

TECHNICAL REPORT STANDARD PAGE

1. Report No. 433		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle Analysis of Louisiana Vehicular Input Data for MOBILE 6		5. Report Date June 2008			
		6. Performing Organization Code			
7. Author(s) Chester G. Wilmot Srinivas Varanasi Srividya Vadlamani		8. Performing Organization Report No. 433			
9. Performing Organization Name and Address Department of Civil and Environmental Engineering Louisiana State University Baton Rouge, LA 70803		10. Work Unit No.			
		11. Contract or Grant No. LTRC Project No. 03-2SS, State Project No. 736-99-1108			
12. Sponsoring Agency Name and Address Louisiana Transportation Research Center 4101 Gourrier Avenue Baton Rouge, LA 70808		13. Type of Report and Period Covered Final Report, 2003 – 2007			
		14. Sponsoring Agency Code			
15. Supplementary Notes					
16. Abstract The purpose of this study was to identify sources of data for MOBILE 6 and set procedures to prepare the data in the format required for use in MOBILE 6. The Environmental Protection Agency (EPA) has provided a comprehensive set of documents describing proper use of MOBILE 6 and has also provided a full set of default values to assist in operation of the model. However, the EPA warns that local data should be used wherever possible, and efforts were made in this study to develop procedures that allow local data to be gathered for use in MOBILE 6. A proof-of-concept study was conducted on the use of Global Positioning System (GPS) instruments to estimate vehicle miles traveled (VMT) and soak times. From data of GPS-equipped vehicles in Lexington, Kentucky, gathering information on VMT and soak times in this manner appears to be feasible, and the results showed that local data is sometimes considerably different from that portrayed in MOBILE 6 default tables. Another method of local data collection tested in this study was a method used to estimate VMT from Highway Performance Monitoring System (HPMS) data without using volume groups. Volume groups are difficult to assign to links where no traffic count data exist. The method produces less accurate results than the current method (root mean square error of 1643 vehicles per day overall, versus 543 in the conventional method), but once established, it is easier to apply.					
17. Key Words Emissions, mobile, data.			18. Distribution Statement Unrestricted. This document is available through the National Technical Information Service, Springfield, VA 21161.		
19. Security Classif. (of this report) Unclassified		20. Security Classif. (of this page) Unclassified		21. No. of Pages 67	
				22. Price	

Analysis of Louisiana Vehicular Input Data for MOBILE 6

By

Chester G. Wilmot
Srinivas Varanasi
Srividya Vadlamani

Louisiana Transportation Research Center and
Department of Civil and Environmental Engineering
Louisiana State University
Baton Rouge, Louisiana 70803-6405

LTRC Project No. 03-2SS
State Project No. 736-99-1108

conducted for

Louisiana Department of Transportation and Development
Louisiana Transportation Research Center

The contents of this report reflect the views of the author/principal investigator who is responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the views or policies of the Louisiana Department of Transportation and Development or the Louisiana Transportation Research Center. This report does not constitute a standard, specification, or regulation.

August 2008

ABSTRACT

The purpose of this study was to identify sources of data for MOBILE 6 and set procedures to prepare the data in the format required for use in MOBILE 6. The Environmental Protection Agency (EPA) has provided a comprehensive set of documents describing proper use of MOBILE 6 and has also provided a full set of default values to assist in operation of the model. However, the EPA warns that local data should be used wherever possible, and efforts were made in this study to develop procedures that allow local data to be gathered for use in MOBILE 6.

A proof-of-concept study was conducted on the use of Global Positioning System (GPS) instruments to estimate vehicle miles traveled (VMT) and soak times. From data of GPS-equipped vehicles in Lexington, Kentucky, gathering information on VMT and soak times in this manner appears to be feasible, and the results showed that local data is sometimes considerably different from that portrayed in MOBILE 6 default tables.

Another method of local data collection tested in this study was a method used to estimate VMT from Highway Performance Monitoring System (HPMS) data without using volume groups. Volume groups are difficult to assign to links where no traffic count data exists. The method produces less accurate results than the current method (root mean square error of 1643 vehicles per day overall, versus 543 in the conventional method), but once established, it is easier to apply.

IMPLEMENTATION STATEMENT

The findings of this research will assist Metropolitan Planning Organizations (MPOs) in Louisiana in conducting air quality conformity analyses as required by the Clean Air Act and its amendments. The state Department of Transportation and Development (DOTD) will also find the input data for MOBILE 6 useful in preparing their State Implementation Plan (SIP) each year. MOBILE 6 can be linked to TransCAD, the transportation planning package used by each MPO in Louisiana as well as the DOTD, to provide graphic output that is not available from running MOBILE 6 on its own.

The input data to MOBILE 6 that is the most demanding to collect is the vehicle miles traveled (VMT) by vehicle class, hour of the day, and facility type. The common method of obtaining this data is to either use Highway Performance Monitoring Study (HPMS) data with links classified into volume groups, or to use transportation planning model link volume estimates. Each of these methods have their shortcomings. This report describes an alternative method of estimating VMT that could be used to obtain local values. Similarly, while it is common for users of MOBILE 6 to adopt the default values for vehicle soak times, this study demonstrates a method using GPS in vehicles to estimate soak times locally. The report documents the soak times observed in Lexington, Kentucky, using the method described in this report, and the values obtained there were considerably different to the default values. The reason for the difference can be traced back to the difference in the site where the default values were obtained (Spokane, WA, and Baltimore, MD) from and Lexington, Kentucky. It is doubtful whether cities in Louisiana are similar to those used to develop the default values either, so establishing local values is important. State and local engineers can use the methods described in this report to estimate VMT and vehicle soak times.

TABLE OF CONTENTS

ABSTRACT.....	V
IMPLEMENTATION STATEMENT	VII
TABLE OF CONTENTS	IX
LIST OF TABLES	XI
LIST OF FIGURES	XIII
INTRODUCTION.....	1
OBJECTIVES	3
SCOPE	5
METHODOLOGY	7
TASK 1: LITERATURE REVIEW	7
TASK 2: IDENTIFICATION OF DATA INPUT REQUIREMENTS	7
TASK 3: IDENTIFICATION OF SOURCES OF DATA	7
TASK 4: PROCEDURES TO PREPARE THE DATA IN THE FORMAT REQUIRED BY MOBILE 6....	8
DISCUSSION OF RESULTS	9
INPUT REQUIREMENTS FOR MOBILE 6	9
EXTERNAL CONDITIONS.....	9
Calendar Year	9
Month of Evaluation	9
Altitude	9
Temperature	10
Humidity	10
Other Air Conditioning Related Input Options.....	10
VEHICLE FLEET CHARACTERISTICS	11
Distribution of Vehicle Registrations	12
Diesel Fractions	13
Annual Mileage Accumulation Rates	13
Natural Gas Vehicle Fraction.....	13
VEHICLE ACTIVITY.....	14
Fraction of vehicle miles traveled (VMT) by vehicle class.....	14
VMT by Facility	14
VMT by hour	14
Speed VMT	14
Average Speed	15
Starts Per Day	15
Distribution of Vehicle Starts During the Day	16
Soak Distribution	16

Hot Soak Activity	17
Diurnal Soak Activity	17
Weekday and Weekend Trip Length Distribution	18
VEHICLE GASOLINE SPECIFICATIONS	18
Effects of Stage 2 Refueling	18
Sulfur Content of Gasoline	19
Oxygenated fuel programs	20
Gasoline Volatility	20
REVIEW OF PAST STUDIES	20
VEHICLE FLEET CHARACTERISTICS	21
Update of Fleet Characteristics for MOBILE 6	21
EPA's Technical Report M6.FLT.007	21
Locality Specific Fleet Characteristics: Current Practice, Problems, and Alternatives..	22
Modeling Vehicle Age Distribution	23
Effect of County Level Income on Vehicle Age Distribution	24
Vehicle Activity Data	25
SOAK DISTRIBUTION INPUTS TO MOBILE SOURCE EMISSIONS MODELING	28
Guidance for the Development of Facility Type VMT and Speed Distributions	29
Implementing Soak-time Distribution Models in a GIS Framework	32
COMPARISON BETWEEN MOBILE 6 AND MOBILE 5	34
USING GLOBAL POSITIONING SYSTEM (GPS) TO ESTIMATE VMT	36
Background to Investigation of Using GPS to Estimate VMT	36
Data	37
VMT by Facility Class	38
VMT by hour of the day	45
VMT by Speed	48
Conclusions Regarding the Use of GPS to Estimate VMT	55
AN ALTERNATIVE APPROACH TO ESTIMATING VMT	56
CONCLUSIONS	59
REFERENCES	61

LIST OF TABLES

Table 1: Mobile 6 vehicle classifications.....	11
Table 2: Composite vehicle types	12
Table 3: Comparison of MOBILE 5 and MOBILE 6.....	35
Table 4: VMT breakdown by functional classes in Lexington.....	41
Table 5: VMT breakdown by functional classes for MOBILE 6	42
Table 6: VMT by hour of the day input for MOBILE 6.....	46
Table 7: VMT distribution by speed for freeways (GPS Values).....	49
Table 8: VMT distributions by speed for freeways (default values)	50
Table 9: VMT distribution by speed for arterials (GPS values)	52
Table 10: VMT distribution by speed for arterials (default values)	53

LIST OF FIGURES

Figure 1: Lexington network divided into four functional classes	39
Figure 2: Functional classification of the GPS travel in Lexington area	41
Figure 3: Distribution of VMT by freeways for MOBILE 6	43
Figure 4: Distribution of VMT by arterials for MOBILE 6	43
Figure 5: Distribution of VMT by local streets for MOBILE 6	44
Figure 6: Distribution of VMT by ramps for MOBILE 6	44
Figure 7: VMT distribution by hour of the day for MOBILE 6	47
Figure 8: Graphical interpretation of speed distribution results for freeways	51
Figure 9: Graphical interpretation of speed distribution results for arterials	54

INTRODUCTION

The Clean Air Act amendments (CAAA) of 1990 require that non attainment and air quality maintenance areas regularly conduct regional emissions analyses to demonstrate conformity to emissions budgets established for these areas. In Louisiana, Baton Rouge and Lake Charles are ozone non attainment areas, while New Orleans is an air quality maintenance area. Eleven parishes in the state are currently non compliant with respect to air quality, and the number is respected to increase over time. Thus, emissions analysis is expected to become a more important issue in the future.

The Environmental Protection Agency (EPA) requires that its MOBILE models be used to estimate vehicle emissions in all states except California, where the use of emission factor (EMFAC) is permitted. The most recent version of the EPA's vehicle emissions model, MOBILE 6, was released in January of 2002. Following the release of MOBILE 6, a grace period of two years in which MOBILE 5 could be used was permitted, but after that the use of MOBILE 6 became mandatory.

Moving from MOBILE 5 to MOBILE 6 involves providing more detailed input data. This places greater demands on an agency conducting an analysis, and it is in response to this need that the research reported in this document has been conducted. The research conducted in this study was aimed at identifying the most efficient and reliable way to obtain the input data for MOBILE 6. Formalizing the data collection and preparation process allows for the review and assessment of alternative procedures; selection of those procedures that are the most efficient and reliable; and establishment of a procedure that can be subsequently reviewed, assessed, and, if necessary, amended.

OBJECTIVES

The objectives of this study are to:

1. identify sources of data that can provide or assist in providing input to MOBILE 6,
2. develop procedures to prepare the input data for MOBILE 6, and
3. document the procedures.

SCOPE

The research conducted in this study has been directed specifically at the estimation of data input for metropolitan areas within Louisiana. Thus, data sources particular to the state and procedures applicable to metropolitan area sizes and the technical expertise existing in the different metropolitan areas have influenced the choice of procedures.

The developers of the EPA's MOBILE models emphasize that, while default parameter values are provided in the model, local values should be used wherever possible. In this study, considerable effort was dedicated to the estimation of local VMT in the form required by MOBILE 6 as well as the estimation of local starts and soak times.

METHODOLOGY

The study was conducted by executing the five tasks listed below. The results of the tasks are described in the “Discussion of Results” in the following section.

Task 1: Literature Review

The literature on the topic of the data input requirements of MOBILE 6 and the procedures used to prepare the data in the required format was reviewed. Particular attention was given to literature distributed by the Environmental Protection Agency (EPA) and the Travel Model Improvement Program (TMIP) and research published by agencies such as the Transportation Research Board (TRB) and California Air Resources Board (CARB). Other sources of literature accessed included the Transportation Research Information System (TRIS); Elsevier’s new search engine for scientific publications, Scirus; and papers presented at the Transportation Research Board’s annual meetings.

Task 2: Identification of Data Input Requirements

Information from the literature review was used to identify the input data requirements of MOBILE 6. Much of the information was obtained from the comprehensive list of publications issued by the EPA. This task included identification of the data items needed and the format in which the data is required as input to MOBILE 6.

Task 3: Identification of Sources of Data

Records of data sources, such as the “Directory of Transportation Data Sources,” published by the Bureau of Transportation Statistics (BTS, 2000) and other sources (Wilmot and Stopher, 2002), were consulted in an effort to identify the most appropriate source of data for each input item. Data sources used by agencies conducting emissions analysis in the past were also reviewed and assessed. Past practice in establishing vehicle miles traveled (VMT) by vehicle type, time of day, and the sources of data used to establish these values was noted.

Task 4: Procedures to Prepare the Data in the Format Required by MOBILE 6

Procedures identified in the literature review in Task 1 were reviewed for application in Louisiana. New procedures were also considered. One such new procedure involved assessing the feasibility of using Global Positioning System (GPS) instruments in a sample of vehicles to estimate VMT, speed, starts, soak times, and trip length distributions of vehicular travel in an area. Another new procedure involved using a detailed classification of roadway segments and observed traffic volumes from the Highway Performance Monitoring System (HPMS) to estimate VMT.

DISCUSSION OF RESULTS

Input Requirements for MOBILE 6

The input requirements for MOBILE 6 can be broadly classified into four categories (EPA, 2002a):

1. external conditions,
2. vehicle fleet characteristics,
3. vehicle activity, and
4. vehicle gasoline specifications.

The input required in each classification is described below. A detailed description of each input item is described in the Environmental Protection Agency's requirements (EPA, 2002b) and is summarized below.

External Conditions

The following are the external conditions of the vehicle fleet and transportation network that influence vehicle emissions. The conditions that are permitted to influence MOBILE 6's estimate of vehicle emissions are described in detail.

Calendar Year

MOBILE 6 is set up to model emission factors for the calendar years 1952 to 2050. The calendar year for which emission factors are to be calculated is given to MOBILE 6 by the analyst. A four digit value of the calendar year of evaluation must be entered.

Month of Evaluation

Using the information provided, emission factors for January 1 or July 1 of the calendar year of evaluation can be calculated. Users should choose the month of evaluation with respect to the pollution concerned. For summer day emissions of VOC, July is appropriate, while January is appropriate for the winter CO emissions. By default, MOBILE 6 calculates for January.

Altitude

This input data permits the estimation of the effect of altitude on vehicle emissions. A low

altitude region is any region up to 500 feet above mean sea level.

Temperature

Temperature is a mandatory input for MOBILE 6, and users must supply either minimum and maximum temperatures or the hourly temperature during a typical day. Hourly temperature is not required unless the user is trying to model a specific exceedance episode for air quality analysis.

Daily Temperature Range MOBILE 6 requires minimum and maximum daily temperatures in several calculations. The allowable minimum temperature range is 0° F to 100° F. The allowable maximum temperature range is 10° F to 120° F.

Ambient Temperature for Each Hour of the Day This command is required if the Min/Max temperature command is not selected. It allows for the entering of temperature values for each hour of the day.

Humidity

Humidity is measured as absolute humidity or humidity ratio. Humidity ratios may be available as output from a meteorological model. Humidity should be expressed in grains of water per pound of dry air when input to MOBILE 6, and the units of the meteorological model should be converted accordingly. For users without access to humidity ratios for the exceedance days being modeled, the humidity ratios can be calculated using temperature, relative humidity, and barometric pressure. MOBILE 6 allows users to enter values between 20.0 and 528.0 grains of water per pound of dry air. The EPA suggests using the lowest humidity ratio value of the day to prevent unreasonable relative humidity values.

Other Air Conditioning Related Input Options

EPA specifies three factors that affect air conditioning usage:

1. average percent cloud cover,
2. period of peak sun, and
3. sunrise/sunset times.

Vehicle Fleet Characteristics

MOBILE 6 uses a new vehicle classification system that allows more accurate emission estimates to be made. Fleet characteristic commands require input on vehicle age, power source, and level of activity. The detailed vehicle classification results in 32 categories of vehicles, as shown in Table 1 (EPA, 2002b). If no distinction is made between gasoline and diesel powered vehicles, the 16-category classification shown in Table 2 is used (EPA, 2002b).

Table 1
Mobile 6 Vehicle Classifications

Number	Abbreviation	Description
1	LDGV	Light-Duty Gasoline Vehicles (Passenger Cars)
2	LDGT1	Light-Duty Gasoline Trucks 1 (0-6,000 lbs. GVWR, 0-3750 lbs. LVW)
3	LDGT2	Light-Duty Gasoline Trucks 2 (0-6,000 lbs. GVWR, 3751-5750 lbs. LVW)
4	LDGT3	Light-Duty Gasoline Trucks 3 (6,001-8,500 lbs. GVWR, 0-5750 lbs. ALVW)
5	LDGT4	Light-Duty Gasoline Trucks 4 (6,001-8,500 lbs. GVWR, 5751 lbs. and greater ALVW)
6	HDGV2B	Class 2b Heavy-Duty Gasoline Vehicles (8501-10,000 lbs. GVWR)
7	HDGV3	Class 3 Heavy-Duty Gasoline Vehicles (10,001-14,000 lbs. GVWR)
8	HDGV4	Class 4 Heavy-Duty Gasoline Vehicles (14,001-16,000 lbs. GVWR)
9	HDGV5	Class 5 Heavy-Duty Gasoline Vehicles (16,001-19,500 lbs. GVWR)
10	HDGV6	Class 6 Heavy-Duty Gasoline Vehicles (19,501-26,000 lbs. GVWR)
11	HDGV7	Class 7 Heavy-Duty Gasoline Vehicles (26,001-33,000 lbs. GVWR)
12	HDGV8A	Class 8a Heavy-Duty Gasoline Vehicles (33,001-60,000 lbs. GVWR)
13	HDGV8B	Class 8b Heavy-Duty Gasoline Vehicles (>60,000 lbs. GVWR)
14	LDDV	Light-Duty Diesel Vehicles (Passenger Cars)
15	LDDT12	Light-Duty Diesel Trucks 1 and 2 (0-6,000 lbs. GVWR)
16	HDDV2B	Class 2b Heavy-Duty Diesel Vehicles (8501-10,000 lbs. GVWR)
17	HDDV3	Class 3 Heavy-Duty Diesel Vehicles (10,001-14,000 lbs. GVWR)
18	HDDV4	Class 4 Heavy-Duty Diesel Vehicles (14,001-16,000 lbs. GVWR)
19	HDDV5	Class 5 Heavy-Duty Diesel Vehicles (16,001-19,500 lbs. GVWR)
20	HDDV6	Class 6 Heavy-Duty Diesel Vehicles (19,501-26,000 lbs. GVWR)
21	HDDV7	Class 7 Heavy-Duty Diesel Vehicles (26,001-33,000 lbs. GVWR)
22	HDDV8A	Class 8a Heavy-Duty Diesel Vehicles (33,001-60,000 lbs. GVWR)
23	HDDV8B	Class 8b Heavy-Duty Diesel Vehicles (>60,000 lbs. GVWR)
24	MC	Motorcycles (Gasoline)
25	HDGB	Gasoline Buses (School, Transit and Urban)
26	HDDBT	Diesel Transit and Urban Buses
27	HDDBS	Diesel School Buses
28	LDDT34	Light-Duty Diesel Trucks 3 and 4 (6,001-8,500 lbs. GVWR)

Table 2
Composite Vehicle Types

Number	Abbreviation	Description
1	LDV	Light-Duty Vehicles (Passenger Cars)
2	LDT1	Light-Duty Trucks 1 (0-6,000 lbs. GVWR, 0-3,750 lbs. LVW)
3	LDT2	Light-Duty Trucks 2 (0-6,000 lbs. GVWR, 3,751-5,750 lbs. LVW)
4	LDT3	Light-Duty Trucks 3 (6,001-8,500 lbs. GVWR, 0-5,750 lbs. ALVW*)
5	LDT4	Light-Duty Trucks 4 (6,001-8,500 lbs. GVWR, 5,751 lbs. and greater ALVW)
6	HDV2B	Class 2b Heavy-Duty Vehicles (8,501-10,000 lbs. GVWR)
7	HDV3	Class 3 Heavy-Duty Vehicles (10,001-14,000 lbs. GVWR)
8	HDV4	Class 4 Heavy-Duty Vehicles (14,001-16,000 lbs. GVWR)
9	HDV5	Class 5 Heavy-Duty Vehicles (16,001-19,500 lbs. GVWR)
10	HDV6	Class 6 Heavy-Duty Vehicles (19,501-26,000 lbs. GVWR)
11	HDV7	Class 7 Heavy-Duty Vehicles (26,001-33,000 lbs. GVWR)
12	HDV8A	Class 8a Heavy-Duty Vehicles (33,001-60,000 lbs. GVWR)
13	HDV8B	Class 8b Heavy-Duty Vehicles (>60,000 lbs. GVWR)
14	HDBS	School Buses
15	HDBT	Transit and Urban Buses
16	MC	Motorcycles (All)

Distribution of Vehicle Registrations

This option allows for providing vehicle registration distributions by vehicle age for any of the 16 composite (combined gas and diesel) vehicle types (Table 2). Vehicle registration data for each of the 25 vehicle ages (from vehicle age 1 through vehicle age 25) for all the 16 composite vehicle types can be specified. This option requires that an external data file containing the age distributions be provided. Below is an example of an external data file for this option.

REG DIST

*LDV

1 0.57 0.99 0.33

*LDT1

2 0.57 0.97 0.333.....

The first line features the command name. The next line contains the composite vehicle type, followed by 25 age fractions representing the fraction of vehicles of that age in that composite class in July. The composite vehicle type is represented by an integer between 1 and 16. The vehicle age fractions are represented by fractions between 0 and 1 for each of the 25 model years. The sum of the 25 fractions should be equal to 1.

Diesel Fractions

Command name: DIESEL FRACTIONS

This command allows users to supply locality specific diesel fractions for 14 of the 16 composite vehicle categorized by vehicle age. All urban/transit buses are assumed to be diesel fueled, and all motorcycles are assumed to be gasoline fueled, so those categories do not require a diesel fraction. That leaves 14 composite vehicle categories for which diesel fractions can be specified. The diesel fraction represents the percent of diesels in a composite vehicle category for any vehicle age. Separate input for each calendar year to be modeled must be provided.

Diesel fractions categorized by the age of the vehicles and for each of the 14 vehicle types must be provided with this command. With 14 vehicle categories and 25 vehicle ages, 350 separate diesel fractions must be provided.

Annual Mileage Accumulation Rates

Command name: MILE ACCUM RATE

This command allows the annual mileage accumulation rates by vehicle age for any or all of the 28 vehicle types to be provided. The mileage accumulation rate represents the total annual travel accumulated per vehicle of a given age and individual vehicle category.

The format of this command is as follows:

1. After the command name, the next line should contain the vehicle type and mileage accumulation data. The vehicle type is represented by the number associated with each of the 28 vehicle types (classified in Table 1). This value should be followed by the mileage accumulation rates for each of the 25 ages, starting with the youngest.
2. The values are input as miles divided by 100,000 (e.g., 12,000 miles is input as 0.12).

Natural Gas Vehicle Fraction

Command name: NGV FRACTION

This command indicates the percentage of vehicles that operate on compressed or liquefied natural gas. By default, the fraction of NGVs is zero. This command requires an external file containing the NGV fractions for each vehicle type and model year from 1994 through 2050. In the external file, the command name should be followed by 28 blocks of NGV fractions, with each block containing 57 values of NGV fleet penetration percentages. The values can range from 0.0 to 100.0.

Vehicle activity

Vehicle activity relates to input requirements regarding vehicle miles traveled distributions (VMT), vehicle engine start distributions, vehicle trip length distributions, soak length distributions etc.

Fraction of vehicle miles traveled (VMT) by vehicle class

Command name: VMT FRACTIONS

The VMT mix is the fraction of total highway VMT accumulated by each of the 16 composite vehicle types (Table 2). The input must consist of 16 fractional values, which are the fraction of total highway VMT accumulated by each of the 16 composite vehicles. All values must be between 0 and 1, and the sum of all the values must add up to 16

VMT by Facility

Command name: VMT BY FACILITY

The VMT should be allocated to various roadway or facility types (4 types) by each of the 28 vehicle classes for each of the 24 hours in the day. An external file containing the command name followed by the vehicle class number (1 to 28) and the 96 fractions representing the fraction of travel on each roadway type at each hour of the day for that specific vehicle class is needed. For a given vehicle class, the VMT fractions must be entered as 24 values, starting from 6 a.m. for every hour of the day. There should be 4 times 24 value sets corresponding to the 4 facility types in the order: freeway, arterial, local, and ramp.

VMT by hour

Command name: VMT BY HOUR

The fraction of total VMT that occurs each hour of the day can be allocated by using this command. This command requires values that are independent of facility type. The external file that we supply should contain 24 VMT fractions, one value for each of the 24 hours in the day.

Speed VMT

Command name: SPEED VMT

This command allocates VMT by average speed (preselected ranges) on freeways and arterial roads for each of the 24 hours in the day. This command requires an external data file, which contains the following information:

1. command name;

2. an integer for facility type, where 1= freeway, 2= arterial;
3. An integer for the hour of the day; and
4. fourteen numbers representing VMT fractions in each of the 14 average speed categories. The 14 average speed fractions must add up to 1. The first of the 14 speeds is 'idle,' and the next 13 range from 5 mph to 65 mph in 5 mph increments.

Average Speed

Command name: AVERAGE SPEED

By using this command, users can designate a single average speed to use for different roadways for the entire day. Using this command will override the default and the user-supplied distribution of VMT by facility type. The average speed command includes up to six data elements. The first two data elements are always required. The elements are:

1. average speed (can range from 2.5 to 65 mph),
2. roadway scenario, and
3. distribution of VMT by facility type (4 values).

The roadway scenario indicates the type of driving that the user intends for the average speed input to model, e.g.,

AVERAGE SPEED: 25.8 Arterial

This example sets the average speed of arterials at all hours to 25.8 mph and assigns 100 percent of VMT to the arterial type.

AVERAGE SPEED: 48.9 freeway 90.0 0.0 0.0 8.6 (order is freeway, arterial, local, freeway ramp)

This example sets the combined average speed of freeways and freeway ramps to 48.9 mph. The four values following the indication of a roadway scenario are the distribution of VMT by facility used. They show 90 percent of VMT on freeways and 8.6 percent of VMT on freeway ramps.

Starts Per Day

Command name: STARTS PER DAY

By using this command, the average number of engine starts (trips) per vehicle per day for specific vehicle classes for all days can be specified. Emission estimates for heavy duty diesel fueled vehicles and buses are not affected by the number of starts per day in MOBILE 6.

The external data file should contain the following:

1. first value which indicates whether the engine starts per day values apply to the weekdays (1) or the weekend (2),
2. next value, which indicates the vehicle class, and
3. the next 25 values which indicate the average number of vehicle starts per day by vehicle age.

E.g., STARTS PER DAY,

1 1 10.9 10.98 7.925 values

1 2 10.9 10.98 7.925 values

Distribution of Vehicle Starts During the Day

Command name: START DIST

By using this command, engine starts by hour of the day can be allocated. 48 values must be supplied; each value represents the average fraction of all engine starts that occur in each hour of a 24 hour day, for both weekdays and weekends. The sum of all 24 values must be equal to one.

The external data file should contain the following information for this command:

1. The first 24 values are the distribution for the weekdays, and
2. the next 24 values are the distribution for the weekend.

E.g., START DIST,

0.08 0.09 0.07

Soak Distribution

Command name: SOAK DISTRIBUTION

The soak time affects exhaust start and exhaust running emissions. A vehicle is considered to be soaking if its engine is not running. Soak time is the length of time between when an engine is turned off and the next time when it is started. We use soak time in MOBILE 6 to determine, for each hour of the day, the percentage of vehicles that have been soaking for a given amount of time prior to an engine start. There will be 70 soak durations (explained later) for each of the 24 hours in the day.

The external data file for this command should contain the following information:

1. The external data file must first index the 24 hours and then the 70 soak times.

2. The first 24 values are the soak time distribution values for only the first soak time interval.
3. The second set of 24 values is for the second soak time interval, and the sets continue until all 70 sets of 24-soak time intervals have been included.

The soak time intervals are (EPA, 2002b):

<u>Interval number</u>	<u>Interval range (N=interval number)</u>
1	(Greater than 0.1) to 1.0 minutes
2 to 30	(greater than N-1) to N minutes
31 to 45	(greater than 2N-32) to (2N-30) minutes
46 to 67	(greater than 30N-1320) to (30N-1290) minutes
68	greater than 720 minutes
69	(greater than 0) to 0.01 minutes (restarts)
70	zero minutes (stalls, not used)

Hot Soak Activity

Command name: HOT SOAK ACTIVITY

This command allows the user to specify a hot soak duration distribution for each of the 14 daily time periods. When fuel vapors escape from a hot vehicle that has just been turned off, hot soak emissions occur. MOBILE 6 assumes that hot soak durations range from a minimum of 1 minute to a maximum of 60 minutes. The hot soak time distributions represent the number of vehicles experiencing a hot soak of a given duration (1-60 minutes) at each hour of the day. MOBILE 6 divides the day into 14 time periods - one for each hour between 6 a.m. and 7 p.m., plus one for the hours from 7 p.m. through 5 a.m. the next day.

The external data file should contain the following information: After the command name, the hot soak activity values must be arranged in columns of 60 fractions, with one column for each time period, starting with 6 a.m. The 60 rows represent the time periods from 1 to 60 minutes. When using this command, 840 values representing the fraction of vehicles experiencing a hot soak of each duration (1 to 60 minutes) at each time period of the day are to be entered.

Diurnal Soak Activity

Command name: DIURN SOAK ACTIVITY

By using this command, we can specify a diurnal soak time distribution for each of the 18 daily time periods. MOBILE 6 assumes that diurnal soak times range from a minimum of 1

hour to a maximum of 72 hours. It is assumed that from 11 p.m. to 6 a.m., there will be no diurnal emissions. When choosing this command, 1296 values representing the fraction of vehicles experiencing a diurnal of each duration for their time periods in the day must be entered. So the external data file should contain 18 columns (time periods) and 72 rows (fraction of vehicles experiencing diurnal) in each time period.

Weekday and Weekend Trip Length Distribution

Weekday Trip Length Distribution Command name: WE DA TRI LEN DI.

Using this command, the fraction of weekday VMT that occurs during trips of various durations at each hour of the day can be specified. If this command is used, percentages must be entered for 6 trip length ranges and 14 hourly groups, resulting in 84 values. The trip length ranges are: 10 minutes or less; 11 to 20 minutes; 21 to 30 minutes; 31 to 40 minutes; 41 to 50 minutes; and 51 or more minutes. The percentage of VMT for each of these 6 ranges is required for each hour of the day from 6 a.m. to 7 p.m. plus an average distribution for the 11-hour block from 7 p.m. to 6 a.m. is also required.

The external data file should have the following format:

1. the first line is command name;
2. the second line is the percentage of VMT for the first hour (6 a.m. to 7 a.m.) for each of the 6 trip length ranges in the increasing order of trip length;
3. the next 12 lines provide the corresponding information from hours 2 through 13; and
4. the last line provides the information for the 11-hour block of time.

Weekend Trip Length Distribution Command name: WE EN TRI LEN DI

Using this command, one can specify the fraction of weekend VMT. This requires the same kind of data as that of the weekday trip length distribution; the only difference is that it is weekend trip length distribution data.

Vehicle Gasoline Specifications

Effects of Stage 2 Refueling

Command name: STAGE 2 REFUELING

By using this command, one can model the impact of a stage 2 (at the pump) vapor recovery system requirement on refueling emissions (also known as stage 2 emissions) from gasoline-fueled vehicles. If this command is not specified, MOBILE 6 will not calculate the impact

of a stage 2 program. Stage 2 systems reduce HC emissions by reducing the amount of gasoline vapors that escape into the atmosphere during refueling and fuel spillage. The reduction is high if the vehicle has an onboard recovery system, and it is low with uncontrolled emissions. The uncontrolled emissions are calculated from inputs such as RVP, fuel economy, and various fuel temperature parameters.

This command requires the following data entry (EPA, 2002b):

1. The first line is the command name.
2. The second line should contain the following four items:
 - a. The first number should be the last two digits of the calendar year in which stage 2 programs began or will begin. The eligible calendar years are 1989 through 2050.
 - b. The second number should be a one digit integer for the number of phase-in years of the program. This value should range from one through nine.
 - c. The third number is a three digit integer for the percent efficiency for the LDGVs (Table 1) and LDGTs (Table 1) in the program. This value can range from 0 through 100.
 - d. The fourth number is a three digit integer for the percent efficiency for the HDGVs (Table 2) in the program (ranges from 0 through 100).

Sulfur Content of Gasoline

Command name: SULFUR CONTENT

By using this command, one can enter the sulfur content of gasoline for calendar years through the year 1999. For calendar years after the year 1999, MOBILE 6 will ignore this input. MOBILE 6 accepts sulfur contents for 30 ppm through 600 ppm.

E.g., SULFUR CONTENT: 50.0

For years after the year 2000, the “FUEL PROGRAM” command should be used. The user should place integer 4 on the command line after the fuel program command. On the following lines, the user must enter 32 values, indicating average and maximum sulfur contents. The first 16 values are average fuel sulfur contents for gasoline in calendar years 2000 through 2015. The next 16 values are the maximum sulfur levels of these years. Values from 30 to 1000 ppm are valid input.

E.g.,

FUEL PROGRAM: 4

333 334 242 356.....

Oxygenated fuel programs

Command Name: OXYGENATED FUELS

By using this command, one can model the effects of oxygenated gasoline on exhaust emissions for all gasoline fueled vehicle types. This command requires that the following five numbers be provided by the user:

1. Ether blend market share. The range is 0.000 to 1.000.
2. Alcohol blend market share. The range is 0.000 to 1.000.
3. Average oxygen content of ether blend fuels (percent weight expressed as a decimal fraction). The range is 0.000 to 0.027.
4. Average oxygen content of alcohol blend fuels. The range is 0.000 to 0.035.
5. A number that specifies whether a Reid Vapor Pressure (RVP) has been granted to allow “splash blending” of alcohol based oxygenates. If a waiver has not been granted, then value=1; otherwise, value=2.

Gasoline Volatility

The RVP (Reid Vapor Pressure) of the dispensed fuel indicates the volatility of gasoline. This command is required. There is no MOBILE 6 default. The RVP value entered must reflect the average in-use RVP of gasoline in the region of the country being modeled. The RVP value ranges from 6.5 psi to 15.2 psi.

Review of Past Studies

Several past studies by various authors and EPA’s technical documents regarding the input requirements for MOBILE 6 were reviewed, and the most pertinent information is presented in this section.

Data regarding external conditions and the vehicle gasoline specifications can be obtained from local agencies, and little further analysis is needed. The most tedious and time consuming part of data input acquisition is getting vehicle fleet characteristics and vehicle activity data. Most of the literature deals with acquisition of the complicated input.

Vehicle Fleet Characteristics

Modeling vehicle emissions requires input data characterizing the fleet of the vehicles in use in the study area. Economic and demographic changes cause significant changes in the vehicle fleet characteristics such as vehicle age distribution, vehicle type distribution, and mileage accumulation rates, which makes updating the fleet characteristic important.

Update of Fleet Characteristics for MOBILE 6

In 1998, Acurex Environmental Corporation, under contract to the EPA, used the latest available data to update national fleet characteristics for the MOBILE 5a model (Arcadis et al., 1998). They used vehicle registrations in 1996 to characterize the fleet. They suggested that the results can also be used in MOBILE 6. Their report presents a review of the sources of the vehicular data and states that the only centralized source of nationwide vehicle registration data is from R.L. Polk & Company. This company compiles the Department of Motor Vehicle registration data from each state into its database on a quarterly basis. The data include the make, model year, gross vehicle weight, and fuel type. Along with this source, the company also reviewed:

1. Truck Inventory and Use Survey by the U.S Bureau of Census, which gives data regarding light duty and heavy duty trucks. The data includes vehicle age, gross vehicle weight, and fuel type, and, most importantly, it gives the miles driven in the calendar year 1992.
2. Nationwide Personal Transportation Survey – this data does not contain the data regarding fuel type and vehicular weight, but it gives data regarding light-duty vehicle type, age, and annual mileage.
3. Federal Transit Administration data – section 15 of the Urban Mass Transportation Act of 1970 imposed a uniform accounting system among transit agencies that receive federal aid, and, since the 1979 fiscal year, this data has been compiled into annual reports that are a valuable source of information. Arcadis and associates compiled bus inventory data from this source, including model year and vehicle miles traveled from this source (1998). They also performed several manipulations and adjustments to the data to get it into the form required by MOBILE 5.

EPA's Technical Report M6.FLT.007

The EPA explains a methodology to convert the July 1996 registration data, prepared by

Arcadis and associates, into registration data by age as well as methods to project the future vehicle fleets by vehicle category (EPA, 2001b). The EPA fits curves through the 1996 registration data for each vehicle class and fuel type category. Several types of curves, such as linear, polynomial, exponential, and Weibull were explored to find the best fit. The best fit analysis resulted in age distribution estimates for vehicle ages 1 through 25. To estimate the VMT mix, the MOBILE model requires the total vehicle population (count) by vehicle class for each calendar year. The EPA technical report M6.FLT.007 gives the following formula for determining the total vehicle count that can be determined for each vehicle category for future calendar years:

$$VC_x = (VC_{x-1} + Sales_x) \times (1 - SR_x)$$

where

VC = vehicle count

Sales = number of vehicles sold

SR = % of in-use fleet that is scrapped

x = the calendar year in question

Locality Specific Fleet Characteristics: Current Practice, Problems, and Alternatives

A recent study by Granell, Guensler, and Bachman (2001) describes the problems in the current practice of using the locality specific fleet distributions and proposes some alternatives. They reviewed certain EPA documents and found that the Emissions Inventory Improvement Program (EIIP) provides various procedures for estimating vehicle characteristic using two data sources:

1. Inspection and maintenance (I/M) programs: I/M programs require light duty, medium duty, and heavy duty vehicles to undergo annual or biannual I/M tests. The I/M test database can be helpful in registration distributions. These I/M programs collect information regarding vehicle class, model year, vehicle identification number (VIN), and odometer reading data.
2. License plate readings from video cameras used in large remote sensing programs: These video cameras record vehicle license plates. The license plate number should be matched to the registration database to obtain vehicle characteristics in fleet distribution development.

A third alternative data source, proposed by Granell and colleagues (2001), is the tax revenue data, which is commonly used among MOBILE modelers. They describe how the local data for the state of Georgia was developed using a sample from the 1999 Georgia Department of Revenue registration data. The vehicle identification numbers (VIN's) of all the vehicles registered in the state were obtained from the database. Software developed by the Eastern Research Group (version 2000.01) is capable of decoding VINs of vehicles sold in the US between model years 1972 and 2000. The results of VIN decoding gave important information, such as vehicle model, body, engine type, model year, etc. In applying the software, Granell and associates found the following problems, and proposed the following solutions, for VIN decoding:

1. Lack of uniformity of VIN numbers: Older vehicles and European vehicles gave problems because of a lack of uniformity of VIN numbers. A graphical display of year distributions is the suggested solution for this problem.
2. The VIN decoder weight information is inadequate for MOBILE 6, which requires detailed weight information. The registration database is a good solution for obtaining missing data.
3. VIN decoding sometimes led to inconsistencies in type of fuel and type of vehicle. Garner and colleagues suggested the inconsistencies could be avoided by using the registration distribution.

Modeling Vehicle Age Distribution

Qiao, Yu, Li, and Wang (2001) proposed models and corresponding software to forecast the age distribution of vehicles for specific localities. They developed two types of models:

1. Modeling Type 1 (MT1) models the number of vehicles of a particular vehicle type in each age group and uses the results to forecast the future age distribution.
2. Modeling Type 2 (MT2) models the future age distribution directly. Prediction of age distribution by these models includes additional information, such as socio-economic indices and the local distribution of the recent years.

The process of projecting age distribution includes the following steps:

1. calibrating parameters for each model,
2. examining of input indices,
3. choosing model types and structures, and
4. projecting of vehicle age for the target year

A FORTRAN program by the name of MOFAD (Modeling and Forecasting Age Distribution) implements the whole modeling process. MOBILE 6 needs vehicle age distributions with 25 ages and 16 vehicle types to be modeled and forecasted, making a total of 400 models to be built. For each model, the paper gives two model types, and for each model type, the authors tested five kinds of candidate models. One was a linear regression model, three were different non linear models, and the last was a time series model. Age distributions for the eight counties in Houston, Texas were used for the validation of model structures and parameters. Qiao and associates found that their model produces more reliable results than the use of default values.

Effect of County Level Income on Vehicle Age Distribution

Miller, Davis, Reed, and Doraiswamy (2001) proposed the effect of county level income on vehicle age distribution and emissions. Vehicle registration data were analyzed on a county by county basis in the process of performing a statewide emission inventory of on road mobile sources in Tennessee. The types and ages of vehicles were compared to national statistics and average personal income of county residents. Median vehicle age was observed to correlate strongly with the average personal income for each county. Counties with the lowest income were observed to have the oldest vehicle fleets, while the counties with the highest income had the newest vehicle fleets. The median vehicle age of the lower income counties was 10.8 years, while that of the higher income counties was only 5.9 years. Low income counties also showed 86 percent more light duty trucks per capita than the national average, and these trucks were five years older than the national average. The conclusion is that county level personal income is a good predictor of vehicle age and can be used as a readily obtainable criterion instead of using the local vehicle registration data.

An objective of this study was to investigate ways to improve the accuracy of emission inventories for mobile sources. County level vehicle age distributions were developed from the registration data provided by the Tennessee Department of Safety's Title and Registration Section. The data contained information on the county, registration class code, model year, and body type for vehicles of model year 2001 and earlier. County codes are 2-digit numbers, and the model year was given as the last two digits of the year in which the vehicle was manufactured. The registration class code is a four digit number that identifies passenger vehicles, heavy duty trucks, motorcycles, and trailers. The data set was filtered to remove all but passenger vehicles and light trucks. For the purpose of this paper, the data were grouped into four area sub groups. Vehicle age distributions were developed for two

major vehicle categories: light duty vehicles and light duty trucks. Using the vehicle counts, the fraction of vehicles in each age group was calculated and plotted. These fractions represent the registration distribution by age. A simple single measure of vehicle age is the median vehicle age, calculated as the 50th percentile of the cumulative age. Tennessee counties showed higher vehicles per capita than the national average.

From the U. S. Department of Commerce's Bureau of Economic Analysis (BEA) website (www.bea.doc.gov), data on county-level per capita personal income for Tennessee counties were obtained. The most recent data available were for 1999. Research found that there is a consistent pattern of newer vehicles in high income counties and older vehicles in low income counties. A significant finding of this research is that there is a strong relationship between personal income and median vehicle age at the county level. The correlation coefficient for median vehicle age versus personal income was 0.996 for LDVs and 0.979 for LDTs. Therefore, personal income may be a good surrogate for vehicle age. Analyzing vehicle registration data is not difficult, but it can be a time consuming process.

Vehicle Activity Data

This category of MOBILE input contains information regarding VMT distributions, soak length distributions, trip length distributions, and vehicle engine starts. The following are the studies reviewed in forming a background to understand the vehicle activity concept for obtaining MOBILE 6 input.

Alternative Methodology for Collecting Local Vehicle Start Information. Everett and Sacs (2001) proposed an alternative methodology for collecting local vehicle start information for use in MOBILE 6. They used a small, inexpensive (\$100) data logger developed by GeoStats, Inc. Installation of the device is very simple, and in order to reduce recruitment costs, use of the data logger was linked to the Knoxville MPO's planned household survey of 1,600 households. The device records the on/off engine events as well as the date and duration of each event. The duration of the survey was expected to be between seven and ten weeks. Only weekdays were eligible travel days in the household (HH) survey, so a maximum of 50 travel days was available. Using 1,600 households and 50 travel days, the potential pool for the "start study" was estimated at 32 households per day. Assuming 20 percent of the eligible households were willing to participate in the start study resulted in an estimate of six households per day. Assuming an average of two vehicles per household meant that 12 devices per day were needed. The devices needed to be installed 12

hours prior to the beginning of a household's travel day. This is due to the EPA's definition of engine soak distributions, which include a 720 minute category. Finally, individuals were instrumented for a minimum of two days. Sixty electronic data loggers (EDLs) were considered sufficient to serve the expected number of households.

Households were contacted by telephone to schedule a delivery date and time. Prior to scheduling a delivery, the scheduler verified which vehicles in the household were qualified and whether or not the participant's family members were willing to have an EDL installed. Generally, to qualify for the study, a vehicle had to be considered a part of the household's regularly used fleet. EDLs were delivered to the households between 5.30 p.m. and 8.30 p.m. Upon completion of the survey, participants were requested to unplug the EDLs themselves and leave the EDLs and info-cards in the bag outside in an agreed upon location for pick up. The entire process, including installation, required less than 15 minutes. When a driver collected an EDL with data, it was necessary to download the data into laptop computers before the EDL could be delivered to another household. The data were trimmed to eliminate non starts such as engine stalls or other non-trip related activity. Durations less than 30 seconds were eliminated as engine stalls. The last step before analysis was to place each valid installation vehicle into its proper EPA vehicle type classification. In total, the study collected data from 377 vehicles from 200 households.

The results of the survey were compared with the profiles of the default values proposed for MOBILE 6. The number of vehicle starts obtained using the EDLs were found to be considerably less than in the EPA default data. The total cost of gathering the data and performing the initial analysis was approximately \$36,000, or about \$85 per vehicle data set.

VMT Mix Modeling for MOBILE Source Emissions Forecasting Bhat and Nair (2001) present a fractional split model that predicts the VMT mix on links as a function of the functional roadway classification of the link, the physical attributes of the link, the operating conditions on the link, and the attributes of the traffic analysis zone in which the link lies.

The EPA requires the metropolitan planning areas, rated as serious or higher in non-attainment designation for ozone and CO, estimate their mobile source emissions using network based transportation models. This involves the estimation of volumes and speeds on each network link in the metropolitan area from travel demand models, followed by the

computation of link specific emissions based on:

1. link VMT,
2. vehicle speed on the link,
3. vehicle class-specific emission factors, and
4. VMT mix fractions.

Of these items, the link VMT and link speeds are obtained from the network based travel demand models. The vehicle class specific emission factors are obtained from the emission factor models. The VMT fraction by the vehicle class is a supplementary traffic related parameter that is to be provided by the analyst. Bhat and Nair (2001) proposed a multinomial logit model to estimate the fraction and used the following information to test the model in the Dallas/Fort Worth region in Texas:

1. vehicle classification counts conducted in Dallas area by TXDOT's Regional Planning organization;
2. the 1996 GIS-based road network file for the Dallas area; and
3. zonal-level land use characteristics on file for the Dallas area.

The count data obtained were 16 hour data, and the authors expanded to a 24 hour period to obtain 24 hour vehicle classification counts. These 24 hour counts formed the basis for estimating VMT mix on links. The fraction of counts of each vehicle type represented the VMT mix at an individual link level and was the dependent variable in current analysis. The 1996 GIS road network file included information on the characteristics of each link in the Dallas area. The attributes available in the network file included length of link, traffic direction, functional classification, number of lanes, free speed, capacity, and whether the link is divided or not. Next, the land use data were obtained from zonal-level land use characteristics. Data assembly was conducted to assign link/zone characteristics to each link vehicle classification count observation. To accomplish this, authors spatially overlaid the 1996 GIS road network file and the 1996 Zonal coverage file. Then, each link at which the vehicle counts were recorded was manually queried in the network database using the street name and the names of the cross streets at the end nodes. After spatially locating the links, its identification number was extracted from the network file, and the traffic analysis zone that spatially contains the link was identified. Using the link and zonal identifier fields, the relevant link and land use characteristics were mapped to each vehicle count observation.

The authors included the following five sets of independent variables in the model to predict

the VMT mix on links:

1. link functional classification involving five roadway classes;
2. link physical attributes—number of lanes and whether the lane is divided or not.
3. link free speed variables ; and
4. degree of urbanization of the zone in which the link lies, which is characterized by classifying the zone as a central business district, an urban residential area, or a suburb/rural area; and
5. zonal land use characteristics—An airport presence variable, an institution presence variable, zone acreage in retail and office space variable, and acreage in manufacturing plants and warehousing variable.

The authors observed VMT mix in six vehicle types. They compared the actual observed VMT mix with the corresponding model predicted VMT mix and suggested that the same model could be applicable for other areas, also.

Soak Distribution Inputs to MOBILE Source Emissions Modeling

In the study by Venigalla and Pickerell (2002), the methods for deriving soak period distributions from travel survey data are discussed. They have conducted a comprehensive analysis of the 1995 NPTS data to derive soak times.

In deriving soak times from the 1995 NPTS data, the two most relevant files were the vehicle file and the travel day trips file. However, not all the data from these files is necessary for computing the soak times. According to the authors, the start time and the end time of each trip made by the household vehicle is enough to determine the soak time. But information on other variables is needed to compute engine soak period for each trip. The following are the other variables:

1. household identification number,
2. vehicle identification number,
3. model year of the vehicle,
4. geographic classifications,
5. start time and length of the trip, and
6. trip purpose variables.

The above variables were extracted from the NPTS database to create a smaller and more

manageable database for the analysis. The data was then closely examined in an effort to derive the soak times. The examination included:

1. separating highway trips from transit trips,
2. isolating vehicle trips from person trips, and performing
3. consistency checks and taking into account additional considerations.

The soak times derived were then categorized into appropriate MOBILE 6 soak interval categories.

The findings from the study include:

1. the variation in soak times was best explained by the trip purpose variable, closely followed by the time of the day variable;
2. individual MSAs have considerable influence on the soak period distribution; and
3. if local data are not available, or the NPTS local data is insufficient for accurate local estimates, agencies should choose the next level of aggregation to derive soak periods, as these estimates would be an improvement over national default values.

Guidance for the Development of Facility Type VMT and Speed Distributions

The EPA provides guidance on how to estimate VMT on different classes of roadways by time of day and speed and how to investigate other vehicle activity characteristics (EPA, 1999).

One source of VMT information includes state and local transportation and traffic management agencies that collect traffic count data at both permanent and temporary locations. Count data sets are maintained by site descriptors (county, route, distance from a reference point, direction, lane facility type, segment length, and area type) and time features (date, day of week, time of day, duration, and vehicle count). The estimation total VMT of using count data is relatively straightforward but biases in the calculation may exist. The steps are:

1. calculation of the sum of counts in each functional class,
2. determination of the sample size in each functional class,
3. determination of the average volume by dividing the total count by the sample size,
4. the obtaining of miles of facility in each class (from DOTs or GIS databases), and

5. the calculation of VMT by class as the average volume multiplied by the number of miles of facility resulting in a distribution of VMT by facility type.

MOBILE 6 requires that VMT for arterials and freeways be disaggregated by speed. Two methods available to achieve those results are:

1. Use of the Highway Capacity Manual (HCM) in estimation of link speeds.
However, the HCM method requires more information (facility specific) than is likely to be available.
2. Use of the volume/capacity relationship expressed in Bureau of Public Roads (BPR) curves:

$$S = S_f / 1 + a(v / c)^b$$

where

S = predicted mean speed

S_f = free flow speed

V = volume

c = practical capacity (practical capacity = 80 percent of capacity)

a = 0.05 for signalized facilities (arterials)

a = 0.2 for freeways

b = 10

By entering coded capacities by facility type or using default look-up tables along with link volumes from traffic count data, link speeds can be predicted with the BPR equation. VMT within each functional class can be grouped by speed.

Another source of estimating vehicle activity by functional class, time of day, and speed is travel demand models. Travel demand models provide comprehensive regional VMT estimates and avoid uncertainties associated with extrapolation of traffic volumes from count data at selected locations. Agencies with access to travel demand models can readily obtain VMT distributions from the link level traffic volumes. Travel demand model inputs are prepared for specific time periods, and simulation results provide a single assignment representation of that period. For average travel assignments, description of hourly variation in congestion and speeds within the model is not possible. Thus, the external calculation of speeds may be preferable. Post processing software that uses HCM procedures and BPR

curves to calculate hourly congested speeds is available. One speed post-processing procedure is the Direct Travel Impact Model (DTIM2), which operates on daily or peak-period hourly link volumes as follows:

1. The procedure distributes link level volumes by hour of the day using user provided or default temporal distributions (usually from count data sets).
2. The procedure allows for the hourly calculation of VMT through the multiplication of link distribution and hourly volume.
3. The procedure allows for the finding of v/c ratio through the use of link specific capacities.
4. The procedure allows for the application of BPR curve using link specific free flow speed or look up tables to arrive at hourly congested speeds.

A Travel Model Improvement Program (TMIP) workshop on air quality analysis (TMIP, 1994) provides useful information on using travel demand modeling output in supplying input data to MOBILE models. Diurnal factors are used to estimate traffic by hour of the day as a percent of total daily traffic. These factors are estimated using either hourly traffic counts or trip departure times reported in household travel surveys. When using travel demand modeling output in air quality analysis, the following factors should be kept in mind:

1. The error in traffic volumes of a specific link may range from 15 percent to 50 percent, depending on the total traffic volume on the link.
2. Aggregating the link VMT over the entire system may reduce the error to around 5 percent, however, all streets in the area are not covered by the network, so it can be assumed that around 15 percent of the local VMT may not be accounted for by travel model output.
3. Trips made by automobile in search of parking, are not assigned to the network, and some models fail to assign the bus vehicle trips to the network.
4. Link speeds that are output directly from travel models are bound to be in error for various reasons such as the use of unrefined traffic volumes and the use of simplified volume/capacity to speed relationships in modeling travel time.
5. Information on vehicle type can't be directly obtained from travel model output.

The workshop suggests the following methods to refine the traffic volume outputs and make them compatible to MOBILE modeling:

1. Processing traffic volume outputs:

- a. Local missing VMT can be derived from the number of intrazonal trips and intrazonal trip distances. For this we can rely on:
 - i. Intrazonal trip lengths in minutes and output from the trip distribution model.
 - ii. Assumed slower average speeds for intrazonal trips, as they are made on local streets.
 - b. If park/ride lots are not assigned to the network, the missing VMT can be estimated on the basis of the number of such trips (based on output of mode choice models) and the average distance to P/R lots.
 2. Processing link speeds: Speed values must be adjusted, as the speeds output from the traffic models do not reflect the actual speeds. The workshop suggests two methods—a detailed method and a quick method.
 - a. Detailed method:
 - i. Daily traffic data is allocated to a typical peak hour and a typical off-peak hour of the day based on traffic count data for various facility functional classes and area types.
 - ii. V/C ratios are recalculated for each typical hour.
 - iii. Link specific speeds are estimated using V/C relationships based on highway capacity manual.
 - b. Quick method: In this method, VMT is estimated by aggregating network link VMT by facility class/area type. The average adjusted speeds by facility class and area types for two representative time periods, peak and off-peak, can be obtained by using average V/C ratios by facility class and area type. E.g.: An urban area with five facility classes and four area types would have to estimate only 40 adjusted speeds (five classes*four area types*two time periods).
 3. Developing vehicle type data: Vehicle type data can be obtained using two methods, namely, link based and trip based methods. Link based methods require advanced technologies and resources, so we will deal with the trip based methods only. VMT by vehicle type can be obtained by using the trip tables, trip purpose, and distance skims in conjunction with the travel survey data. If a trip based method is used for MOBILE emission modeling, a single region wide distribution of percent VMT will be obtained.

Implementing Soak-time Distribution Models in a GIS Framework

The objective of this work by Sivakumar (2000) is to implement the soak time distribution

models in a GIS framework with the following features:

1. user interactive input
2. output format suitable for MOBILE 6.

Soak time is the duration time in which the vehicle's engine is not operating and which precedes a successful vehicle start, i.e., one that does not result in a stall. A vehicle trip start that occurs after more than 12 hours of soak time is known as a cold start, and engine starts with shorter soak times are called hot starts. MOBILE 6 takes the soak time distribution as input, and divides it into 70 soak time bins. Soak time bins of 1 to 30 minutes are split into bins of 1 minute intervals, and bins of 30-60 minutes are split into bins of 2 minute intervals. Soak times of 60-720 minutes are split into bins of 30-minute intervals. Sivakumar says that the research team at the University of Texas in Austin has developed a model that predicts the zone specific soak time distributions of each combination of time of day and activity purpose. Given the zone to zone production/attraction matrices by trip purpose, one can obtain the number of trip starts by using the zone of origin, the activity purpose prior to trip start, and the time of day. This can be used to compute the zone specific soak time distributions using the soak time distribution models. First, a binary logit model is applied to estimate the fraction of vehicle starts that are first trip starts. Log linear soak time models are then applied to estimate the fraction of trip starts in each soak time bin. From these models, soak time duration depends on the time of day of the trip start, the activity purpose preceding the trip start, the land use, and the demographic characteristics of the zone trip start and trip characteristics. The author implemented the soak time distribution models in a GIS framework so that the output format is suitable for MOBILE 6.

The project focuses on the Dallas area at the TAP (traffic analysis process) zonal level. Each TAP zone is made of several disaggregate TSZs (traffic survey zones, or, synonymously, traffic analysis zones (TAZs)). Zonal land use and demographic characteristics were obtained from North Central Texas Council of Governments (NCTCOG) at the TSZ level data and were aggregated to the TAP zonal level. Shape files of the Dallas area were also obtained from the same source. The following data files and shape files were used:

1. TSZ-TAP equivalence,
2. land use data file,
3. zonal characteristics data file,
4. estimated coefficients,
5. TSZ shape file, and

6. TAP shape file.

The shape files were imported to TransCAD. The coefficient matrix for the soak-time models, together with the land use and demographic models, was then used to apply them in Excel. This yields the soak time distribution by origin activity purpose for each of the six time periods and for each TAP zone in Dallas. The soak time distribution for each time period was then imported to TransCAD as a .dbf file and stored as a TransCAD dataview. This was done for each of the 919 TAP zones. These data views were then linked to corresponding zones so that the attribute table for a zone also included the soak time distributions by time period. This was then input into MOBILE 6. The attribute table for a zone includes links to the soak time distribution tables by time period. Also included was a distribution chart that plots the soak time distribution by activity purpose.

Comparison Between MOBILE 6 and MOBILE 5

MOBILE 6 is the latest of the models developed by EPA to calculate estimates of current and future emissions from highway motor vehicles. MOBILE 6 has many new features that allow for very detailed modeling of specific local situations. The EPA requires states to supply their locality specific transportation and geographic data to MOBILE 6 to attain the vehicle emissions in a more detailed way. Though MOBILE 6 has default values for most of its input requirements, they are based on the national aggregates and may not represent local conditions accurately.

The difference between MOBILE 5 and MOBILE 6 is summarized in Table 3. MOBILE 6 has been the required vehicle emissions model to apply since January 2004, following a grace period of two years in which either MOBILE 5 or MOBILE 6 could be applied. MOBILE 6 is expected to produce more accurate assessments of vehicle emissions, but it requires extra data. The difference in data requirements can be seen in the comparison of features in Table 3.

Table 3
Comparison of MOBILE 5 and MOBILE 6

MOBILE 6	MOBILE 5
1. MOBILE 6 allows the user to enter many activity parameters individually for each hour of the day. For example, MOBILE 6 allows users to enter 14 trip length distributions for separate hourly groups.	1. MOBILE 5 allows a single activity entry only. For example, MOBILE 5 allows users to enter a single trip length distribution.
2. MOBILE 6 allows users to supply 28 individual vehicle classes, and in some contexts, it allows 16 composite vehicle classes.	2. MOBILE 5 allows the user to supply only eight individual vehicle classes.
3. MOBILE 6 allows users to enter 16 VMT values by combined vehicle class. I.e., MOBILE 6 allows users to supply VMT values for gasoline and diesel fueled vehicles combined.	3. MOBILE 5 allows users to enter eight VMT values by individual vehicle class. I.e., MOBILE 5 allows users to supply VMT values for gasoline and diesel fueled vehicles separately.
4. MOBILE 6 allows users to enter registration distribution for 16 vehicle classes	4. MOBILE 5 allows users to enter registration distribution for eight vehicle classes
5. MOBILE 6 uses diesel sales fractions by model year and vehicle type to distribute the overall vehicle counts by combined vehicle class into gasoline fueled and diesel fueled vehicle classes.	5. MOBILE 5 uses diesel sales fractions for light duty passenger vehicles and light duty trucks only.
6. MOBILE 6 allows users to enter mileage accumulation rates for each of the 28 vehicle types.	6. MOBILE 5 allows for the entering of mileage accumulation rates for each of the eight individual vehicle classes.
7. MOBILE 6 does not use any operating mode; instead, it controls the effect of engine starts using the soak distribution and starts per day commands.	7. Each of MOBILE 5's scenarios has three operating modes: cold start mode for non catalyst vehicles, cold start mode for catalyst vehicles, and hot start mode.
8. MOBILE 6 allows the use of a single average speed. It is not possible to model a separate average speed for each vehicle type in a single MOBILE 6 run.	8. MOBILE 5 allows the entry of single average speed or entry of separate average speed for each of the eight vehicle types.
9. There are four roadway types simulated by MOBILE 6 – freeways, arterials/collectors, local roads, and freeway ramps.	9. MOBILE 5 does not differentiate vehicles by roadway type.

Using Global Positioning System (GPS) to Estimate VMT

Background to Investigation of Using GPS to Estimate VMT

Traditionally, VMT-related input are supplied by travel demand models, traffic count data, or personal travel surveys. However, there is always a question of accuracy of these data sources because these methods are not exclusively developed for air quality modeling purposes and are seldom at the detailed level required in emissions modeling. Developing accurate travel related input data for air quality models like MOBILE 6 is becoming a major challenge for the transportation planners involved in air quality modeling. Research is ongoing in this field, and there is a keen interest in testing the use of Global Positioning System's (GPS) data as an alternative means of estimating vehicle activity. More accurate and extensive travel information, which can be obtained using GPS, makes it an attractive alternative means of data acquisition. However, this field is new, and the concept of using GPS instruments to obtain data for air quality models like MOBILE 6 has not been fully tested.

An independent investigation was conducted as a part of this study to test the feasibility of using a sample of vehicles equipped with GPS to supply input requirements for MOBILE 6 (Varanasi, 2003). Instead of conducting a GPS survey, existing survey data is used, as described in greater detail below. The study investigates the supplying of major input requirements, namely VMT by facility type, VMT by hour, and VMT by speed. MOBILE 6 requires data for 28 vehicle classes. However, in this study, the data is limited to light duty gasoline vehicles (LDGV) because GPS instruments were only installed in vehicles of this type. However, the same procedures may be adopted in obtaining the data for other vehicle classes.

There has been considerable interest in using the GPS technology for travel surveys in recent times, despite its complications and technical problems (Stopher, et al., 2003). Significant difficulties with recruiting the households were observed by Stopher, et al. (2003) in a GPS pilot survey done in Sydney, Australia. Research by Wolf, et al. (2003) done on a GPS survey in California, indicates that GPS has great potential to enhance travel surveys. The survey was conducted in three different regions, and GPS VMT values were compared with regular CATI modeled estimates. Trip underreporting in CATI modeled estimates was found to occur in Alameda and Sacramento regions. However, there was slight trip underreporting in GPS values in the San Diego region. A similar study by Pierce, et al. (2003) reported considerable trip underreporting with regular travel surveys, compared to GPS survey values.

Data

The Lexington Study was the first study to collect in-vehicle GPS data and involved a survey period of six days, with 100 households participating. The specific objective of the survey was to test the feasibility of GPS in improving personal travel survey data.

Each household that participated in the survey was provided with a single GPS instrument to be installed in a vehicle used by the household. Along with the primary driver, other drivers in the household were also allowed to use the vehicle outfitted with the GPS instrument (Battelle, 1997).

The raw GPS data collected in Lexington were “absolute” GPS points. That is, they were not differentially corrected for selective availability. Selective availability, which was in operation at the time of the Lexington study, was a purposeful degradation of the GPS signals that could move GPS readings as much as 100 meters in a random manner. Because of the inaccuracy in the GPS data, it needed further processing to match it to roadway nodes and links to identify the accurate distance traveled by the vehicle based on the link length (Battelle, 1997).

The map matching analysis was performed on the Lexington data using software developed by TransCore (formerly JHK & Associates). The results of the map matching analysis (referred to hereafter as “processed data”) provided valuable information such as day of travel, time of day, link ID, distance traveled within that link, travel time, speed, latitude and longitude, user ID, trip number, trip purpose, and the location of the trip.

Matched data for all the households was unavailable because of some problems with the equipment and its use, such as the malfunction of the equipment, signal losses, and occasional failure of the equipment to switch on properly. Considering 100 vehicles, data from 80 vehicles were found to be useful for this analysis. Processed data from all six days of the survey period were available for only 50 percent of the cases. Out of the 80 vehicles, 40 vehicles had 6 days of travel data, 12 vehicles had 5 days of data, 14 vehicles had 4 days of data, 11 vehicles had 3 days of data, and 3 vehicles had 2 days of data.

Another problem encountered with the GPS data was that some of the GPS trips extended out of the Lexington street network. These “out of area” trips were deleted from the database

because they constituted trips that would not normally be included in an air quality analysis of a non attainment area. The GPS layer was attached to the Lexington network, and all of the travel lying outside of the network area was highlighted. A total of 3.21 percent of GPS observations involved “out of area travel,” and this travel was deleted from the parent file. The remaining data were used for further analysis.

VMT by Facility Class

This input option for MOBILE 6 requires that hourly VMT data be supplied for each of the four facility classes. The four facility classes include freeways, arterials, local roads, and freeway ramps, respectively. Under each of the facility classes, 24 hourly VMT values were supplied, starting at 6 a.m. Travel between 6 a.m. and 7 a.m. was considered as the first hour of travel, and travel between 5 a.m. and 6 a.m. next day was considered as the 24th hour of travel. The hourly values from 2 to 23 were assigned in increasing order for the remaining hours lying in between 7 a.m. of the first day and 5 a.m. of the next day. GPS time information was used for this purpose, and for each GPS record, the corresponding hour of the day was assigned.

To classify the VMT data by facility class, the GPS data needed to be analyzed in a GIS environment. TransCAD GIS software was used for this specific purpose. First of all, the GPS data was geo-coded in TransCAD using the latitude and longitude values. The processed GPS data file was then opened as a “data view file” in TransCAD. The GPS data file consisted of all the travel information except for the facility class information. To add the facility class information to the GPS “data view” file, a street network with facility class information was needed.

The Lexington street network, provided by the Battelle team, consists of all the facility class information, except the ramp information. It has different facility types, namely freeways, arterial highway, major arterial, minor arterial, collectors, and local through streets, numbered from 1 to 6, respectively. As mentioned earlier, MOBILE 6 requires facility class information in terms of freeways, arterials, local streets, and ramps. To achieve this classification, the information provided with the network was aggregated to represent freeways, arterials, and local streets. Arterial highways, major arterials, minor arterials, and the collectors were grouped together as the “arterials.” The network then contained all the functional class information, with the exception of the ramp information. To assign the ramp classification to the network, all the ramp links were identified manually and classified as

ramps. The network with all four functional classes is shown in Figure 1.

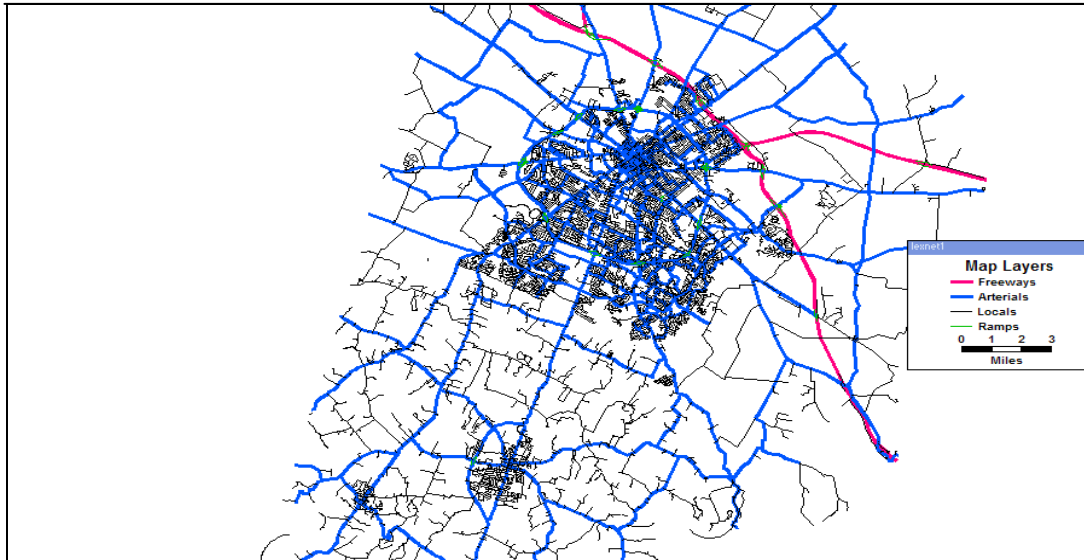


Figure 1
Lexington network divided into four functional classes

To transfer the facility class information into the GPS data, the “Overlay” feature in TransCAD was used. With the use of this feature, a separate data file containing all the features of the GPS layer and the functional classification information of the street network was obtained. The process of overlaying involved the following steps: First, the Lexington network, with a functional classification included, was opened, and the GPS layer was then opened. The process involved assigning the features of underlying layer to the top layer using a bandwidth. Bandwidth is the width of the strip that is to be chosen around the street network, based on which the GIS assigns the street network properties to the GPS layer. For example, if a bandwidth of 60 ft is chosen, the GPS points within 60 ft range of the street will be assigned with the properties of the street.

This procedure is not accurate. Under certain conditions however, if the bandwidths selected for this purpose are too narrow, GPS observations can be lost. Broad bandwidths, on the other hand, will assign the information of surrounding streets to the GPS layer. Hence, the process involved testing various bandwidths and selecting an optimum one. The process works well, provided facilities of different functional classes are not in close proximity to

each other. One such problem area is where ramps merge into freeways or where roads of different functional classifications intersect. To minimize these problems, the street network was divided into four separate layers based on the facility type. This allowed the individual facility class networks to be overlaid separately with GPS layer using different bandwidths. After experimenting with different bandwidth values, a bandwidth of 60 ft was selected for freeways and arterials, and bandwidth of 30 ft was selected for local streets and ramps. This minimized the error but did not eliminate it. In these cases, the roads that were wrongly assigned with respect to ramp classification were manually selected by TransCAD's "Select by Shape" feature and assigned with appropriate facility class information.

After obtaining the GPS layers with functional classification associated with each, VMT values were distributed based on the hour of the day. For this purpose, all the GPS layers were transferred into an access database, and using SQL queries, the hourly VMT values were estimated. In each hour of the day, VMT fractions were obtained for each facility class by dividing the VMT of a specific facility class by the total VMT of all the facility classes. The VMT classification of the GPS data is shown in Figure 2.

Table 4 shows the VMT breakdown according to functional classes in the Lexington area. The Lexington survey value represents the result obtained from the analysis of the Battelle team. The table shows that there are very few freeway miles in the area, which is the reason for relatively little freeway travel in Lexington. The table can also be used as a check to this innovative method of classifying the GPS VMT data into facility types. In observing the table, one may observe that the aggregated results of study are more or less similar to the VMT breakdown developed by the Battelle team. The fact that Lexington network, the ramp values were unavailable at the time of analysis by the Battelle team in the Lexington network should be noted. These values were introduced manually by identifying the ramp links and modifying the network in the present study. Higher percentages of arterial and freeway travel were found by the Battelle team than found in this study. This is probably the result of including ramps with the arterials and freeways in their analysis.

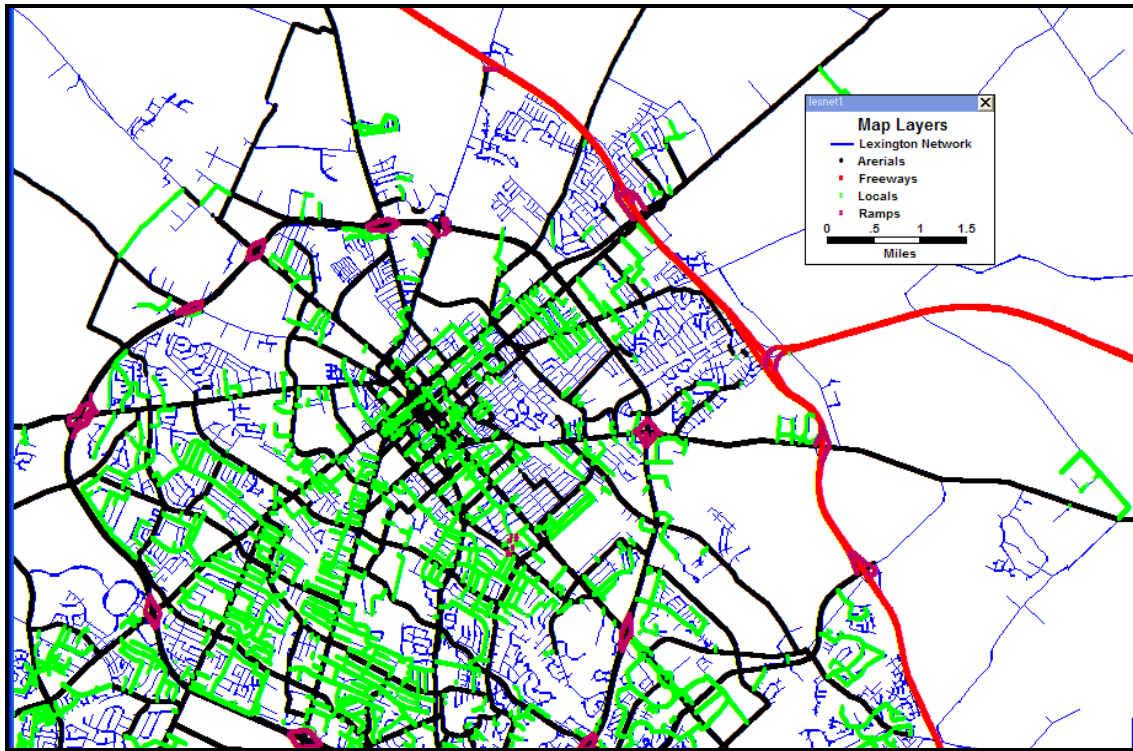


Figure 2
Functional classification of the GPS travel in Lexington area

Table 4
VMT breakdown by functional classes in Lexington

Functional Class	VMT	
	Lexington Survey Value (%)	Overlay Procedure Value (%)
Freeways	2.87	2.31
Arterials	81.62	81.12
Locals	15.5	15.47
Ramps	Unavailable	1.1

In Table 5, the VMT by facility class and hour of the day from GPS are shown. The default values are also shown, for comparison purposes. The results are shown graphically in Figures 3-6. The GPS and default values in Table 5 are clearly different. From GPS

observations, one may observe that most of the travel in Lexington was conducted on arterials, and relatively little travel was on freeways when compared with the default values. The reason could be that there are relatively few freeways in the Lexington area; longer distance trips in the urban area must, therefore, be made on arterials. VMT by facility class can appear to vary from locality to locality and depends on local network conditions. Travel on local streets and ramps is generally similar to the default values.

Table 5
VMT breakdown by functional classes for MOBILE 6

	Freeway		Arterial		Local		Ramp	
Hour	GPS	Default	GPS	Default	GPS	Default	GPS	Default
1	0.01	0.39	0.98	0.46	0.01	0.12	0.01	0.03
2	0.01	0.34	0.84	0.5	0.15	0.13	0.01	0.03
3	0.00	0.34	0.85	0.5	0.13	0.14	0.01	0.03
4	0.00	0.35	0.84	0.49	0.16	0.13	0.00	0.03
5	0.04	0.35	0.77	0.5	0.17	0.13	0.01	0.03
6	0.08	0.33	0.76	0.51	0.15	0.13	0.01	0.03
7	0.02	0.32	0.83	0.52	0.14	0.13	0.01	0.03
8	0.03	0.33	0.82	0.51	0.14	0.13	0.01	0.03
9	0.01	0.33	0.84	0.51	0.14	0.13	0.01	0.03
10	0.03	0.32	0.80	0.52	0.16	0.13	0.01	0.03
11	0.01	0.33	0.78	0.51	0.20	0.14	0.01	0.03
12	0.03	0.31	0.80	0.52	0.15	0.14	0.02	0.03
13	0.01	0.3	0.83	0.54	0.15	0.14	0.01	0.03
14	0.02	0.31	0.84	0.53	0.13	0.14	0.01	0.03
15	0.01	0.33	0.82	0.51	0.16	0.13	0.01	0.03
16	0.04	0.34	0.81	0.5	0.14	0.13	0.00	0.03
17	0.03	0.38	0.86	0.46	0.10	0.13	0.01	0.03
18	0.00	0.41	0.90	0.44	0.09	0.12	0.00	0.04
19	0.00	0.43	0.77	0.42	0.21	0.12	0.02	0.04
20	0.00	0.44	0.86	0.4	0.10	0.12	0.03	0.04
21	0.02	0.46	0.94	0.39	0.03	0.11	0.01	0.04
22	0.00	0.46	0.94	0.39	0.06	0.11	0.00	0.04
23	0.00	0.45	0.80	0.4	0.20	0.11	0.00	0.04
24	0.07	0.42	0.86	0.43	0.01	0.11	0.06	0.04

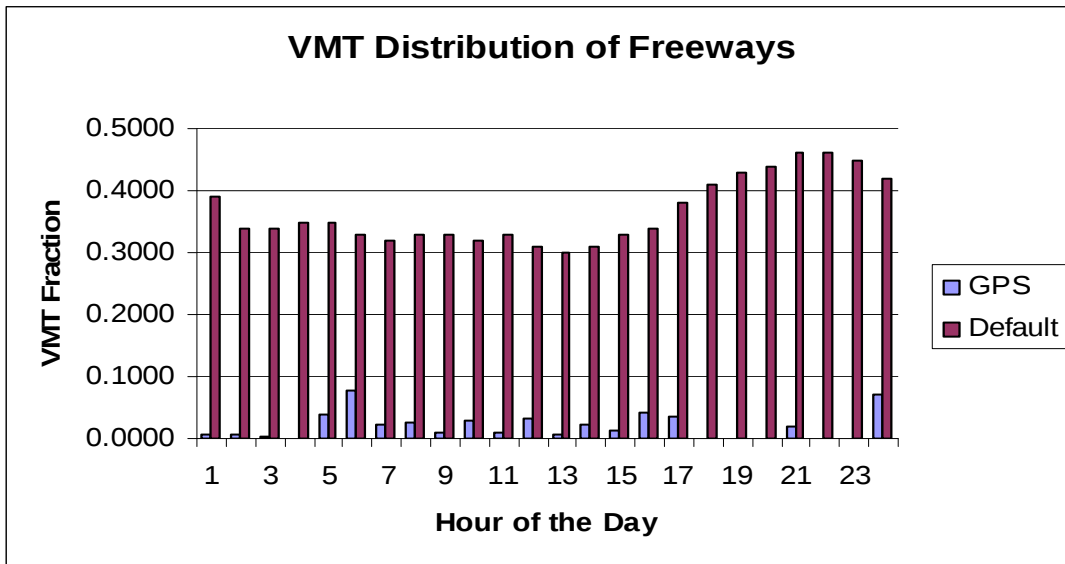


Figure 3
Distribution of VMT by freeways for MOBILE 6

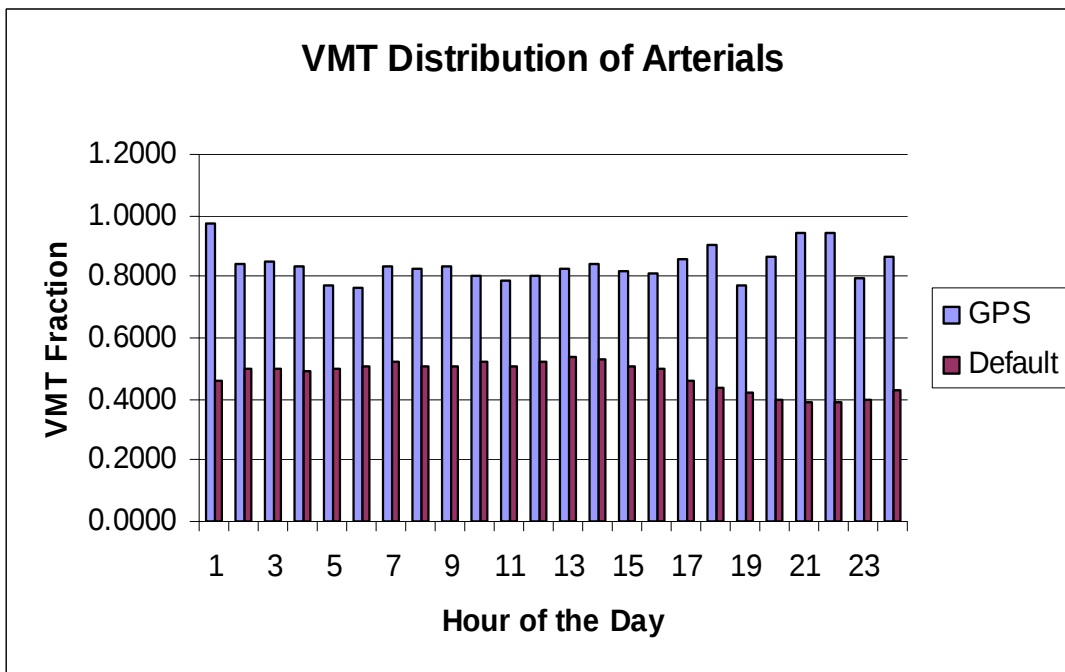


Figure 4
Distribution of VMT by arterials for MOBILE 6

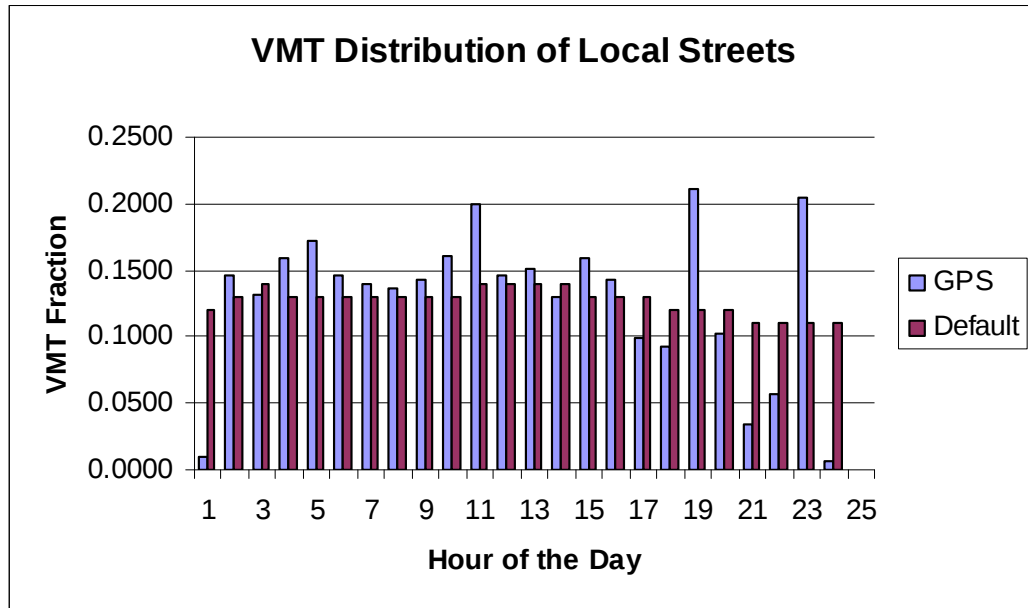


Figure 5
Distribution of VMT by local streets for MOBILE 6

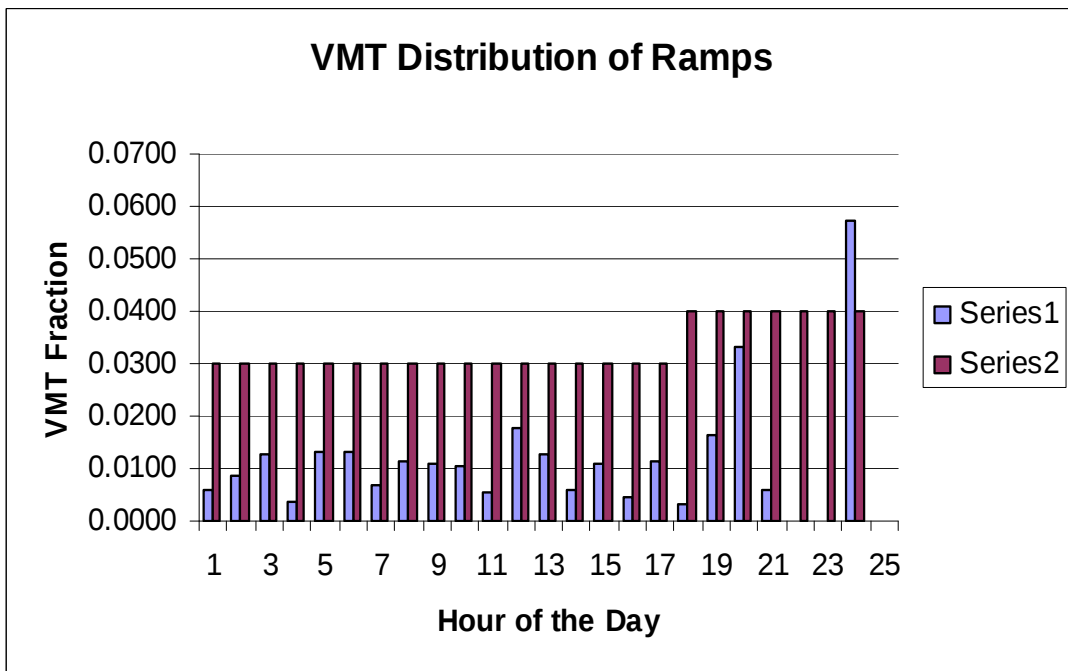


Figure 6
Distribution of VMT by ramps for MOBILE 6

The Chi-square test statistic, found in M.S. Excel software, was used in this study. The software calculates the Chi-square value and provides a chi-squared probability value for the given degrees of freedom. If the Chi-squared probability value does not exceed 0.05, then one can state that the expected and observed values are different at 95 percent confidence limit. The software calculates the Chi-square value using the following formula:

$$\chi^2 = \sum_{i=1}^r \frac{(A_{ij} - E_{ij})^2}{E_{ij}}$$

where

A_{ij} = actual observations in the i-th row (GPS distribution values)

E_{ij} = expected observations in the i-th row (default distribution values)

r = number of rows

Hypothesis testing was conducted using the obtained Chi-squared probabilities at a five percent level of significance. The null hypothesis was that the default and GPS values were equal, while the alternative hypothesis was just the opposite. The Chi-square test was applied separately on the results of each facility class. To apply the Chi-square test on the results, each cell value in the expected observations should be greater than five, and there should not be any null values. To make sure that the individual cells in the expected range (default values) have values greater than five, the VMT in the individual cells was used instead of using the VMT fractions. The total VMT in the individual facility classes of the default input data was assumed to be equal to those of the GPS input data for the purpose of obtaining the VMT by hour of the day values of the default data. The VMT fractions in each cell in the individual facility classes of the default input data were multiplied by the total VMT in those facility classes to obtain the corresponding VMT values in each cell. A Chi-squared probability value of zero was obtained, with 23 degrees of freedom for the VMT by facility class distribution results. Since Chi-squared probabilities for the VMT distribution fractions were less than 0.05, the null hypothesis was rejected. Hence, the GPS and default values are different at a five percent level of significance.

VMT by hour of the day

MOBILE 6 also requires that 24 VMT fractions be supplied for each hour of the day for all facilities combined. The only difference between this input and that in the previous section is

that the VMT is aggregated by facility class. The VMT was distributed into 24 hours of the day, and fraction of travel in each hour of the day was obtained. The results are shown in Table 6. The results are interpreted graphically in Figure 7.

In observing Figure 3 one can see that the GPS and default distributions follow similar patterns. The graph shows that there are two major peaks and a minor peak in both GPS and default distributions. The only difference is that the peak periods are more pronounced in the GPS data, and there is a slight difference in the time when the peaks are attained. The morning peak period in the GPS distribution is between 8 and 9 a.m. while it is between 7 and 8 a.m. in the default distribution. The evening peak period for the GPS distribution is between 5 and 6 p.m., while it is between 4 and 5 p.m. for the default distribution. Between the morning and evening peaks, a small peak can be observed at 12 p.m. in both the distributions. The temporal variations can be attributed to the fact that, the default data was collected from cities comparatively bigger than Lexington. Another difference is that the default distribution has so-called peak spreading, which typically occurs in bigger cities but may not be applicable in Lexington.

Table 6
VMT by Hour of the Day Input for MOBILE 6

Hour of the Day	GPS Value	Default Value
1	0.0117	0.0569
2	0.0458	0.0740
3	0.0827	0.0655
4	0.0356	0.0555
5	0.0382	0.0540
6	0.0515	0.0582
7	0.0783	0.0608
8	0.0766	0.0571
9	0.0666	0.0598
10	0.0630	0.0636
11	0.0790	0.0777
12	0.1062	0.0730
13	0.0694	0.0501
14	0.0432	0.0389
15	0.0446	0.0308
16	0.0396	0.0264

17	0.0233	0.0194
18	0.0123	0.0144
19	0.0108	0.0108
20	0.0098	0.0086
21	0.0053	0.0081
22	0.0019	0.0080
23	0.0017	0.0098
24	0.0029	0.0186

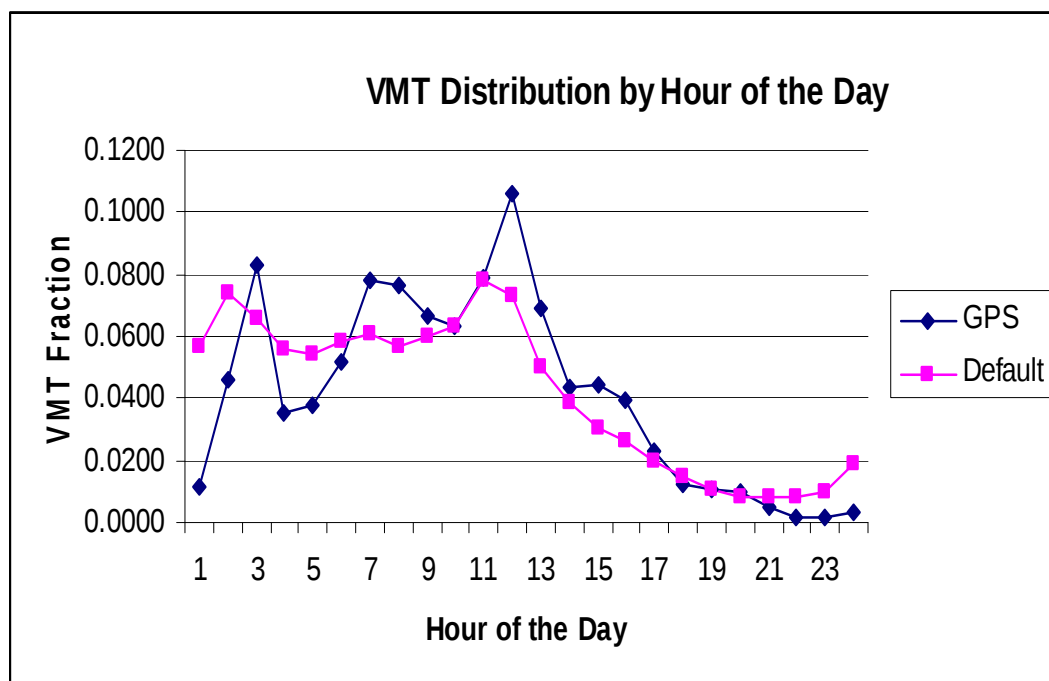


Figure 7
VMT Distribution by hour of the day for MOBILE 6

Hypothesis testing was conducted using the Chi-squared test. The null hypothesis was that the default and GPS values were equal. The Chi-squared test was applied using the GPS values as observed results and the default values as expected results. To make sure that the individual cells of the default input data have values greater than five, the frequency of vehicles in the individual cells of the GPS and default input values were used in the statistical test, rather than the input fractions. To obtain the number of vehicles in the individual hours of the day of the default input data, the assumption that the total number of vehicles traveling in a particular day was equal in both GPS and default data sets was made. The input fractions in the individual cells of the default data were multiplied by the total number of vehicles

traveling in a day in an effort to obtain the frequency of vehicles in the corresponding individual cells. The null hypothesis was rejected at a five percent level of significance, meaning that the GPS and default values are significantly different.

VMT by Speed

The VMT by speed command requires VMT by average speed (preselected ranges) on freeways and arterial roads for each hour of the day. Within each facility type and each hour of the day, travel in terms of VMT is divided into 14 speed bins. The 14 speed ranges are allocated as follows: The first bin consists of “idle,” which speeds are below 2.5 mph. The second bin consists of speeds between 2.5 mph and 5 mph. From 5 mph to 60 mph, there are 11 speed bins at 5 mph intervals. The last bin contains all speed values greater than or equal to 65 mph. The GPS VMT data were first taken on freeways, and the speed data were classified into 14 speed bins using SQL queries in the MS Access database. Speed VMT fractions were obtained for each hour of the day, such that all 14 fractions add up to one. This process was done by dividing each speed bin VMT by the total VMT on freeways of all 14 speed bins in a particular hour of the day. The same procedure was applied to the arterial VMT data. The results are shown in Tables 7 and 9 for freeways and arterials, respectively. The number of vehicles in the sample in each hour of the day is also shown to illustrate the low level of travel on freeways. The corresponding default values are shown in Tables 8 and 10. Graphical results for freeways and arterials are shown in Figures 8 and 9, respectively.

Table 7
VMT distribution by speed for freeways (GPS Values)

	Speed Bin														# of Vehs
Hr	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
1									0.00	0.01	0.01	0.05	0.08	0.86	6.00
2					0.00	0.01	0.01	0.01	0.02	0.03	0.09	0.35	0.28	0.20	10.00
3							0.01	0.08	0.09	0.10	0.33	0.39			6.00
4	0.07	0.93													2.00
5							0.01	0.00	0.01	0.03	0.04	0.07	0.26	0.59	8.00
6										0.00	0.32	0.20	0.41	0.07	5.00
7											0.00	0.20	0.75	0.05	4.00
8								0.00	0.00	0.02	0.05	0.53	0.07	0.34	7.00
9	0.00	0.00	0.01	0.01				0.09	0.02	0.06	0.15	0.12		0.54	11.00
10	0.00	0.00	0.01	0.02	0.01	0.01	0.01	0.02	0.03	0.04	0.06	0.11	0.13	0.54	15.00
11		0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.02	0.13	0.66	0.15	13.00
12						0.00	0.00	0.01	0.01	0.03	0.03	0.27	0.19	0.46	9.00
13			0.00	0.02					0.01	0.01	0.05	0.12	0.14	0.65	8.00
14							0.00	0.00	0.00	0.02	0.07	0.15	0.05	0.70	8.00
15											0.02	0.06	0.50	0.42	4.00
16								0.00	0.01	0.01	0.06	0.55	0.10	0.27	7.00
17											0.01	0.14	0.24	0.61	4.00
18															0.00
19															0.00
20															0.00
21											0.12	0.36	0.37	0.15	4.00
22															0.00
23															0.00
24								0.00	0.01	0.02	0.04	0.06	0.17	0.69	7.00

Table 8
VMT distributions by speed for freeways (default values)

	Speed Bin													
Hr	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	0.008	0.027	0.021	0.022	0.022	0.038	0.034	0.054	0.061	0.07	0.251	0.115	0.255	0.021
2	0.026	0.007	0.008	0.016	0.028	0.033	0.034	0.036	0.036	0.044	0.245	0.173	0.302	0.013
3	0.026	0.003	0.006	0.006	0.013	0.028	0.034	0.035	0.041	0.037	0.218	0.107	0.44	0.013
4	0.015	0.01	0.002	0.002	0.004	0.017	0.023	0.037	0.042	0.045	0.225	0.119	0.442	0.018
5	0.008	0.009	0.005	0.003	0.004	0.016	0.023	0.036	0.038	0.042	0.235	0.117	0.445	0.018
6	0.007	0.003	0.004	0.01	0.012	0.024	0.029	0.033	0.04	0.039	0.229	0.101	0.454	0.014
7	0.01	0.002	0.006	0.009	0.015	0.028	0.034	0.033	0.035	0.035	0.229	0.096	0.455	0.013
8	0.008	0.008	0.005	0.004	0.005	0.018	0.026	0.038	0.038	0.042	0.226	0.112	0.451	0.018
9	0.011	0.007	0.005	0.002	0.004	0.021	0.028	0.036	0.038	0.052	0.215	0.115	0.448	0.018
10	0.016	0.008	0.003	0.004	0.008	0.027	0.032	0.036	0.032	0.039	0.212	0.064	0.5	0.018
11	0.016	0.041	0.023	0.02	0.028	0.032	0.05	0.049	0.045	0.056	0.222	0.109	0.296	0.015
12	0.019	0.011	0.005	0.011	0.018	0.026	0.049	0.038	0.031	0.053	0.224	0.124	0.374	0.017
13	0.018	0.006	0.001	0.002	0.003	0.016	0.019	0.032	0.036	0.052	0.213	0.067	0.518	0.017
14	0.014	0.004	0.003	0.001	0.001	0.009	0.018	0.026	0.026	0.055	0.206	0.098	0.521	0.018
15	0.009	0.003	0.003	0	0.001	0.007	0.017	0.022	0.026	0.048	0.217	0.105	0.523	0.02
16	0.005	0.002	0.002	0	0.001	0.005	0.016	0.018	0.025	0.04	0.228	0.112	0.525	0.023
17	0.003	0.001	0.001	0	0.001	0.003	0.015	0.015	0.025	0.035	0.235	0.116	0.526	0.025
18	0.001	0	0.001	0	0.001	0.002	0.014	0.013	0.024	0.033	0.239	0.119	0.527	0.026
19	0	0	0.001	0	0.001	0.001	0.014	0.012	0.024	0.03	0.242	0.121	0.527	0.026
20	0	0.001	0	0	0	0.001	0.012	0.01	0.02	0.024	0.245	0.129	0.527	0.032
21	0	0	0.001	0	0	0	0.01	0.009	0.018	0.021	0.246	0.132	0.527	0.035
22	0	0.001	0	0	0	0	0.011	0.008	0.017	0.02	0.245	0.134	0.527	0.036
23	0.002	0	0	0.001	0	0.001	0.012	0.01	0.021	0.022	0.245	0.127	0.527	0.031
24	0.003	0	0	0.001	1E-04	0.001	0.013	0.012	0.024	0.027	0.24	0.123	0.527	0.028

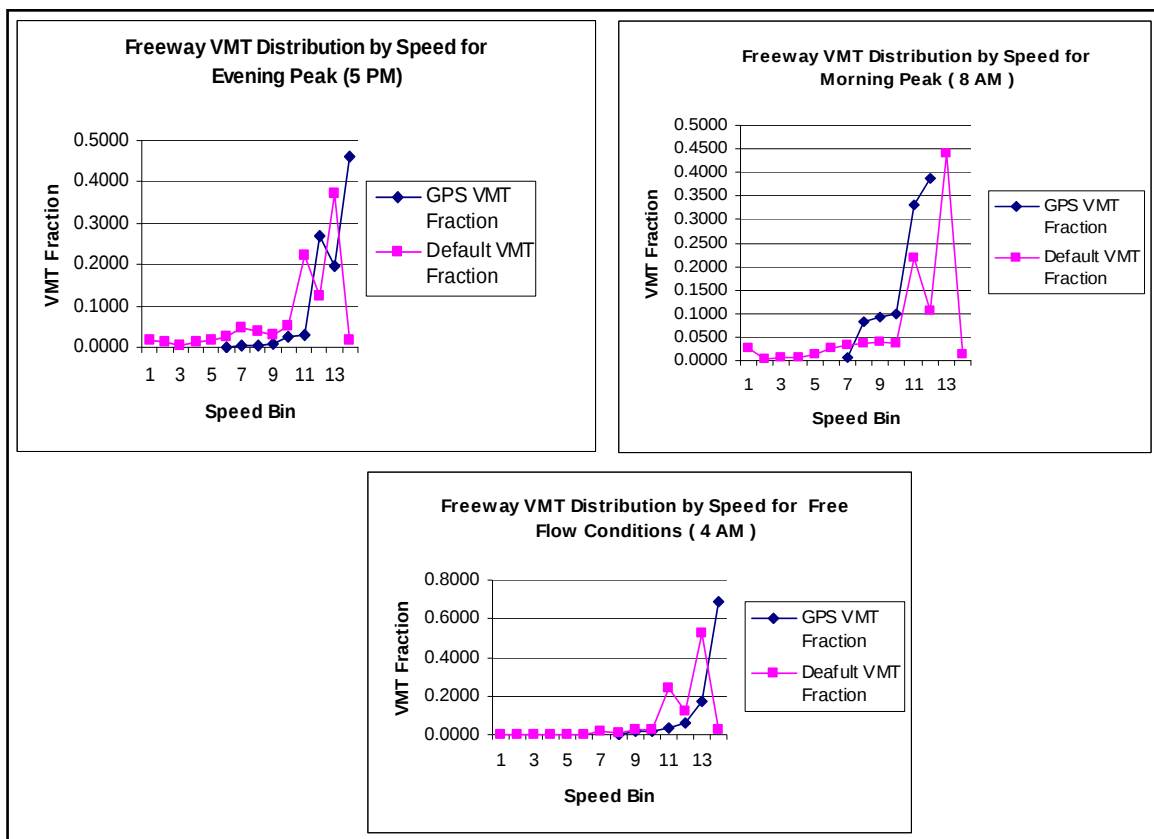


Figure 8
Graphical interpretation of speed distribution results for freeways

In viewing Figure 8, one may observe that the GPS and default values follow similar distributions, although the GPS distributions are relatively erratic due to the small sample sizes. The graphs are drawn for three different time periods in an effort to display the variation for different times of the day, although distributions are available for all 24 hours of the day. The morning peak period starts at 8 a.m. and ends at 9 a.m.; the evening peak period starts at 5 p.m. and ends at 6 p.m.; and the free flow period has been selected as starting at 4 a.m. and ending at 5 a.m.

In viewing Figure 8, one can see that the default values represent more congested flow conditions (as represented by lower speeds) in the morning and evening peak periods than the GPS values from Lexington. Considering the morning peak graph, one may observe that the travel speeds of the GPS vehicles range from 30 mph to 60 mph (speed bins 7 and 12). Almost 70 percent of the total travel was made with speeds greater than 50 mph. The default

distribution, on the other hand, shows more congested conditions, in that only 45 percent of the travel was made with speeds greater than 50 mph. The default travel speeds range from 0 mph to 65 mph (speed bins 1 and 13) in this particular peak hour.

The speed distribution values can be a good measure of the level of service, and one can conclude that the level of service of the Lexington freeways is above the national average. Considering the evening peak graph, a similar conclusion can be drawn. More than 90 percent of the travel by the GPS vehicles was made at speeds greater than 50 mph. The travel speeds range from 20 mph to 65+ mph (speed bins 3 and 14) in the GPS class. The default vehicle speeds range from 0 mph to 65 mph, and more than 70 percent of the travel was made at speeds greater than 50 mph.

Table 9
VMT distribution by speed for arterials (GPS values)

Hr	Speed Bin														# of vehs.
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
1	0.00	0.01	0.01	0.02	0.04	0.09	0.10	0.08	0.07	0.09	0.15	0.23	0.09	0.01	168
2	0.01	0.02	0.04	0.05	0.07	0.10	0.12	0.12	0.11	0.11	0.10	0.09	0.05	0.02	840
3	0.01	0.01	0.02	0.03	0.04	0.05	0.07	0.10	0.12	0.12	0.13	0.16	0.12	0.02	1341
4	0.01	0.02	0.03	0.04	0.06	0.08	0.10	0.15	0.14	0.15	0.10	0.09	0.03	0.01	747
5	0.02	0.03	0.04	0.06	0.08	0.09	0.13	0.15	0.13	0.08	0.05	0.05	0.06	0.04	933
6	0.01	0.03	0.04	0.05	0.06	0.09	0.13	0.15	0.11	0.10	0.09	0.09	0.04	0.02	1094
7	0.02	0.03	0.04	0.05	0.07	0.09	0.12	0.14	0.11	0.12	0.10	0.07	0.04	0.01	1499
8	0.01	0.02	0.03	0.04	0.06	0.09	0.12	0.13	0.12	0.11	0.11	0.10	0.06	0.01	1679
9	0.03	0.03	0.04	0.05	0.06	0.09	0.12	0.12	0.10	0.08	0.09	0.10	0.07	0.02	1508
10	0.01	0.03	0.03	0.05	0.07	0.09	0.12	0.12	0.10	0.10	0.09	0.11	0.08	0.01	1259
11	0.03	0.03	0.03	0.05	0.06	0.08	0.12	0.11	0.11	0.12	0.09	0.09	0.06	0.01	1589
12	0.01	0.02	0.04	0.04	0.05	0.07	0.11	0.12	0.11	0.12	0.11	0.09	0.07	0.02	1893
13	0.01	0.02	0.03	0.04	0.04	0.07	0.10	0.11	0.12	0.13	0.15	0.12	0.06	0.02	1380
14	0.02	0.02	0.03	0.04	0.06	0.09	0.13	0.15	0.11	0.11	0.10	0.11	0.03	0.00	942
15	0.02	0.01	0.02	0.04	0.04	0.07	0.09	0.13	0.12	0.12	0.14	0.13	0.06	0.00	995
16	0.01	0.02	0.02	0.04	0.06	0.11	0.15	0.15	0.15	0.10	0.09	0.07	0.02	0.01	831
17	0.00	0.00	0.00	0.01	0.01	0.01	0.03	0.03	0.05	0.08	0.06	0.03	0.01	0.00	390
18	0.00	0.01	0.02	0.03	0.04	0.07	0.13	0.16	0.11	0.16	0.12	0.11	0.04	0.02	226
19	0.00	0.03	0.03	0.07	0.04	0.06	0.14	0.15	0.11	0.09	0.12	0.10	0.04	0.01	166
20	0.00	0.01	0.03	0.03	0.05	0.10	0.09	0.12	0.11	0.11	0.08	0.16	0.09	0.01	138
21	0.00	0.00	0.01	0.01	0.02	0.04	0.09	0.10	0.17	0.18	0.21	0.08	0.05	0.02	89
22	0.00	0.01	0.06	0.04	0.04	0.05	0.07	0.08	0.36	0.17	0.04	0.06			22
23	0.01	0.04	0.06	0.06	0.11	0.09	0.14	0.17	0.19	0.11	0.02	0.01			44
24	0.01	0.01	0.01	0.04	0.04	0.05	0.07	0.12	0.15	0.19	0.14	0.16	0.03	0.01	70

Table 10
VMT distribution by speed for arterials (default values)

	Speed Bin													
Hr	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	0.000	0.005	0.006	0.005	0.016	0.085	0.321	0.138	0.28	0.06	0.063	0.01	0.01	0
2	0.004	0.003	0.006	0.023	0.074	0.111	0.284	0.095	0.263	0.04	0.07	0.011	0.017	0
3	0.003	0.002	0.003	0.009	0.044	0.113	0.291	0.108	0.284	0.042	0.072	0.009	0.02	0
4	0.003	0.002	0.001	0.002	0.018	0.1	0.291	0.125	0.301	0.054	0.074	0.009	0.02	0
5	0.003	0.001	0.000	0.002	0.018	0.101	0.29	0.125	0.302	0.054	0.075	0.009	0.02	0
6	0.003	0.002	0.002	0.005	0.034	0.109	0.289	0.113	0.293	0.046	0.074	0.009	0.021	0
7	0.004	0.002	0.003	0.008	0.043	0.113	0.286	0.108	0.289	0.043	0.072	0.009	0.021	0
8	0.004	0.003	0.002	0.002	0.022	0.103	0.283	0.124	0.302	0.052	0.074	0.009	0.02	0
9	0.004	0.002	0.002	0.003	0.025	0.105	0.284	0.122	0.299	0.049	0.075	0.009	0.021	0
10	0.005	0.003	0.003	0.009	0.045	0.115	0.282	0.102	0.284	0.042	0.078	0.01	0.023	0
11	0.005	0.017	0.009	0.022	0.065	0.122	0.281	0.096	0.256	0.041	0.065	0.01	0.013	0
12	0.006	0.007	0.008	0.022	0.068	0.117	0.277	0.092	0.264	0.039	0.071	0.011	0.019	0
13	0.004	0.002	0.002	0.004	0.026	0.101	0.285	0.121	0.3	0.05	0.076	0.01	0.021	0
14	0.004	0.002	0.002	0.002	0.012	0.073	0.292	0.122	0.317	0.064	0.079	0.01	0.021	0
15	0.004	0.002	0.001	0.002	0.01	0.056	0.304	0.107	0.331	0.07	0.082	0.01	0.021	0
16	0.004	0.002	0.000	0.001	0.011	0.053	0.306	0.106	0.332	0.071	0.083	0.01	0.021	0
17	0.003	0.001	0.001	0.002	0.01	0.053	0.307	0.106	0.333	0.071	0.083	0.01	0.021	0
18	0.003	0.001	0.002	0.002	0.01	0.053	0.306	0.106	0.333	0.07	0.083	0.01	0.021	0
19	0	0	0	0	0.009	0.05	0.33	0.105	0.331	0.07	0.073	0.01	0.021	0
20	0	0	0	0	0.008	0.05	0.33	0.106	0.329	0.07	0.076	0.01	0.021	0
21	0	0	0	0	0.008	0.049	0.331	0.106	0.33	0.069	0.076	0.01	0.021	0
22	0	0	0	0	0.008	0.049	0.329	0.106	0.332	0.069	0.076	0.01	0.021	0
23	0	0	0	0	0.008	0.05	0.329	0.106	0.331	0.07	0.076	0.01	0.021	0
24	0	0	0	0	0.009	0.05	0.327	0.105	0.332	0.07	0.075	0.01	0.021	0

Considering the free flow period graph, GPS travel can be seen to have been made with higher speeds than the default values. More than 95 percent of the GPS travel was made with speeds greater than 50 mph, while only 85 percent of the default travel was made with speeds greater than 50 mph. The reason for the more congested conditions in the default observations could be attributed to the fact that the cities of default data collection have more congested freeway travel than Lexington. Another possible reason for this outcome could be that only about 4.7 percent of the road network in Lexington consisted of freeways, meaning that most travel would have occurred on arterials.

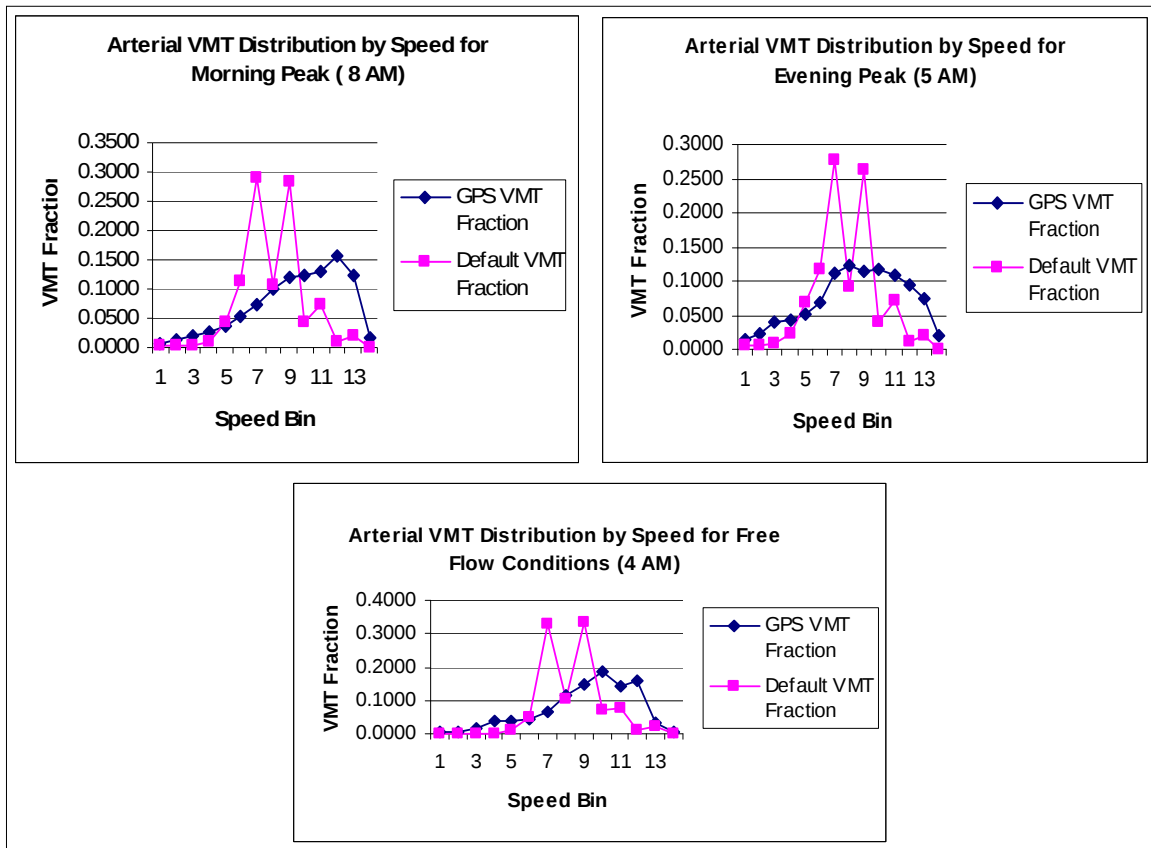


Figure 9
Graphical interpretation of speed distribution results for arterials

Similar to the freeway speed distribution graphical analysis, the arterial results also can be viewed using Figure 9. Unlike the situation with the freeway travel, the GPS values in arterial travel show some signs of congestion, although it is limited. Considering the morning peak and the evening peak of the GPS curves, one can observe that most travel was made at speeds ranging from 30 mph to 60 mph (speed bins 6 to 12). The default values display two peaks at the 7th and 9th speed bins. Overall, speeds in the default distributions appear to be lower than in the GPS distributions. The free flow conditions graph also shows similar travel conditions. One can observe that a higher percentage of travel is noticeable with greater speeds in the GPS distributions than in the default distributions. The reason could be that the Lexington road network has a comparatively higher number of arterial highways (with fewer traffic signals) than the cities in which the default data were collected. The congestion in the arterials may not be the reason for lower speeds in the default data

because the free flow graph has similar distribution when compared with the other two peak period graphs.

Hypothesis testing was conducted using the obtained Chi-squared probabilities at a five percent level of significance. The null hypothesis was that the default and GPS values were equal, while the alternative hypothesis was just the opposite. The results were statistically compared using a Chi-squared test, as applied in the previous sections. The statistical test was applied on the selected hours of the day values, and the graphical comparison was also made on them. Hour of the day columns numbered 3, 12, and 24 in both freeway and arterial results were used for this purpose. Instead of being compared with fractions in the input data, the corresponding VMT values were used to apply the statistical test to make sure that the individual cell values of the default input data were greater than five. For this purpose, the total VMT in a particular hour of the day was assumed to be equal in both GPS and default input data. The VMT in the individual speed bins in each hour of the day of the default input data were obtained by multiplying the total VMT in those hours of the day by the corresponding VMT fractions.

The null values in the individual cells of the default input data were eliminated by merging the adjacent cells. Corresponding cells in the GPS data were also merged to obtain equal matrix size for the GPS and default input data sets. In this process, hour-of-the-day columns for the freeway resulted in 13 rows, and those of the arterial resulted in 9 rows. The Chi-squared test was applied on these columns individually, and probability values of zero were obtained for all hours of the day. Since Chi-squared probabilities for the VMT distribution fractions were less than 0.05, the null hypothesis was rejected. Hence, the GPS and default values are different at a 5 percent level of significance for both freeway and arterial classes.

Conclusions Regarding the Use of GPS to Estimate VMT

In this research project, a new method of estimating VMT by functional classification using GPS data from Lexington, Kentucky was tested. Using the overlay procedure within TransCAD, the street network attributes were attached to the GPS data, and the functional classification analysis was performed. The analysis produced very similar results to those obtained by the Battelle team on the same data. However, the VMT estimates obtained from this method differed considerably from the national default values. A low proportion of freeway VMT and a high proportion of arterial travel were found. This observation was

attributed to the network conditions in Lexington. The analysis reinforces the EPA's statement that local values should be used wherever possible instead of applying default values.

VMT by hour of the day was developed by aggregating the VMT by functional classification. Although the statistical results indicated that the GPS values were different from the default values, the graphical results showed similar distributions. The only difference observed was that the peak periods differed by time of day.

The VMT distribution by speed input was obtained for freeways and arterials as required by the MOBILE 6. On both freeways and arterials the observation that the GPS speeds were higher than the default vehicle travel was made. Freeway travel in Lexington does not have VMT values in the lower speed bins, indicating a high level of service on those facilities. These outcomes were attributed to the congestion levels and the road network conditions in the area.

The study results indicate that GPS can be effectively used in supplying VMT related inputs to air quality models like MOBILE 6. The use of microscopic data sources like GPS to supply macroscopic data to MOBILE 6 was proven to be feasible. Although the process of collecting and analyzing the GPS data is expensive, the high accuracy of the data makes it an attractive alternative. The procedure can be helpful in future research, as the EPA is introducing microscopic models to estimate vehicle emissions in the near future.

An Alternative Approach to Estimating VMT

VMT estimates for input to MOBILE 6 are typically obtained either from traffic counts or travel demand models. The EPA is of the opinion that sample traffic counts conducted as part of the Highway Performance Monitoring System (HPMS) are the preferred source of data on which to base these estimates (EPA, 1992). The HPMS method of VMT estimation involves identifying volume counts on road sections distinguished by their functional classification, location, and volume group (FHWA, 2000). Specifically, a functional classification of interstate, other freeways or expressways, major arterial, minor arterial, and collector are used. Location is divided into rural, small urban areas (< 50,000 population), urbanized areas with population between 50,000 and 200,000, large urbanized areas (> 200,000), and the "donut" area, the area beyond the metropolitan boundary but within the

National Ambient Air Quality System area. There are 13 volume groups in each functional group and location type, varying according to the type of functional class and location (FHWA, 2000, Appendix C). For example, for rural interstates, the first nine volume groups are in 10,000 vehicles per day (vpd) increments from 0 to 90,000; the next three increments are 15,000 vpd up to 135,000, and the last volume group is all traffic volumes in excess of 135,000 vpd. For lower level roads, the volume group intervals are smaller.

VMT is obtained in the HPMS method by assuming that the annual average daily traffic (AADT) observed within each category of road (distinguished by functional class, location, and volume group) applies to all road sections within that group, and the VMT is obtained by multiplying the AADT by the respective length of the sections in the group and summing the product over all sections. The method produces accurate estimates, but it does rely on the rather odd premise that the volume group is known. Thus, each road section must be classified into a volume group before the volume on that section can be estimated. They argue that by rotating traffic counts among road sections and observing trends in traffic growth, it is not too difficult to assign a road section to the appropriate volume group. However, the process does involve judgment and does require maintenance.

The question arises, is there an alternative method of estimating VMT from HPMS data that does not require the use of volume groups? Vadlamani investigated a procedure that replaced the volume group classification in the existing method with a classification based on the characteristics of the road, such as number of lanes, lane width, and speed limit (2005). A regression tree classification procedure was used to identify which features most distinguished link volume. A distinction between urban and rural locations, functional classification, number of lanes, lane width, and speed limit was found to capture 80 percent of the variation in the volume count data. This turned out to be not as good as what is achieved with the conventional method using the volume group classification; although, the difference was not that great. For example, the correlation coefficient between estimated and observed AADT values on the links that had traffic counts was 0.93 with the method developed in this study. It was 0.99 with the conventional method using volume group. The root mean square error of AADT estimates from the method developed in this study was 1,918 vehicles per day, while it was 548 with the conventional method. Clearly, the conventional method performs better when compared with road sections where the correct volume group is known (because traffic counts exist), but how well it would do among road sections where the volume group is not known and must be guessed is not known. To say that

the conventional method of estimating VMT from HPMS data will work well provided good estimates of the correct volume group for links without traffic counts are made would seem safe. Whenever a link is assigned to the wrong volume group, large errors are likely to ensue. The method developed in this study does not use volume groups and, therefore, does not suffer from that disadvantage.

Daily VMT estimates by functional class from the procedures described above must be further disaggregated for use in MOBILE 6. MOBILE 6 requires VMT by vehicle class, facility type (freeway, arterial, local, and ramp), hour of the day, and speed. To achieve this, default values from past studies are used (EPA, 2001).

CONCLUSIONS

The objectives of this study were to identify sources that can assist in providing input data to MOBILE 6, develop procedures to prepare the input data for MOBILE 6, and document the procedure. The major source of information for input data to MOBILE 6 is the Highway Performance Monitoring Study (HPMS) data, and the procedures to use the data are described in a comprehensive set of documents prepared by the EPA. This document contains a summary of the information from these sources and draws together information from a number of sources.

In this study, several new procedures for acquiring data needed in the application of MOBILE 6 were tested. GPS instruments in vehicles were tested as a means of gathering VMT information and obtaining information on vehicle soaks. The method seems viable and may be an appropriate means of gathering data in conditions where default data values are unlikely to be representative. Another method tested in this study was the use of link characteristics to estimate VMT. The method was computerized and proved that it is a quick and convenient way to obtain VMT in an area and provide it in the form required by MOBILE 6.

REFERENCES

1. Arcadis, G.; Browning, L.; Chan, M.; and Miller, I.; *Update of Fleet Characterization Data for Use in MOBILE 6-Final Report*, Prepared for U.S. Environmental Protection Agency, May 1998. <<http://www.epa.gov/otaq/models/mobile6/m6flt002.pdf>> (accessed 04/15/2002).
2. Battelle, *Lexington Area Travel Data Collection Test Final Report*, A report submitted to Federal Highway Administration, Washington, D.C., September 1997.
3. Bhat, C.R.; and Nair, H.S.; *VMT Mix Modeling For Mobile Source Emissions Forecasting: Formulation and Empirical Application*, presented at the 80th Annual Meeting of the Transportation Research Board, Washington D.C., January 2001.
4. Environmental Protection Agency (EPA); *Section 187: VMT Forecasting and Tracking Guidance*, Report AMS-FRL-4115-8, EPA publication number 420R92008, January 1992.
5. Environmental Protections Agency (EPA); *Guidance for the Development of Facility Type VMT and Speed Distributions*, Technical Report M6.SPD.004, February 1999. <<http://www.epa.gov/otaq/models/mobile6/m6spd004.pdf>> (accessed 04/15/2002).
6. Environmental Protection Agency (EPA); *Soak Length Activity Factors for Hot Soak Emissions*, Technical Report M6.FLT.004, April 2001. <<http://www.epa.gov/otaq/models/mobile6/r01014.pdf>> (accessed 04/15/2002).
7. Environmental Protection Agency (EPA); *Development of Methodology for Estimating VMT Weighting by Facility Type*, Technical Report M6.SPD.003, April 2001.
8. Environmental Protection Agency (EPA); *Fleet Characterization Data for MOBILE 6: Development and Use of Age Distributions, Average Mileage Accumulation Rates and Projected Vehicle Counts for Use in MOBILE 6*, Technical Report M6.FLT.007, September 2001c. <<http://www.epa.gov/otaq/models/mobile6/r01047.pdf>> (accessed 04/15/2002).
9. Environmental Protection Agency (EPA); *Technical Guidance on the Use of MOBILE 6 for Emission Inventory Preparation*, January 2002. <<http://www.epa.gov/otaq/models/mobile6/m6techgd.pdf>> (accessed 04/15/2002).
10. Environmental Protection Agency (EPA); *Users Guide to MOBILE 6 -Mobile Source Emission Factor Model*, January 2002. <<http://www.epa.gov/otaq/models/mobile6/r02001.pdf>> (accessed 04-15-2002).

11. Everett, J.; and Sacs, I.; *An Alternative Methodology for Collecting Local Vehicle Start Information for use in MOBILE 6*, presented at the 80th Annual Meeting of the Transportation Research Board, Washington D.C., January 2001.
12. Federal Highway Administration (FHWA); *Highway Performance Monitoring Field Manual*, U.S. Department of Transportation, Washington D.C., December 2000.
13. Granell, J.L.; Guensler, R.L.; and Bachman, W.H.; *Using Locality Specific Fleet Distributions In Emissions Inventories: Current Practice Problems and Alternatives*, presented at the 80th Annual Meeting of the Transportation Research Board, Washington D.C., January 2001.
14. Jackson, R.T.; *Development and use of Age Distributions, Average Annual Mileage Accumulation Rates, and Projected Vehicle Counts for use in MOBILE 6*, EPA Technical Report M6.FLT.007, September 2001.
15. Miller, T.L.; Davis, W.T.; Reed, G.D.; and Doraiswamy, P.; *Effect of County Level Income on Vehicle Age Distribution and Emissions*, presented at the 80th Annual Meeting of the Transportation Research Board, Washington D.C., January 2001.
16. Pierce, B.; Casas, J.; and Giaimo, G.; *Estimating Trip Rate Under-Reporting: Preliminary Results from the Ohio Household Travel Survey*, presented at Transportation Research Board 82nd annual meeting, Washington, D.C., January 2003.
17. Qiao, F.; Yu, L.; Li, G.; and Wang, X.; *Modeling Vehicle Age distribution for Air Quality Analysis*, presented at the 80th Annual Meeting of the Transportation Research Board, Washington D.C., January 2001.
18. Sivakumar, A., *Implementing Soak Time Distribution Models In a GIS Framework*, Term Project for CE394k.3 GIS in Water Resources, Fall 2000, University of Texas at Austin.
19. Stopher, P.; Bullock, P.J.; and Horst, F.; *Conducting GPS Survey with a Time-Use Diary*, presented at Transportation Research Board 82nd annual meeting, Washington, D.C., January 2003.
20. TMIP; *Workshop on Transportation and Air Quality Analysis*, Publication no. FHWA-141-94-011, September 1994.
<<http://tmip.fhwa.dot.gov/clearinghouse/docs/airquality/wtaq/>> (accessed 05/30/2002).

21. Vadlamani, S.; *Development of an Alternative Approach to Estimate Vehicle Miles Traveled Using a Classification Procedure*, M.S. Thesis, Department of Civil and Environmental Engineering, Louisiana State University, May 2005.
22. Varanasi, S.; *Assessing the Feasibility of Supplying Vehicle Activity Data to MOBILE 6 Using the Global Positioning System (GPS)*, M.S. Thesis, Department of Civil and Environmental Engineering, Louisiana State University, December 2003.
23. Venigalla, M.M.; and Pickerell, D.H.; *Soak Distribution Inputs to Mobile Source Emissions Modelling: Measurement and Transferability*, presented at Transportation Research Board 81st annual meeting, Washington, D.C., January 2002.
24. Wolf, J.; Oliveira, M.; and Thompson, M.; *The Impact of Trip Underreporting on VMT and Travel Time Estimates: Preliminary Findings from the California Statewide Household Travel Survey GPS Study*, presented at Transportation Research Board 82nd annual meeting, Washington, D.C., January 2003.