data collected in Kansas City, varied considerably depending on which county was being modeled. The drive cycles we developed were based on largely urban driving data as discussed in Section 4, and applying these drive cycles to more rural counties in the Houston-Galveston area (as exhibited by fewer VMT driven in those counties) produced emissions that may not be representative. On the other hand, emissions from Harris County, the most urban of the counties modeled (with the corresponding largest daily VMT modeled), did not vary much at all. We believe that any alternate drive cycles developed and used in MOVES should closely match the type of driving activity for the area being modeled for emissions to be considered representative.

Apart from the impact of altering the drive cycles themselves, it is worthwhile to note that the effort required to produce representative drive cycle based on real world driving data is substantial, and may be prohibitive for some users. For this reason, we anticipate that most users will elect to use the default drive cycles provided as part of MOVES for their analyses.

What will be the likely impacts on an emission inventory after conversion from a MOBILE6 basis to a MOVES basis?

As presented in Section 5, this study shows that users can expect a considerable increase in modeled NOx emissions when transitioning from MOBILE6 to MOVES. This is consistent with other known comparisons of modeled MOBILE6 and MOVES results.[32][33] Our study also shows smaller increases of both CO and VOC when transitioning from MOBILE6 to MOVES. Other studies that modeled these pollutants have had mixed results on this point (TTI modeled decreased CO and increase VOC, while EPA has found decreased VOC results). Beyond the expected changes in emissions, users should be aware that the level of effort and computing time required to successfully execute and process a MOVES model run relative to MOBILE6 has increased substantially.

---

[32] TTI's *Production of MOVES On-Road Mobile, Link-Based Emissions Estimates and Document Preparation Technical Note*, submitted to TCEQ in July 2010.
[33] *Update on EPA's Motor Vehicle Emissions Simulator MOVES2010*, presented at CRC by J. Koupal in March 2010