GPS-TDG

User's Manual

Table of Contents

1. Introduction

For nearly fifty years, household travel surveys have been used to document the travel behavior of regional households as part of long-range transportation planning efforts. The survey data are used for general planning and policy analysis, as well as to serve as the foundation for regional travel demand models. Technology advancements have resulted in changes in household travel survey data collection procedures, the most recent being the introduction of Global Positioning Systems (GPS) to record travel patterns. The GPS technology shows promise to minimize costs, while maximizing the volume of travel data collected. However, the data recorded by GPS devices do not directly yield travel information; rather, the outputs from these devices are in the form of navigational streams that have to be processed to derive travel information. Therefore, the success of this new technology as a travel survey instrument depends on the ability of the analyst to derive meaningful trip information from the navigational data streams of GPS devices.

Research project 0-5176 funded by TxDOT focused on the development of a prototype software tool labeled the "GPS-Based Travel Diary Generator" (GPS-TDG) that automates the process of converting navigational data streams collected passively from in-vehicle GPS devices into an electronic travel diary. This derived travel diary comprises a sequence of vehicle trips identified from the GPS streams, with each trip characterized in terms of attributes such as trip-end location, trip purpose (or activity type at destination), time of day, duration, distance, and speed.

The software has been developed in the Java programming language using ArcGIS 9.0 as the platform for GIS processing. The software has been designed to operate either in a basic analysis mode or in an enhanced analysis mode. The basic mode converts the GPS data into a simple trip file that distinguishes among home-based work, home-based other, and non-home-based trips. The enhanced mode utilizes additional land-use data and pre-estimated model parameters to derive more refined trip purpose classification taxonomy.

Finally, the software is also capable of aggregating the derived trip diaries to produce inter-zonal vehicle trip tables and network performance measures (average trip speed, distance, and travel time). These measures can be generated in the overall or for specific trip purposes or for specific times-of-the-day.

Detailed descriptions of the state-of-the-art in GPS-based travel surveys, the algorithms incorporated within the GPS-TDG, software implementation details, and testing and validation results are provided in three research reports produced as part of the project. This document is developed as an instructional guide to help analysts use the GPS-TDG software. The rest of this manual is organized as follows. Chapter 2 describes the system requirements and installation instructions. Chapter 3 provides an overview of the GPS-TDG environment including a long name of the menus and commands available. Chapter 4 discusses inputs and outputs. Finally, Chapter 5 guides the user through the basic operations using sample data files provided with the software. This chapter is also intended as a simple tutorial to help analysts familiarize themselves with the software and to interpret outputs.

2. System Requirements and Installation Instructions

The software describes the minimum system requirements for running the software and instructions for installation. The system requirements are presented in Section 2.1 and the installation procedure is outlined in detail in Section 2.2

2.1. MINIMUM SYSTEM REQUIREMENTS

The software has been developed in Java and is therefore portable across platforms. However, we recommend the following system configuration for which the software has been tested:

- Windows XP operating system
- 2 GHz CPU
- 512MB RAM for the basic mode and 1GB RAM for the enhanced mode

In addition to the above system requirements, the user should also have the following software installed:

- ArcGIS 9.0 (either an ArcEngine runtime license or an ArcView license)
- Microsoft Office 2003 or higher
- Java Runtime Engine 1.4.2 or higher.

2.2. INSTALLATION

The installation of the software comprises three main steps. First, the software provided as a self-extractable file has to be copied onto the computer and unzipped (Section 2.2.1). Next, appropriate environment variables have to be defined (Section 2.2.2). Finally, a Data Source Name (DSN) has to be created (Section 2.2.3).

2.2.1 The Self-Extractable File

The software is provided in a CDROM as a self-extractable zip file (GPSTDG.EXE). Copy this file onto the C drive of your computer and unzip it. This creates the C:\GPSTDG directory in which the software is installed. The C:\GPSTDG directory will contain the following files and folders:

1. The GPSTDG.jar is the java executable file. Running this executable runs the software.
2. The GPSTDG.desc is a short file describing the functionality of the ".jar" file.
3. The GPSTDG.bat file contains the command line code to run the executable file. The user runs the software by simply double clicking the GPSTDG.bat file.
4. A copy of the GPSTDG user's manual (UserManual.pdf)

5. The GPSTDG.options file which contains the algorithm parameters (discussed further in Section 4.2.2)
6. The ModelParameters.csv file which has the empirical specification for the multinomial logit model estimated for activity-type classification using data from Laredo (discussed further in Section 4.2.1).
7. The C:\GPSTDG\java folder containing several java-related sub-folders and files.
8. The C:\GPSTDG\Laredo folder containing sample data files for running the software. These files are discussed further in Chapter 5.
9. The C:\GPSTDG\Trips folder would be the default folder to which the generated trip files will be written out (Also see section 2.2.3). This folder also contains outputs corresponding to the above inputs and using the algorithm parameters from GPSTDG.options file and the MNL model from the ModelParameters.csv file. These output files are discussed further in Chapter 5.
10. The C:\GPSTDG\gprmctest folder which contains sample GPS streams data in the GPRMC format.

Important Note:
    Subsequent instructions for creating environment variables and DSN are provided assuming that the software is installed in folder C:\GPSTDG. If the software is being installed to a different location, these instructions have to be suitably modified. In addition, the user will have to edit the GPSTDG.bat file (right click on the file and select Edit) and add the path to the trip file location to the end of the existing command. For example, if the path to the trips file location is to be F:\GPSTDG\trips, the command needs to be modified to "javaw -jar GPSTDG.jar   F:\GPSTDG\trips". By default, the command on the ".bat" file reads "javaw -jar GPSTDG.jar".
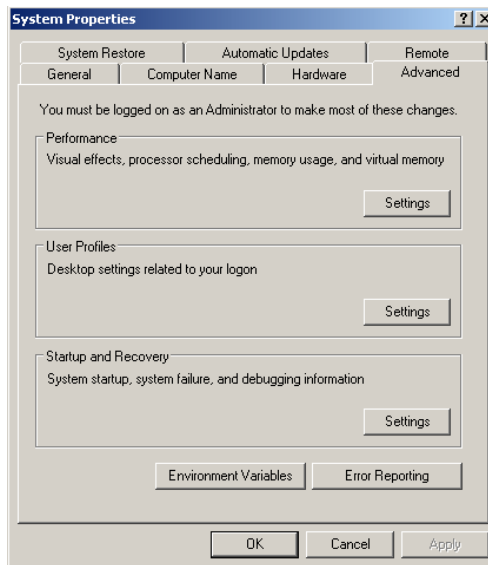
2.2.2 Defining Environment Variables

    Prior to running the software, it is necessary to create or modify three environment variables. These variables are "ARCENGINEHOME", "CLASSPATH", and "PATH". Detailed instructions for defining these variables are presented below.

Right click on *My Computer* and select *Properties*. The *System Properties* window opens.
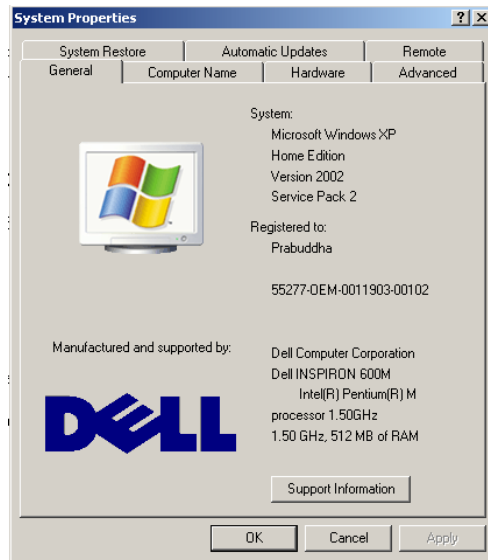

Click on the *Advanced* tab

Click on the *Environment Variables* button. The ARCENGINEHOME variable may be present under System variables with the path to where ArcGIS is installed on the comouter (as shown in figure). If not, click the *New* button under System variables. The *New System Variable* window pops up. Enter ARCENGINEHOME as the *variable name* and the path to where ArcGIS is installed as the *variable value* (for example, C:\Program Files\ArcGIS). Click OK to save the changes.
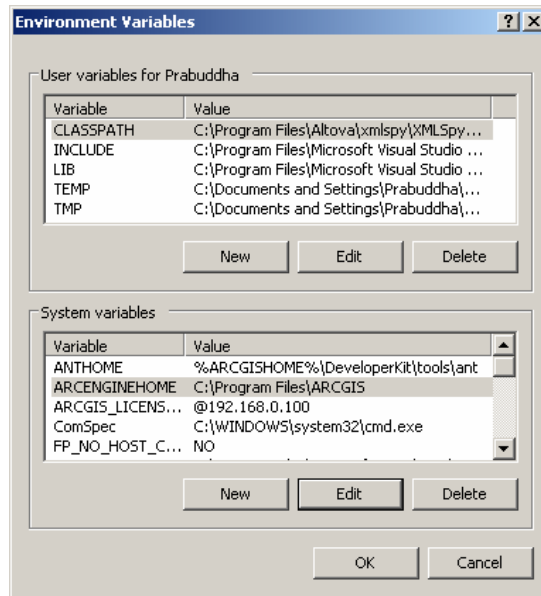
If the variable CLASSPATH exists under User variables, highlight this variable and click *Edit*. The *Edit User Variable* window pops up. To the variable value, add the path to the GPSTDG software, i.e., enter "C:\GPSTDG\;" at the beginning of the text already in this text box. Click *OK* to save the changes. If the CLASSPATH variable does not exist, click the *New* button under User variables. The *New System Variable* window pops up. Enter CLASSPATH as the *variable name* and the path to where GPSTDG is installed as the *variable value* (i.e., C:\GPSTDG\). Click *OK* to save the changes.
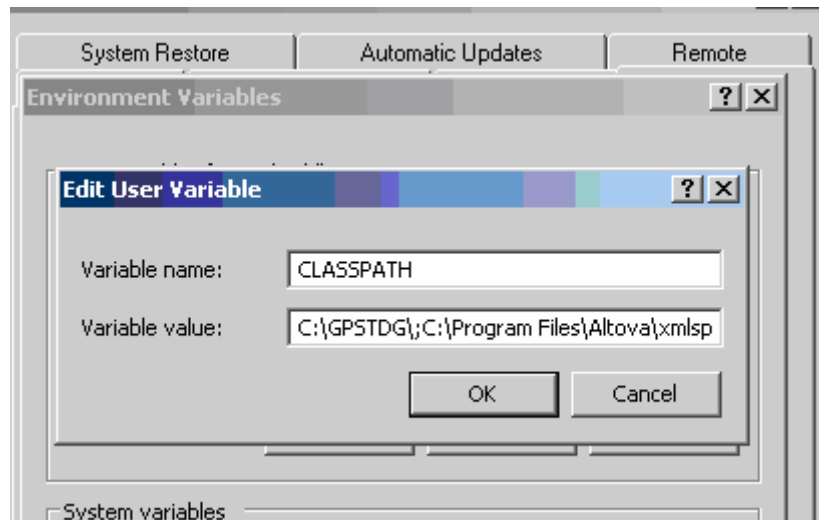

The variable PATH exits under System variables. Highlight this variable and click *Edit*. The *Edit User Variable* window pops up. In the variable value text box, enter "C:\GPSTDG\java\jre\bin;" at the beginning of the text already in this text box. Click *OK* to save the changes.
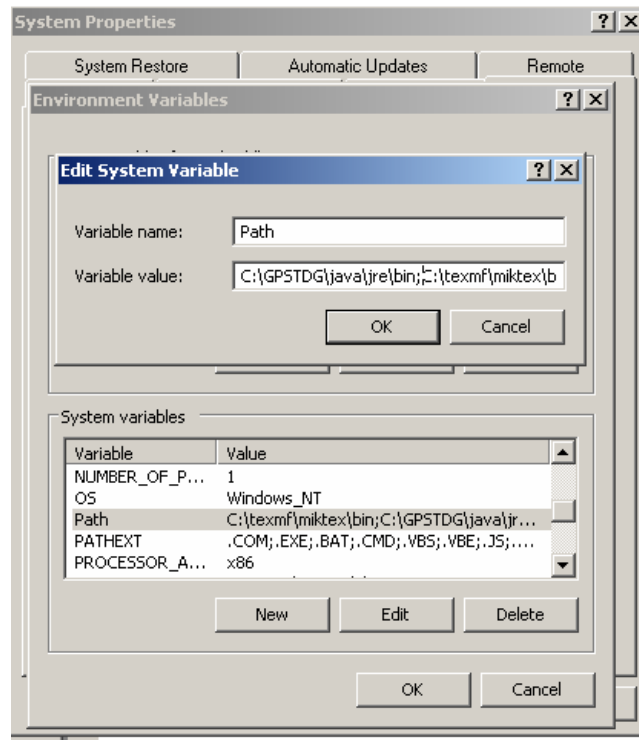
## 2.2.3 Creating the DSN (Data Source Name)

The software requires a DSN to exist in order to perform query operations. To create this DSN, first open the *Control Panel* from the *Start Menu*. Next, open the *Administrative Tools* folder (double-click on the administrative tools icon). Now double click on the *Data Sources(ODBC)* icon. This opens the *ODBC Data Source Administrator* window as shown below:
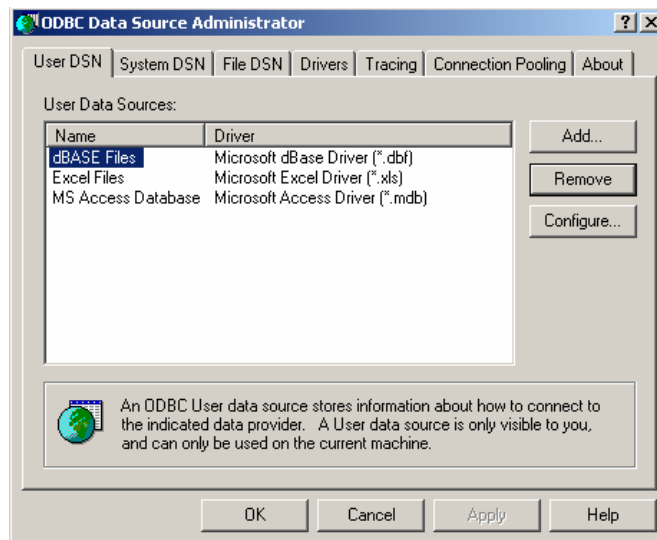
Click on the *Add* button. The *Create New Data Source* window opens up. Scroll down the list and select the as Microsoft Text Driver and Click *Finish*.
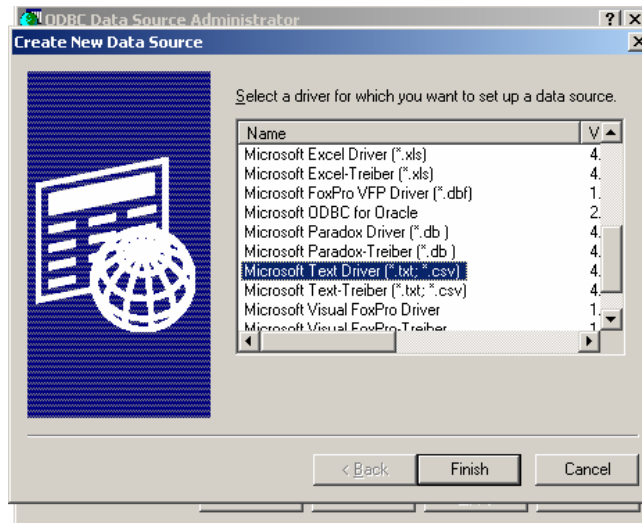
The *ODBC Text Setup* window opens up. Enter TripsDSN as the *Data Source Name*.
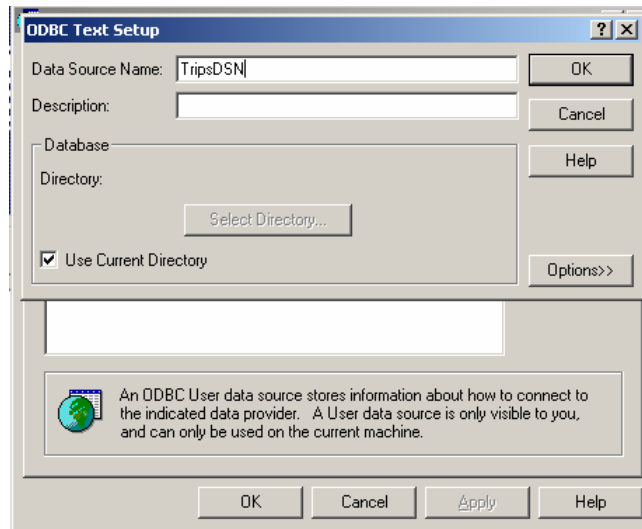
Uncheck the *Use Current Directory* checkbox and click *Select Directory*. The *Select Directory* window opens up. Navigate to C:\GPSTDG\Trips and click *OK*. Click *OK* to save all changes made.
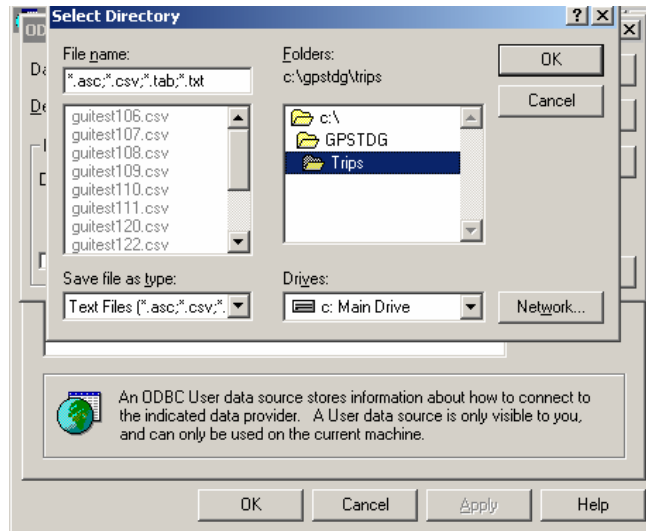
The *ODBC Data Source Administrator* window now appears as shown below:

3. The GPS-TDG Environment

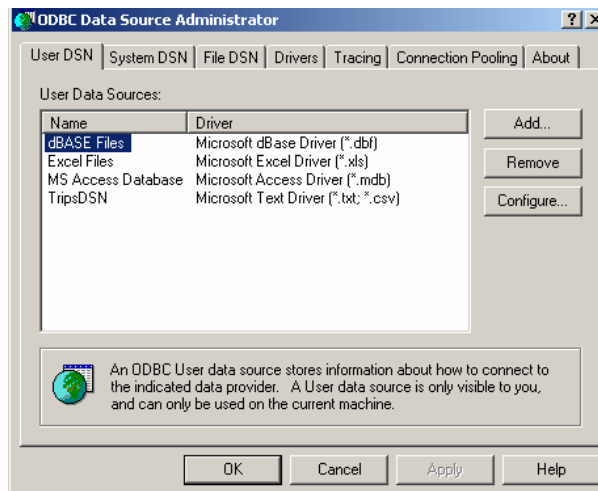The GPS-TDG provides a swing-based graphical interface for the user to interact with the software to derive the travel patterns from GPS streams and to query the derived travel diaries in many ways. This chapter describes the basic elements of the software environment.

To run GPS_TDG, double click on the DOS batch file "GPSTDG.bat". This opens up the main window of the program which remains on screen as long as the software is being used. This main window comprises: (1) The Menu Bar, (2) The Command Area, and (3) The Data Area. Each of these is described in detail in Sections 3.1, 3.2, and 3.3, respectively.
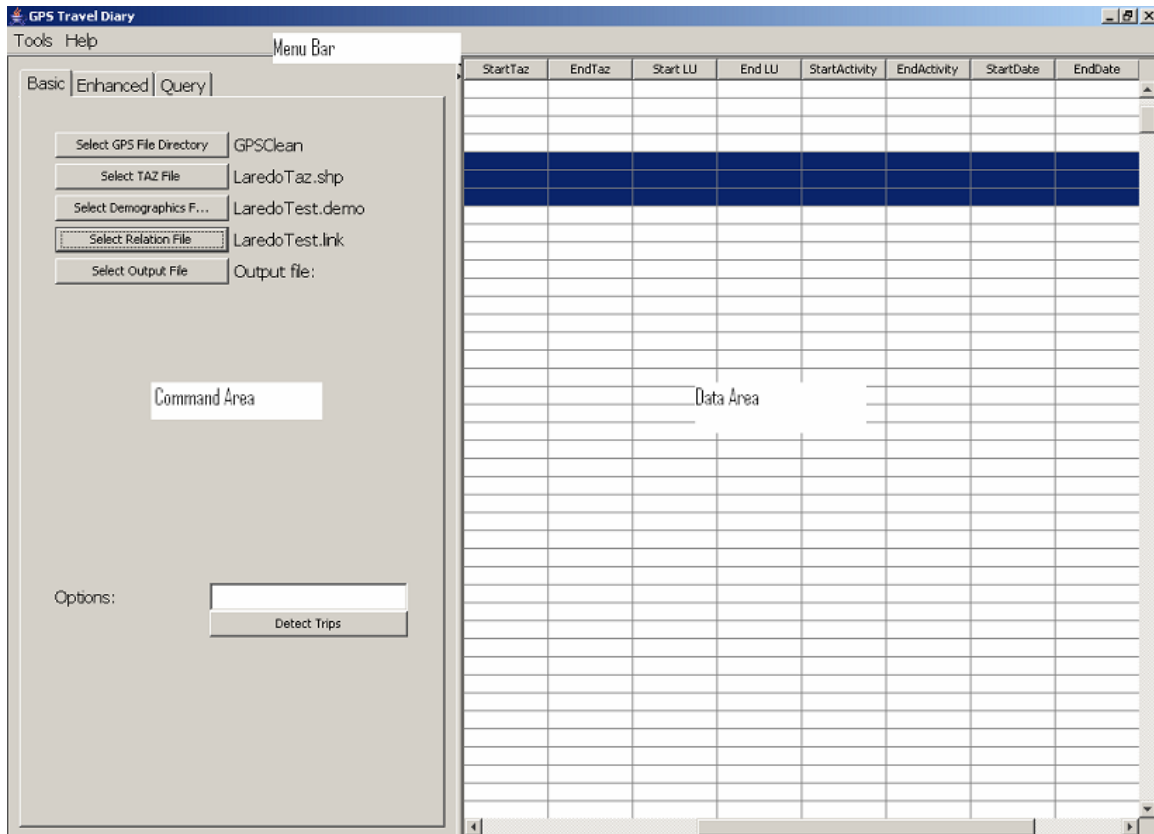
3.1. THE MENU BAR

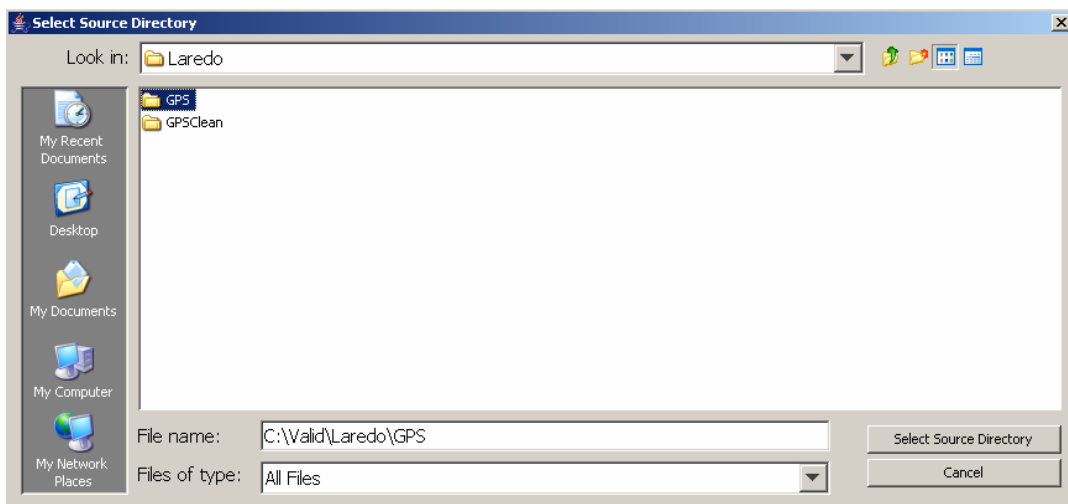The Menu Bar comprises two menu items, (1) Tools, and (2) Help. Each of these items is discussed in detail subsequently.

3.1.1 The Tools Menu

The tools menu allows the user to preprocess the raw GPS streams and convert it into a standard format for further use by the GPS-TDG software. This pre-processing is a requirement and will be the first overall step in using the software for deriving the travel patterns. The software is capable of taking as inputs the raw GPS streams that are in either the GPRMC or in the GeoLogger formats. Correspondingly, the user chooses either *Preprocess GPRMC* or *Preprocess GEO* from the drop-down Tools menu.
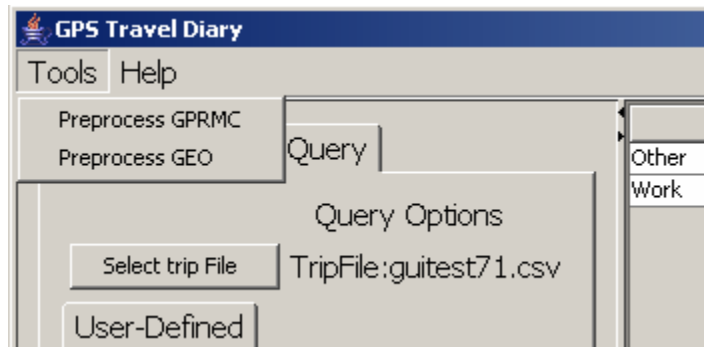
This opens the *Select Source Directory* window which allows the user to specify the directory containing the raw GPS files. This input directory should contain one file per vehicle. Files containing the GPS streams in the GeoLogger format should have a ".dlv" extension. Files containing the GPS streams in the GPRMC format should have a ".gprmc" extension. The detailed structures of these input files are presented in Section 4.1.1. The user can navigate to the directory containing the raw GPS streams and click the *Select Source Directory* button to select the files.

Once the Source Directory is selected (C:\Valid\Laredo\GPS as above), the *Select Target Directory* window opens and the user is prompted to select the sub-directory in which the pre-processed GPS files will be stored. Again, the user can navigate and select the desired output directory by clicking *Select Target Directory* (C:\Valid\Laredo\GPSClean as below).

Once the input and output directories are chosen, the pre-processor runs in the background. A window pops up to indicate completion of pre-processing.

The processing time depends on the number and size of the input files. One output file is generated for each file input (i.e., one for each GPS-equipped vehicle). The output

files have a ".cln" extension. The detailed structure of this preprocessed, output GPS stream file is presented in Section 4.3.1

3.1.2 The Help Menu

The Help Menu has two items. Choosing, the *About* menu item displays the software version and author information. Choosing, the *Online Help* item opens the web browser to the help web page:
http://www.ce.utexas.edu/prof/bhat/GPSTDG/UserManual.pdf

3.2. THE COMMAND AREA

The Command Area provides the fundamental interface for the user to run the software to derive the trips from pre-processed GPS streams. Further, the trip detection can be performed at two levels of sophistication, the basic and the enhanced modes. In the basic mode, the trip-end activity types are classified into "home", "work", and "other". In the enhanced mode, activity purposes classified in the basic mode as "other" are in turn classified into finer categories such as "grocery shopping", "other shopping", and "social". In addition to the trip file generation, the Command Area also provides the user with tools to query the derived trip file in many ways.

The Command Area is organized into three panes, (1) The Basic Pane, (2) The Enhanced Pane, and (3) The Query Pane. Each of these is discussed below.

3.2.1 The Basic Pane

The Basic Pane of the Command Area allows the user to specify inputs, run the trip detection in the basic mode, and view the derived trip file. The Basic Pane comprises several buttons to specify the input data files, a text box to specify optional parameters to control the trip detection algorithm, and a button to run the trip detection algorithm and view the generated trip file. The chosen folders and files containing the input data are also displayed adjacent to the corresponding buttons in the Basic Pane (for example, the figure below indicates that LaredoTaz.shp is the input TAZ file). Each of the buttons and the Options text box is discussed further below:
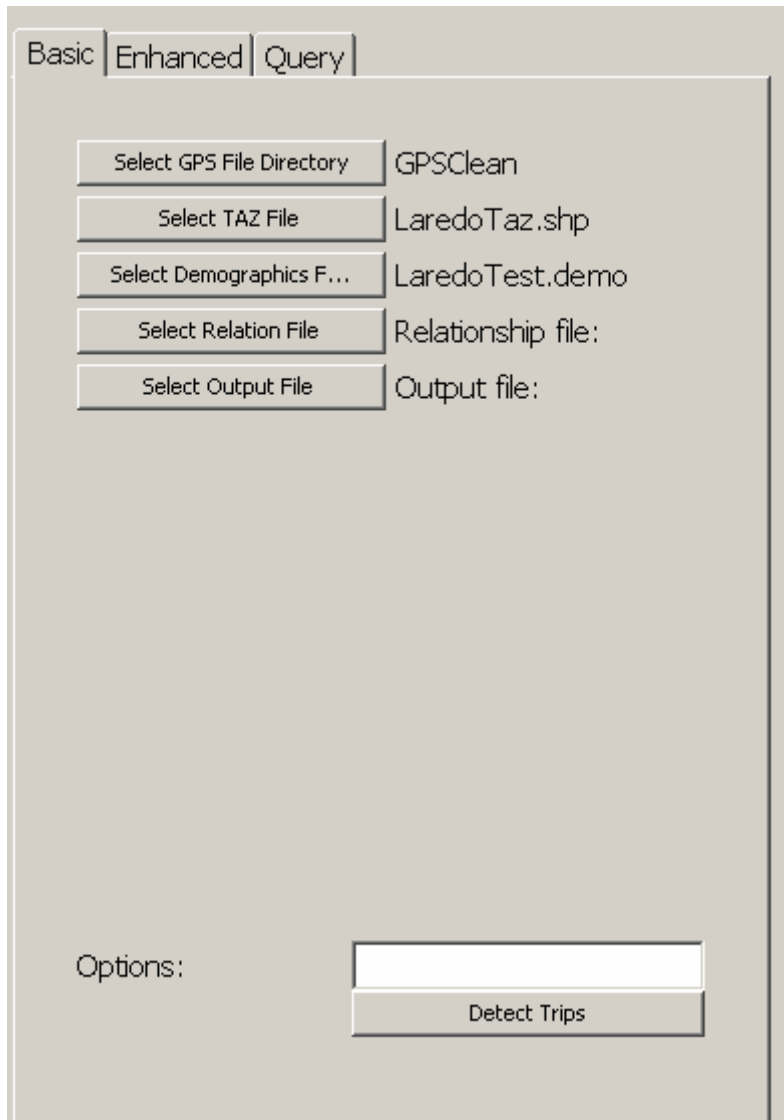
Clicking the *Select GPS File Directory* button opens the *Select GPS Directory* window. The user can navigate to the directory containing the pre-processed GPS streams and click the *Select GPS Directory* button to select the files. This directory (C:\GPSTDG\Laredo\GPSClean below) should contain comma-separated files output from the preprocessor and with ".cln" extensions. The structure of these files is discussed in Section 4.3.1.
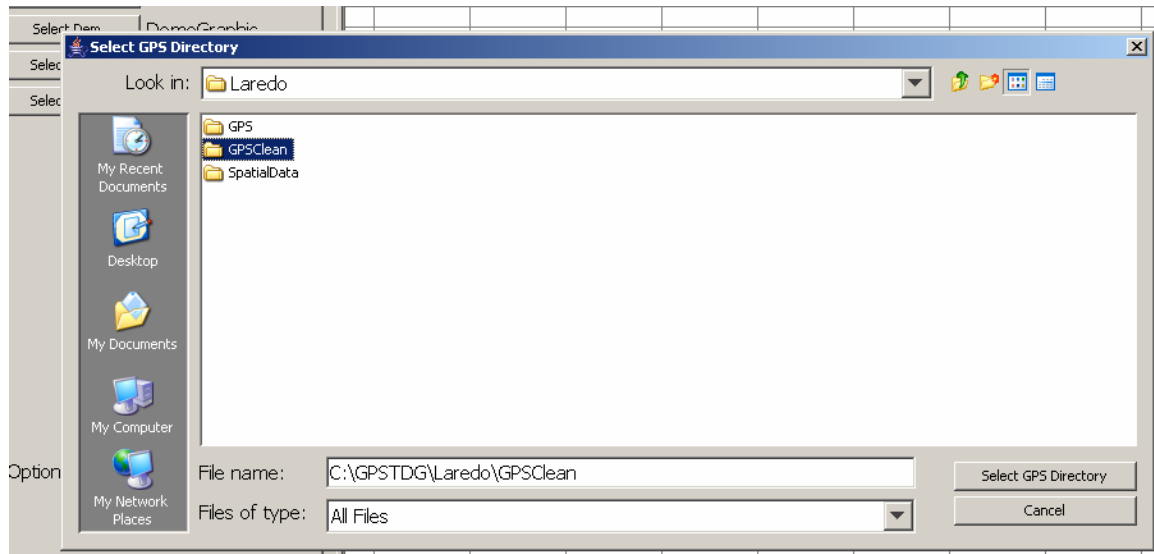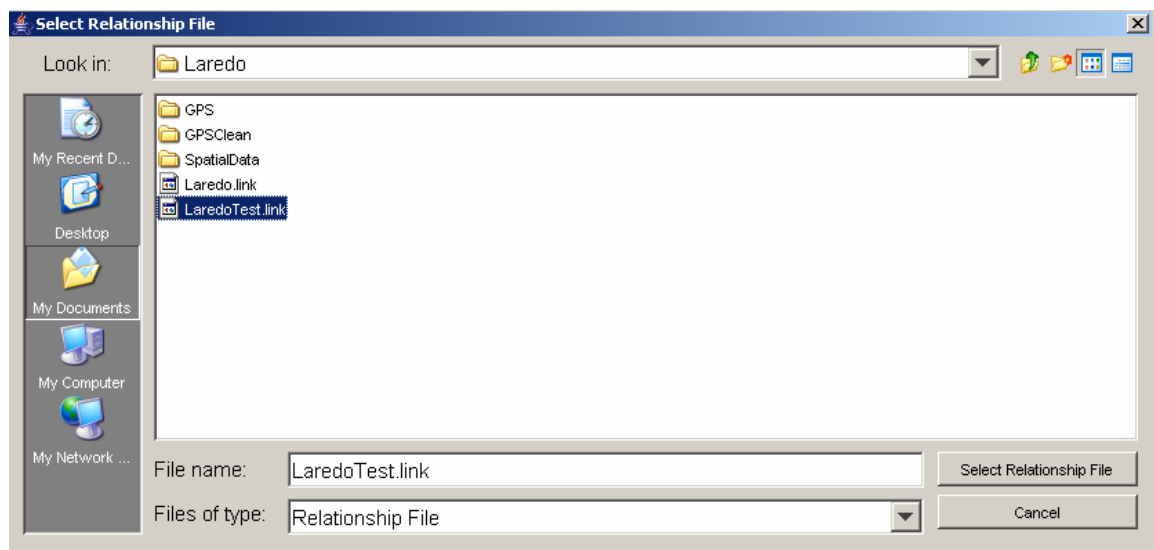
Clicking the *Select TAZ File* button opens the *Select TAZ File* window. The user can navigate to the GIS file containing the boundaries of the traffic analysis zones (TAZs) of the study area and click the *Select TAZ File* button to select the file (the TAZ file is the LaredoTaz.shp below). This GIS file is provided in the ESRI shape file format with a ".shp" extension. Specifications of this file are discussed in Section 4.1.3. The user should note that trip detection can be performed without providing this input. In this case, the software will not determine the trip end locations in terms of TAZs.
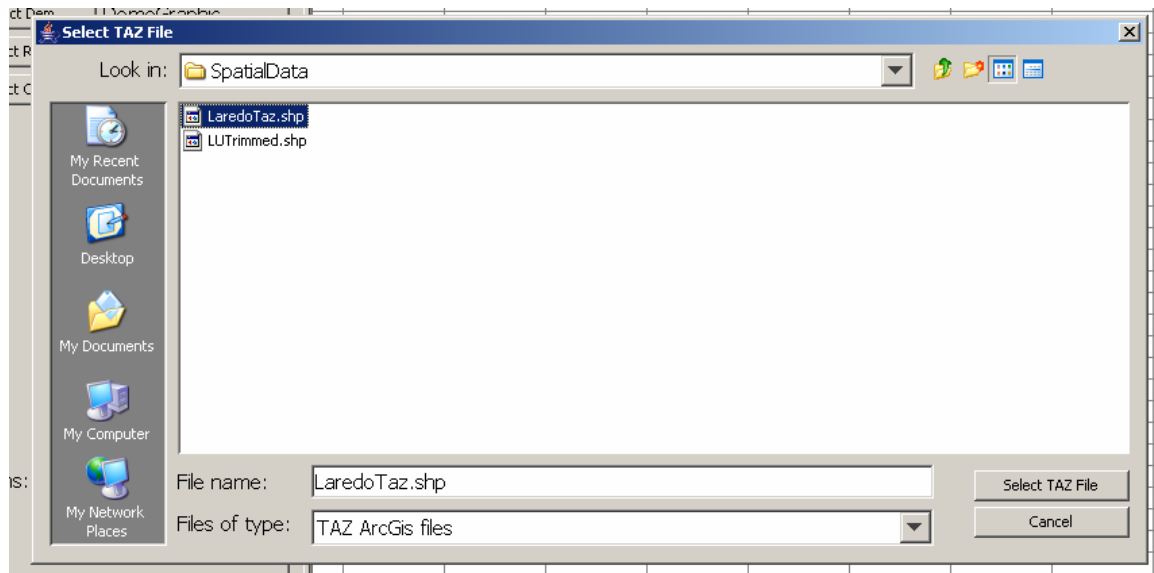

Clicking the *Select Demographics File* button opens the *Select Demographics File* window. The user can navigate to the file containing the household and person-level characteristics of the survey respondents and click the *Select Demographics File* button to select the file. This is a comma-separated file with a ".demo" extension. The detailed structure of this file is discussed in Section 4.1.2.
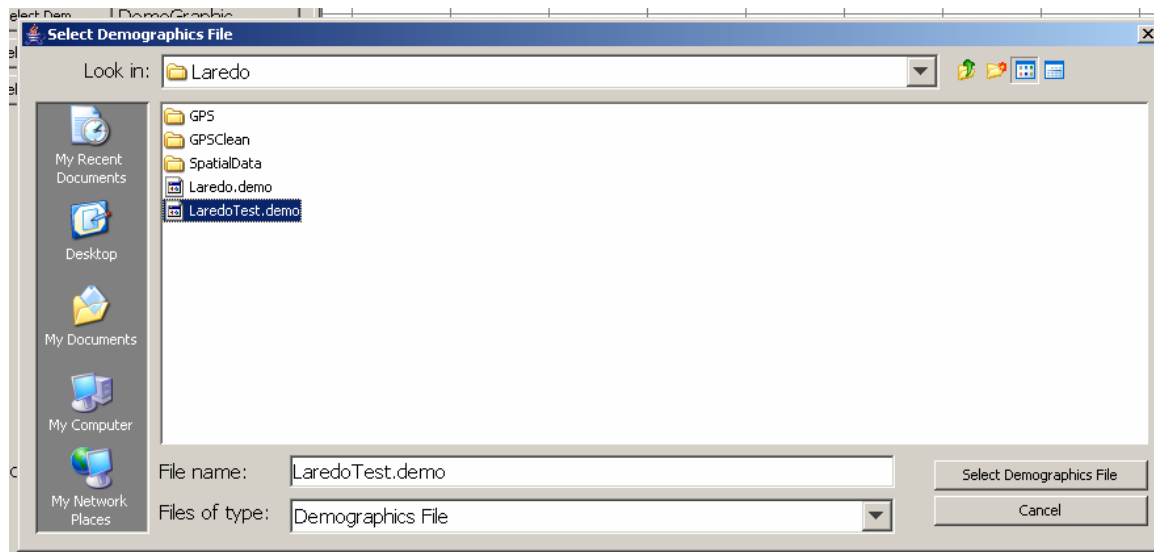
Clicking the *Select Relation File* button opens the *Select Relationship File* window. The user can navigate to the file that associates (or "links") each vehicle with one member of household called the primary driver. Click the *Select Relationship File* button to select the file. This is a comma-separated file with a ".link" extension. The detailed structure of this file is discussed in Section 4.1.2.

Clicking the *Select Output File* button opens the *Select Output File* window and allows the user to specify the output file name. This will be created by default in C:\GPSTDG\Trips directory. (If you have set up the software differently, the file is created at the location specified in the DSN.) Two files with the specified file name and extensions of ".trp" and ".csv" are written out. Both files contain the same data, but differ in their file structure and formats. The structure of these two output files are discussed in Section 4.3.2.
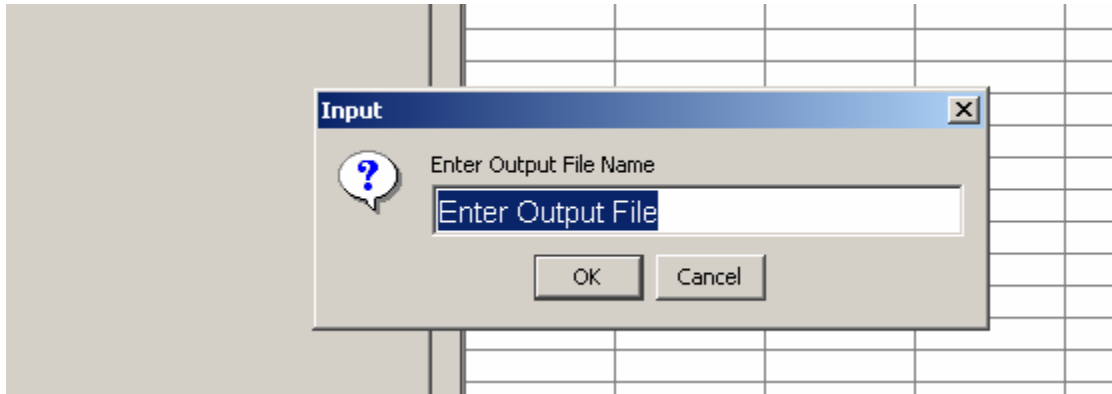
The text box *Options* can be used by the analyst to modify several parameters that control the trip detection and characterization algorithm implemented within the software. A description of these parameters, and the default values are presented in Section 4.2.2. If this box is left blank, the default parameters (provided in the GPSTDG.options file) are used. The user can modify selected parameters by entering the parameter flag and value separated by space into the options box. For example, the analyst can change both the engine power off and non engine power off dwell time thresholds to 45 seconds, by entering "-t 45 -t2 45" in the options box. The complete list of flags associated with the parameters is presented in Section 4.2.2. The analyst can also modify these parameters by making changes to the underlying GPSTDG.options file. It should be noted that the values input via the text box are not saved to the options file. Therefore, the analyst can conveniently use the text box for the purposes of testing parameter values and finally incorporate the preferred values in the GPSTDG.options file for all subsequent analyses.

Clicking the *Detect Trips* button runs the trip detection and characterization algorithm implemented within the software. A Progress Bar pops up which displays the name of the file containing the preprocessed GPS streams of the vehicle currently being processed and the percentage of all vehicles that have been processed (see section 3.2.2. on Enhanced Pane for a progress-bar screen shot). On completion of the run, the ".trp" and ".csv" output files (as discussed above) are generated and the ".csv" file is loaded automatically into the data area. The structure of this spreadsheet is discussed in Section 3.3

3.2.2 The Enhanced Pane

The Enhanced Pane of the Command Area allows the user to specify inputs, run the trip detection in the enhanced mode, and view the derived trip file. This pane has all the elements of the basic pane, and two additional buttons to specify the additional inputs required for the enhanced analysis. The two additional buttons are discussed further here. As in the case of the Basic Pane, the chosen folders and files containing the input data are also displayed adjacent to the corresponding buttons in the Enhanced Pane.
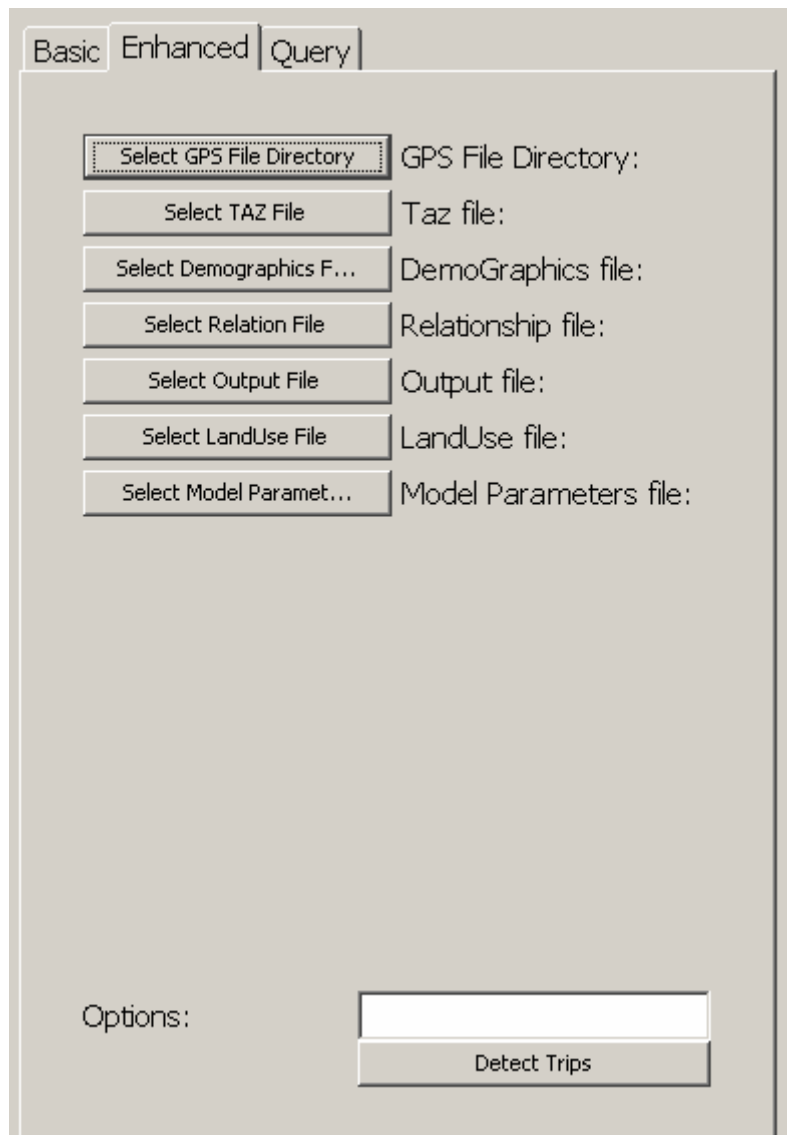
Clicking the *Select LandUse File* button opens the *Select LandUse* window. The user can navigate to the GIS file containing the land use of the study area. Click the *Select LandUse* button to select the file. This GIS file is provided in the ESRI shape file format with a ".shp" extension. Detailed specifications of this file are discussed in Section 4.1.4.
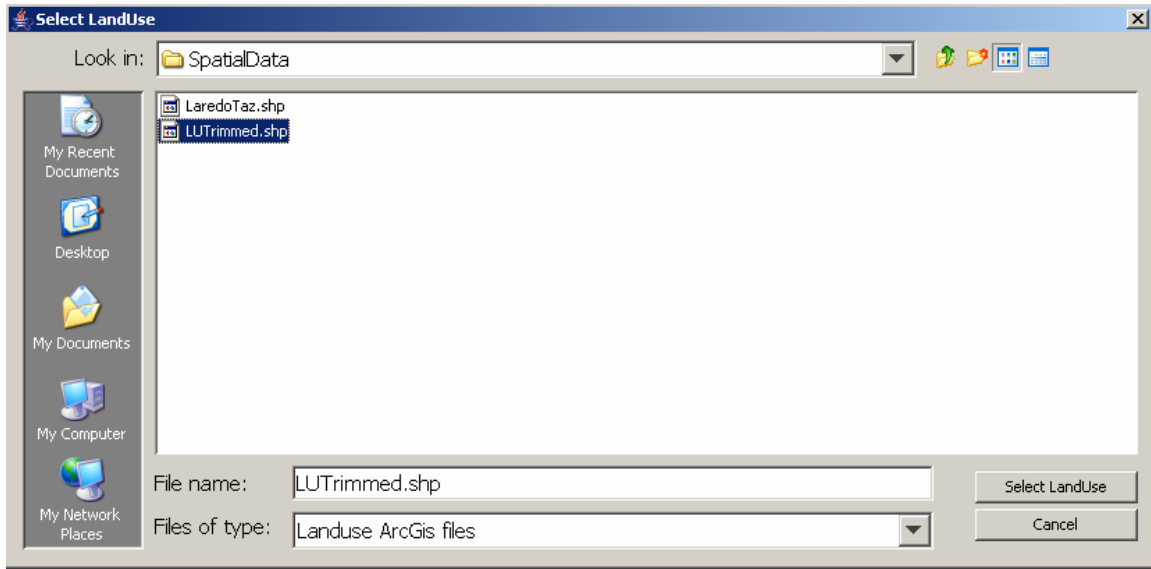
Clicking the *Select Model Parameters File* button opens the *Select Model Parameters* window. The user can navigate to select the empirical specification file for the disaggregate activity-type classification model and click the *Select Model Parameters* button to select the file. This is comma-separated file containing the disaggregate activity-types into which the "other" activities are to be classified and the parameters of a multinomial logit model which will be used by the algorithm to determine the activity purpose probabilistically. Detailed specifications of this file are discussed in Section 4.2.1. Ideally, the analyst should estimate this model using local data from the study region and provide the parameters as inputs. However, if such an exercise cannot be undertaken, the default parameters provided with the software (C:\GPSTDG\ ModelParameters.csv) may be used. These defaults are also presented in Section 4.2.1. The detailed procedure for estimating these model parameters have been presented in Research Report 2.
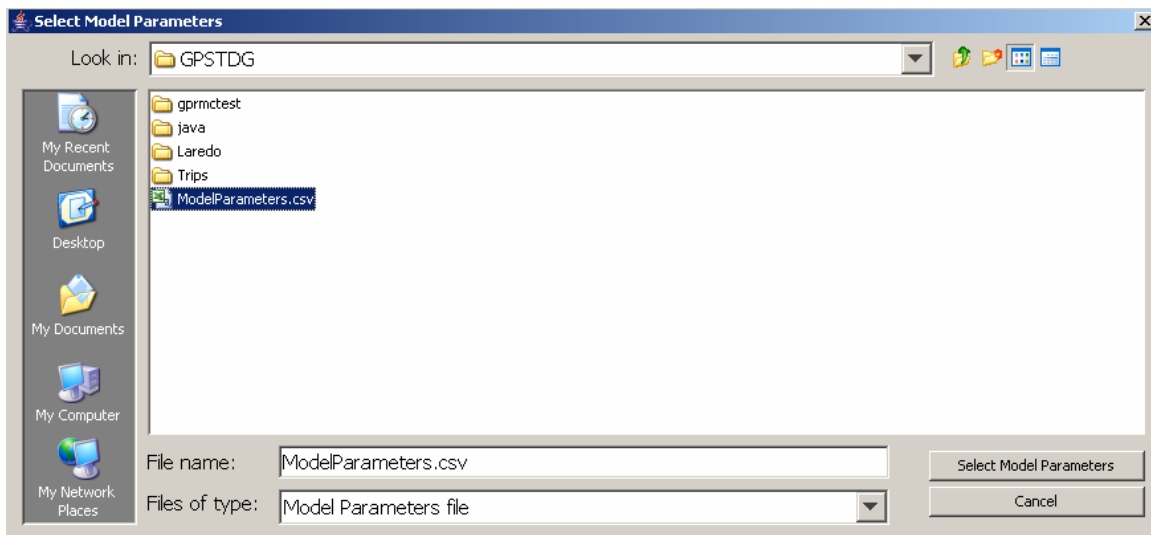
As in the case of the Basic Pane, clicking the *Detect Trips* button runs the trip detection and characterization algorithm implemented within the software. A Progress Bar pops up which displays the name of the file containing the preprocessed GPS streams of the vehicle currently being processed and the percentage of all vehicles that have been processed. On completion of the run, the ".trp" and ".csv" output files (as discussed above) are generated and the ".csv" file is loaded automatically into the data area. The structure of this spreadsheet is discussed in Section 3.3.
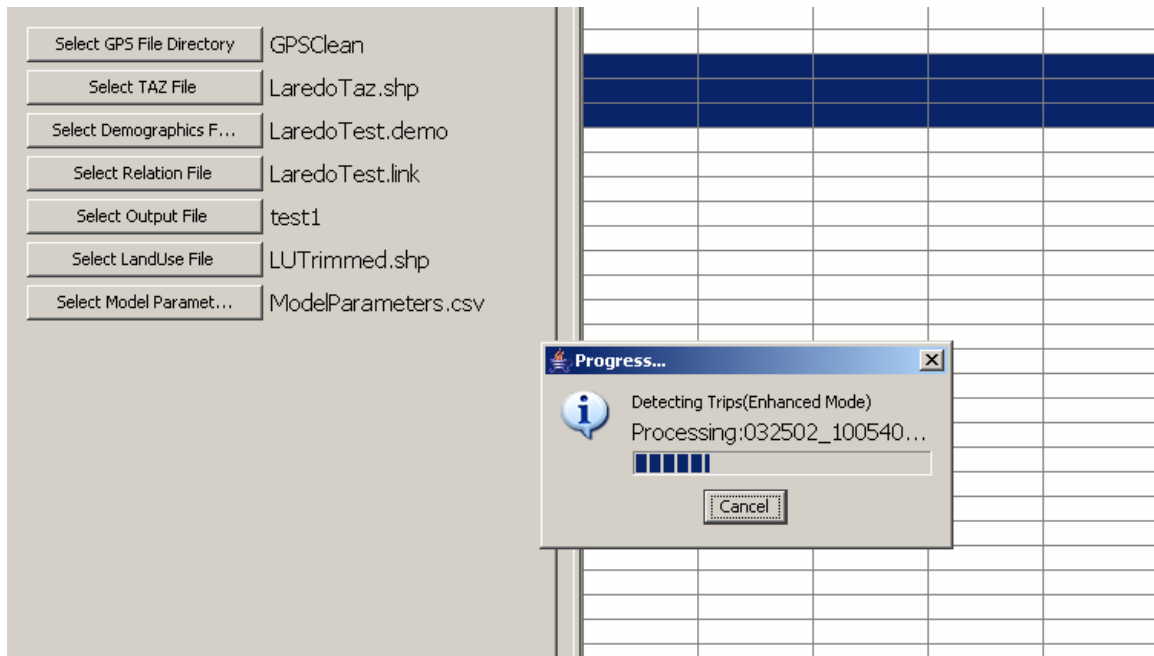
3.2.3 The Query Pane

The Query Pane allows the user to aggregate and analyze the derived trip file in many ways. First, by clicking on the *Select Trip File* button, the user selects a trip file in the .csv format which has been created using the GPS-TDG software (i.e., a file that was generated by clicking "Detect Trips" in either the Basic or Enhanced Pane). The name of the chosen file is displayed adjacent to the button.

The trips in this file can then be aggregated either by trip-end locations and/or trip end activity types. This is done by selecting the appropriate check boxes under the 'Aggregate By" section. The aggregate measures computed are the number of trips, average trip length, average trip duration, and average trip speed. Further, the aggregation can be done over the entire trip file or over a subset of trips satisfying specific criteria. These subsets can be defined by selecting the appropriate boxes under the 'Select By" section. The "Aggregate By" and "Select By" options provided to the user are discussed further here.

The "Aggregate By" Options:

The software allows the user to generate aggregate measures by trip-end locations and/or activity types (these activity types are aggregate if only the basic trip detection analysis was performed and disaggregate if the enhanced trip detection analysis was performed) For example, if the user checks the *By starting TAZ* box, the aggregate measures are computed over trips originating in each TAZ. Hence the output of this query will contain one row of measures (i.e., number of trips, average trip length, average trip duration, and average trip speed) for each TAZ. By checking both *By starting TAZ* and *By ending TAZ* boxes, we get the OD trip table and the inter-zonal level of service measures. If none of the "Aggregate By" boxes are checked (and no "Select By" options are chosen either), we get one row of output, which represent the total number of trips in the file and trip length, duration, and speed averaged over the entire trip file.

The "Select By" Options:

As already discussed, the 'Select By" option allows the user to perform the aggregations over a subset of trips satisfying particular criteria as opposed to the entire trip file. If none of the 'Select By" options are checked, the aggregation is performed over the entire trip file. The 'Select By" options are further categorized and provided under the "user-defined" and "Pre-Defined" tabs.

Using the options under the "User-Defined" tab, the analyst can define a subset of trips for aggregation based on <u>one or more</u> of the following criteria: trip-end location, trip timing, and trip-end activity types. The criteria are chosen by checking the appropriate boxes and entering the required values. For example, to define a subset of trips ending at work, check the *By End Activity* box and choose "Work" as the value from the adjacent drop-down list. The drop-down list is editable and therefore the analyst can also enter disaggregate activity types as appropriate. To define a specific time period for aggregation, (for example, the PM peak period of 3:30 to 6:30 PM), first check the *By start time* box and enter the value (i.e., 15:30) in the editable list adjacent to the check box. Next, check the *By end time* box and enter the value (i.e., 18:30) in the editable list adjacent to the check box. The aggregation is then performed over trips that fall within this period (i.e., trips that start at or after 3:30 PM and end at or before 6:30 PM)

The analyst can also specify selection criteria from the "Pre-Defined" tab. In this case, the user chooses one of the three commonly used trip purposes (i.e., home-based work, home-based other, and non home-based) as the selection criterion.

Once the user has specified the "Aggregate By" and "Select By" options, clicking on the button, *Execute Query* runs the query and the aggregate measures generated (i.e., number of trips, average trip length, average trip duration, and average trip speed) are displayed in the Data Area of the main window. The structure of this spreadsheet is discussed in Section 3.3.

3.3. THE DATA AREA

The Data Area is used to display the outputs generated (in a spreadsheet format) when either the *Detect Trips* or the *Execute Query* button is clicked. As already discussed, the *Detect Trips* generates the derived trips as output. The structure of the output displayed in this case is described in Section 3.3.1. The *Execute Query* generates aggregate measures such as number of trips and average trip speeds as the output. The structure of the output displayed in this case is described in Section 3.3.2.

3.3.1. Trips Display

The trips are displayed in the conventional trip-diary format. Each row of display corresponds to a single vehicle trip. Although, the software generates several attributes for each trip (See section 4.3.2 for the complete details), only the most important and commonly used characteristics are displayed in the Data Area.

The displayed items are listed and briefly described below. The analyst can view all the attributes generated by opening the .csv output file (described in detail in Section 4.3.2) in standard spreadsheet packages like MS Excel.

Column	Col. Name	Description
1	VehID	Identifier for GPS equipped vehicle (HHVehID in the corresponding .trp file)
2	DriverID	Identifier for the primary driver of the equipped vehicle
3	OriginalTripNo	Trip number assigned originally to all potential trips
4	TripNo	Final trip number assigned
5	TripDuration	Trip duration in seconds
6	StartTAZ	TAZ identifier for the trip-start location
7	EndTAZ	TAZ identifier for the trip-end location
8	StartLU	Land-use at the trip-start location
9	EndLU	Land-use at the trip-end location
10	StartActivity	Activity type pursued by the driver of the vehicle at the origin of the trip
11	EndActivity	Activity type pursued by the driver of the vehicle at the destination of the trip
12	StartDate	Date and time of the start of the trip
13	EndDate	Date and time of the end of the trip
14	TripLength	Trip length in miles
15	TripSpeed	Average trip speed in miles per hour

3.3.2. Query Results Display

The data elements displayed for query outputs are the number of trips, average trip length (miles), average trip duration (seconds), and average trip speed (miles per hour). The number of rows of data depends upon the "Aggregate By" criteria employed. If none of the "Aggregate By" boxes are checked we get one row of output. For example, the figure below represents the output displayed in the Data Area by querying the trip file indicated in the previous figure without any "Select By" or "Aggregate By" options chosen.
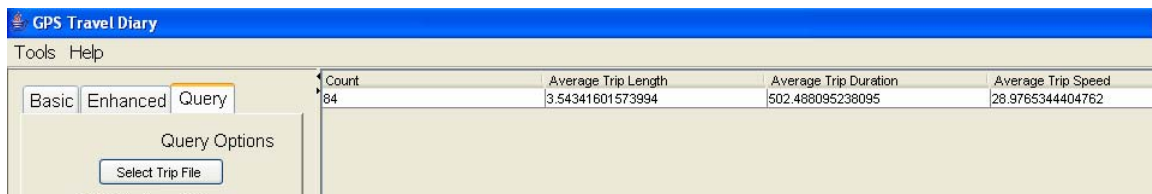
The figure below is the output displayed when the trip file as shown in Section 3.3.1 is aggregated by trip-end activity type.

EndActivity	Count	Average Trip Length	Average Trip Duration	Average Trip Speed
Eat_out	3	1.5807653106217	263.0	24.690667
Grocery_shopping	8	3.16534419003562	497.25	26.65397225
Home	7	2.49879209424674	419.714285714286	27.7753341428571
Other_shopping	1	2.49853956603812	405.0	32.48414
Personal_services	5	2.83712228282214	510.0	27.6844246
School	13	4.52116747441686	573.692307692308	33.3323086923077
Serve_passenger_at_other ...	12	2.51730006166305	410.083333333333	25.5972218333333
Serve_passenger_at_school	17	3.45958006378119	480.764705882353	28.8027478235294
Social	4	7.71681221225867	910.5	33.1788515
Work	9	3.10566422901045	440.666666666667	30.5824796666667
Work_related	5	5.35846686566881	677.8	28.6600802

The figure below is the output displayed when the subset of non home-based trips (from the same trip file as shown in Section 3.3.1) is aggregated by trip-end activity type.

4. Inputs and Outputs

This chapter of the manual describes the various inputs provided to GPSTDG and the structure of the output files that the software generates. This chapter is organized as follows. Section 4.1 describes the data inputs to the software, Section 4.2 describes the parameter inputs to the software, and Section 4.3 describes the outputs.

4.1. DATA INPUTS

There are four main data inputs required for GPSTDG. These are (1) Raw GPS navigational streams, (2) Individual and household characteristics, (3) A GIS layer of TAZ boundaries, and (4) A GIS land use layer. Each of these is discussed in the following sections.

4.1.1 Raw GPS Navigational Streams

The raw GPS navigational streams represent the GPS data largely in the same format in which they have actually been recorded by the in-vehicle devices. It is assumed that the stream from each vehicle is saved as a separate comma-separated file (i.e., one file per surveyed vehicle) with each line comprising a sequence of data elements in a pre-defined format (this sequence is also called a sentence). The software is capable of handling sentences that are either in the "GPRMC" or in the "GeoLogger" format. The raw GPS files are to be named with a ".dlv" extension if they are in the GeoLogger format and with a ".gprmc" extension if they are in the GPRMC format. The sentence structures of the two formats are presented below.

The GeoLogger Sentence Format

Example:
5,273,1005361,1,03/25/02,13:13:24,03/25/02,07:13:24,27.52020,99.46400,00134,0.00,000,01.8,00

Generic Structure:
RecType, GPSID, HHID, VehID, UTC_Date, UTC_Time, Loc_Date, Loc_Time, Lat_Raw, Long_Raw, Elev_Raw, Speed, Heading, HDOP, SATS

1. RecType takes the value 5 indicating a GPS record
2. GPSID is the identifier for the GPS device used.
3. HHID is the identifier for the household.
4. VehID is the identifier for the GPS-equipped vehicles in a household.
5. UTC_Date is the UTD date in the MM/DD/YY format
6. UTC_Time is the UTC military time in the HH:MM:SS format

7. Loc_Date is the local date in the MM/DD/YY format

8. Loc_Time is the local military time in the HH:MM:SS format

9. Lat_Raw is the instantaneous latitude position recording in decimal degrees

10. Long_Raw is the instantaneous longitude position recording in decimal degrees.

11. Elev_Raw is the instantaneous elevation (in meters)

12. Speed is the instantaneous speed recoding in meters per second (0 to 514 m/s).

13. Heading is the direction of movement in degrees (0 to 359.9 degrees).

14. HDOP is the horizontal dilution of precision (0.5 to 99.9)

15. SATS is the number of satellites in view (0 to 12)

The GPRMC Sentence Format

Example:
$GPRMC,040113,A,3653.10,S,17437.47,E,000.6,074.4,190903,,*03

Generic Structure:
"$GPRMC", Time, Status, Latitude, Lat_Hem, Longitude, Long_Hem, Speed, Heading, Date, Blank, CheckSum

1. "$GPRMC" is a record-type identifier

2. Time is the UTC time in HHMMSS.SSS format.

3. Status field takes the value "A" if the record is valid or "V" otherwise.

4. Latitude is the instantaneous latitude position recording (ddmm.mmmm format).

5. Lat_Hem represents the latitude hemisphere (N = North and S = South)

6. Longitude is the instantaneous longitude position recording (dddmm.mmmm format).

7. Long_Hem represents the longitude hemisphere (E= East and W = West)

8. Speed is the instantaneous speed recoding in knots (values between 0 and 999.9).

9. Heading is the direction of movement in degrees (0 to 359.9 degrees).

10. Date is the UTC date in the DDMMYY format.

11. Blank is a blank entry

12. Checksum (the *03 entry in the example)

If data is recorded in the GPRMC format, identifiers for the GPS unit, the household, and the vehicle have to be manually added to files containing the GPS streams. This is accomplished by adding a single line header with the identifiers to each raw GPS stream file. The identifiers are in the following sequence: GPSID, HHID, VehID

4.1.2 Individual and Household Characteristics

Data on individual and household characteristics of the survey respondents are provided as two files: (1) The Demographics File, and (2) The Relationship File. The first file includes the household and work locations of the survey respondents, and other characteristics that are required for activity type determination in the enhanced mode. The second file associates or links each vehicle equipped with the GPS device with one member of the household called the primary driver. The structure of these files is presented below:

The Demographics File

This is a comma-separated file with a ".demo" extension. This file comprises two types of records: the household record and the person record. The surveyed households are represented sequentially in this file. Each household is represented by one household record followed immediately by one or more person records. Every person in the household who has been identified as a primary driver of a GPS-equipped vehicle should be represented by a person record. Individuals who are not primary drivers (for example, adults without a driver's license and children) need not be included in this file.

Example (representation for a household with one primary driver):
DREC,1,1001373,-99.470800, 27.469860,2
DREC,2,1001373, 1,1,-99.479140, 27.571130,1,1

The generic structure of the household record is as follows:
"DREC," RecType, HHID, HomeLocLat, HomeLocLong, NumChild

1. "DREC" is a record-type identifier indicating that this is a record in a demographics file.
2. RecType is the identifier used to distinguish between household and person records in the location and demographics file. Household records have an entry 1.
3. HHID is the identifier for the household.
4. HomeLocLat is the latitude of the home location in decimal degrees.
5. HomeLocLong is the longitude of the home location in decimal degrees.
6. NumChild is the number of children (age < 16) in the household. This is required only for enhanced analysis. If the software is to be run only in the basic mode, this user can specify a value of zero for this field.

The person record has the following generic format:
"DREC," RecType, HHID, PersID, EmpFlag, WorkLocLat, WorkLocLong, Student, Male

1. "DREC" is a record-type identifier indicating that this is a record in the locations and demographics file.

2. RecType is the identifier used to distinguish between household and person records in the location and demographics file. Person records have an entry 2.

3. HHID is the identifier for the household.

4. PersID is the identifier for a person (primary driver) in the household.

5. EmpFlag is a binary flag used to indicate the employment status of the individual (1 if employed).

6. WorkLocLat is the latitude of the work location in decimal degrees (provide a value of zero if the person is not employed).

7. WorkLocLong is the longitude of the work location in decimal degrees (provide a value of zero if the person is not employed).

8. Student is a binary flag used to indicate the student status of the individual (1 if a student. This is required only for enhanced analysis. If the software is to be run only in the basic mode, this user can specify a value of zero for this field.

9. Male is a binary flag used to indicate the gender of the individual (1 if male). This is required only for enhanced analysis. If the software is to be run only in the basic mode, this user can specify a value of zero for this field.

The Relationship File

The relationship file (also called the "link" file) is used to specify an association between each vehicle equipped with a GPS device and a member of the household who is the "primary driver" of this vehicle. Hence, this file provides a link between the files containing the GPS streams and the person-level information contained in the Demographics File. There is one record in this file for each GPS equipped vehicle in the survey. This is a comma-separated file with a ".link" extension.

Example:
LREC,040802_1001373_V1.dlv,1001373,1,1

Generic Format:
 "LREC," GPSFilename, HHID, VehID, PersID
1. "LREC" is a record-type identifier indicating that this is a record in the link file.
2. GPSFilename is the name of file containing the raw GPS stream corresponding to the vehicle VehID.
3. HHID is the identifier for the household.
4. VehID the identifier for the GPS-equipped vehicles in a household.
5. PersID is the identifier for the primary driver in the household associated with the vehicle VehID.

4.1.3 A GIS Layer of TAZ Boundaries

The TAZ boundaries GIS layer is to be provided as an ESRI shape file. The software is capable of handling any co-ordinate system as along as the TAZ Ids are provided in the 7th column called "TAZ" in the underlying database.

4.1.4 A GIS Land-Use Layer

The GIS land-use layer is also to be provided in the ESRI shape file format. Again, the software is capable of handling this spatial data any co-ordinate system. However, the land-use variable must be provided in the 4th column called "P_LANDUSE" in the underlying database. Further, the following zonal land-use classification scheme must be used: "LDR", "MDR", "HDR", "HEAVY INDUSTRIAL", "INSTITUTIONAL", "RETAIL/OFFICE", "AG", "WATER", "PARK/REC. OPEN SPACE", and "WAREHOUSE/LI".

4.2. PARAMETER INPUTS

GPS-TDG takes two types of parameters as inputs: (1) Specification of the MNL model for activity type classification, and (2) Algorithm parameters. These are discussed in further detail below.

4.2.1 MNL Model Specification

The disaggregate activity type classification and the parameters of the MNL model (i.e., coefficients on the explanatory variables) are specified using a comma-separated file. The first row of input defines the structure of this file. Each subsequent row of this file comprises the set of parameters defining the "utility function" for one activity purpose. The parameters are input in the sequence defined in the first line. The analyst can classify the activities into any number of disaggregate purposes and provide the appropriate number of rows of input. In this version of the software, the set of explanatory variables that can be used are pre-determined. These variables and the specification structure are discussed below:

Example (specification of parameters for one of the activity purpose types)

MREC,Serve_passenger_at_otherlocation,0.0,0.0,0.0,0.9835,0.0,0.0,0.0,0.0384,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.1509,5.8298

The first entry in this record is "MREC" identifying this record contains model parameters. The second field has the name of the activity type. The subsequent records represent the coefficients on the explanatory variables corresponding to this activity type. The sequence of explanatory variables and their descriptions are provided in the table below:

Column	Variable Name	Description
3	Industrial_or_warehouse	1 if the trip end falls in a zone with industrial or warehouse land use, 0 otherwise
4	Institutional	1 if the trip end falls in a zone with institutional land use, 0 otherwise
5	Retail/office_space	1 if the trip end falls in a zone with retail or office land use, 0 otherwise
6	LDR	1 if the trip end falls in a zone with LDR land use, 0 otherwise
7	MDR_or_HDR	1 if the trip end falls in a zone with MDR or HDR land use, 0 otherwise
8	No_land_use	1 if the trip end falls in a zone with blank or no land use, 0 otherwise
9	Trip_duration	Duration of the trip in seconds
10	Activity_duration	Duration of the activity at trip end in seconds
11	7_AM_9_AM	1 if trip ends between 7 AM and 9 AM, 0 otherwise
12	9_AM_2_PM	1 if trip ends between 9 AM and 2 PM, 0 otherwise
13	2_PM_6_PM	1 if trip ends between 2 AM and 6 PM, 0 otherwise
14	after_6_PM	1 if trip ends after 6 PM, 0 otherwise
15	Trip_origin_is_home	1 if the trip originates at home, 0 otherwise
16	Trip_origin_is_work	1 if the trip originates at work, 0 otherwise
17	Employed	1 if the primary driver is employed, 0 otherwise
18	Student	1 if the primary driver is a student, 0 otherwise
19	Male	1 if the primary driver is male, 0 otherwise
20	Number_of_children_in_HH	Number of children in the household
21	Constant	Alternative-specific constant

The default MNL model parameters provided with the software were estimated using data from Laredo employing a nine-way classification scheme for the activity types. The complete empirical specification of this model in the input format is presented in Appendix A. The reader is referred to Report 2 for the model estimation details.

4.2.2 Algorithm Parameters

The trip detection and characterization algorithm implemented in the GPS-TDG software is controlled by several user-specified parameters. The values of these parameters can be varied as required by the analyst to suit the requirements of different study areas and for various analysis objectives. A description of all the user-defined parameters is presented in table below. Our prescribed default values are also provided.

Parameter Name	Flag	Description	Default Flag/Value
Debug Options Flag	-debug or -nodebug[1]	Determines whether the software is run in the debug mode (-debug) or not (-nodebug). The analyst will only run the software in the –nodebug mode. The –debug option was developed for software testing purposes	-nodebug
Non Engine Power-off Flag	-nepo or -nonepo[1]	Determines whether the algorithm to determine non engine power off stops is executed (-nepo) or not (-nonepo).	-nepo
Engine Power-off Dwell Time Threshold	-t	Minimum dwell-time gap between successive valid GPS records for signalling an engine power off stop (seconds)	30 secs
Non Engine Power-off Dwell Time Threshold	-t2	Minimum duration for which the speed should be less than the Speed Threshold for signalling a non engine power off stop	60 secs
Speed Threshold for Non Engine-power off Stop Determination	-s2	Value of instantaneous speed below which the vehicle is assumed to be at rest (used in determination of non engine-power off stops)	0.1 m/s
Speed Threshold for Trip Distance Computation	-s	If the value of instantaneous speed is below this threshold, this GPS point is not used in distance computation.	0.5 m/s
Time Threshold for Trip Distance Computation	-i	Minimum time between pairs of GPS points for which the distances are computed and summed to determine the trip length.	1 secs
Home Location Distance Threshold	-h	Maximum distance between a trip-end location and home location for classifying the trip-end activity purpose as "Home"	0.3 miles
Work Location Distance Threshold	-w	Maximum distance between a trip-end location and work location of primary driver for classifying the trip-end activity purpose as "Work"	0.5 miles
Work Duration Threshold	-d	Minimum activity duration at a trip-end for classifying a trip end activity as "Work"	3600 sec
Average Update Rate	-u	Average duration between successive navigational points recorded by the GPS instrument used in the survey	1/sec
Minimum Trip Duration Threshold	-mt	Minimum trip duration for a potential trip to be classified as a real trip	120 secs
Minimum Trip Length Threshold	-ml	Minimum trip length (distance) for a potential trip to be classified as a real trip	1 miles

[1] there is no value following this flag

In order to change the parameters via the Options Box in the Command Area, the parameter flag followed by the appropriate value (if needed) should be provided as already described in Section 3.2. Alternately, the analyst can also make changes to the

GPSTDG.options file, again by simply changing the flags and/or parameter values. This file with the default flags /parameter values is presented in Appendix B.

4.3. OUTPUTS

There are two main outputs generated by the GPS-TDG software. These are (1) the pre-processed GPS navigational stream files, and (2) the derived trips in the travel diary format. Each of these are discussed below.

4.3.1 Pre-Processed GPS Streams

The pre-processed GPS streams are generated by the pre-processor and serve as subsequent inputs to the main trip-detection component of the GPS-TDG software. These files are written out in the comma separated format with a ".cln" extension. There is one file corresponding to each vehicle in the survey. The structure of each row of data for these files is discussed below.

Example:
GREC,273,1005361,1,1017058405000,27.52020,-99.46402,0.00,000,1

Generic Structure:
"GREC", GPSID, HHID, VehID, Time, Latitude, Longitude, Speed, Heading, NumInvRecs

1. "GREC" is a record-type identifier indicating that this is a record in a cleaned GPS stream file.
2. GPSID is the identifier for the GPS device used.
3. HHID is the identifier for the household.
4. VehID is the identifier for the GPS-equipped vehicles in a household.
5. Time is the instantiations time in milliseconds from 1970 (corrected for DST)
6. Latitude is the instantaneous latitude position recording in decimal degrees
7. Longitude is the instantaneous longitude position recording in decimal degrees.
8. Speed is the instantaneous speed recoding in meters per second (0 to 514 m/s).
9. Heading is the direction of movement in degrees (0 to 359.9 degrees).
10. NumInvRecs is the number of invalid records immediately prior to this GPS record that were removed during pre-processing

Each of the pre-processed GPS streams file also contains a header record which provides the parameter values used in pre-processing. For example, the header row of pre-processed GPS file created from raw GeoLogger files could be:

HREC,HDOP=5.0,Min Satellites=3

This indicates that GPS records with an HDOP value greater than 5 or number of satellites less than 3 were considered invalid and removed.

In the case that the raw GPS streams have the GPRMC format, the A/V flag of the GPS sentences are used in pre-processing. Therefore the header row of a corresponding cleaned file does not have any parameter values and simply reads HREC, GPRMC.

4.3.2 Derived Travel Diaries

The derived travel diary constitutes the fundamental output from the GPSTDG software. The file containing these derived travel patterns is written out in two formats. The first file with a ".trp" extension is a compact comma separated file with headers. The second file is a flat file (with a ".csv" extension). This file structure is useful for display and processing queries. Further, the file format can also be opened in EXCEL or other spreadsheet applications with ease for further analysis.

Format of the ".trp" file

The overall format of the ".trp" file is as follows:

"HREC", WorkThresh, HomeThresh, SpeedThresh, DwellTime, MinWorkDur

"VH", HHVehID, DriverID

"TR", OriginalTripNo, TripNo, TripDuration, StartTAZ, EndTAZ, StartLU, EndLU, StartActivity, EndActivity, StartLat, StartLong, EndLat, EndLong, StartDate, EndDate, TripLength1, AvSpeed1, VarSpeed1, TripLength2, AvSpeed2, VarSpeed2, StartTOD, EndTOD, NRecRatio

"VT", HHVehID, NumTrips

....

TREC, TripCount, VehCount,

The first record in this file is a header characterized by the entry "HREC" in its first field. This header record sequentially includes the values of all the threshold parameters used in the travel-diary extraction algorithm. The complete format of this header is as follows:

"HREC", WorkThresh, HomeThresh, SpeedThresh, DwellTime, MinWorkDur

1. "HREC" is a record-type identifier indicating that this is a header record in the output file.
2. WorkThresh is the parameter defining the maximum distance between a trip-end location and work location of primary driver for classifying the trip-end activity purpose as "Work"
3. HomeThresh is the parameter defining maximum distance between a trip-end location and home location for classifying the trip-end activity purpose as "Home"
4. SpeedThresh is the parameter for trip distance computation (If the value of instantaneous speed is below this threshold, this GPS point is not used in distance computation)

5. DwellTime is the parameter defining the minimum dwell-time gap between successive valid GPS records for signalling an engine power off stop (seconds)
6. MinWorkDur is the parameter defining the minimum activity duration at a trip-end for classifying a trip end activity as "Work"

The very last record in the output file is a trailer record with the following format:

"TREC", TripCount, VehCount,

where

1. "TREC" is a record-type identifier indicating that this is a trailer record in the output file.
2. TripCount is the total number of trips derived from GPS streams contained in the output file.
3. VehCount is the total number of vehicles in the output file.

The rest of the output file is comprised of three types of records: the vehicle header records, the vehicle trailer records, and the trip records. The vehicle header and vehicle trailer records, respectively, are included immediately before and after the trip records corresponding to a particular vehicle and thus help to distinguish the trips of one vehicle from another. The vehicle header record has the following format:

"VH", HHVehID, DriverID,

where

1. "VH" is a record type identifier indicating that this is a vehicle header record.
2. HHVehID the identifier for the GPS-equipped vehicles in a household. This is created by appending the vehicle ID to the household ID
3. DriverID is the identifier for the primary driver in the household associated with the vehicle HHVehID. This is created by appending the Person ID to the Household ID

The trailer record has the following format:

"VT", HHVehID, NumTrips,

where

1. "VT" is a record type identifier indicating that this is a vehicle trailer record.
2. HHVehID the identifier for the GPS-equipped vehicles in a household. This is created by appending the vehicle ID to the household ID
3. NumTrips is the total number of trips detected for this vehicle. This is equal to the number of trip records included between the header and the trailer records of a particular vehicle.

Each trip record in this file contains the characteristics of a single trip in the following format:

"TR", OriginalTripNo, TripNo, TripDuration, StartTAZ, EndTAZ, StartLU, EndLU, StartActivity, EndActivity, StartLat, StartLong, EndLat, EndLong, StartDate, EndDate, TripLength1, AvSpeed1, VarSpeed1, TripLength2, AvSpeed2, VarSpeed2, StartTOD, EndTOD, NRecRatio

where

1. TR is a record type identifier indicating that this is a trip record

2. OriginalTripNo is the trip number assigned originally to all potential trips. These trips are subsequently classified as either true trips and retained or as false trips and discarded. Further, a single potential trip could also be broken into two or more trips if non-engine power-off stops are detected.

3. TripNo is a trip identifier. The trips are sequentially numbered over the final set of "true" trips.

4. TripDuration is the duration of trip is seconds.

5. StartTAZ is the TAZ identifier for the trip-start location.

6. EndTAZ is the TAZ identifier for the trip-end location.

7. StartLU is the land-use at the trip-start location.

8. EndLU is the land-use at the trip-end location.

9. StartActivity is the activity type pursued by the driver of the vehicle at the origin of the trip (an aggregate classification scheme is used in the basic analysis and a disaggregate analysis scheme is used in the enhanced analysis).

10. EndActivity is the activity type pursued by the driver of the vehicle at the destination of the trip (the classification scheme is same as that for StartActType).

11. StartLat is the latitude of the trip-start location.

12. StartLong is the longitude of the trip-start location.

13. EndLat is the latitude of the trip-end location.

14. EndLong is the longitude of the trip-end location.

15. StartDate is the date and time of the start of the trip.

16. EndDate is the date and time of the end of the trip.

17. TripLength1 is the trip length (in miles) computed using the instantaneous position information.

18. AvSpeed1 is the average speed of the trip (miles per hour) computed (speed computed as ratio of distance between successive GPS records to the time stamp between successive records)

19. VarSpeed1 is the variance is the speed over the length of the trip (speed computed as ratio of distance between successive GPS records to the time stamp between successive records)

20. TripLength2 is the trip length (in miles) computed using the instantaneous speed information.

21. AvSpeed2 is the average speed of the trip (miles per hour) computed using instantaneous speed data.

22. VarSpeed2 is the variance is the speed over the length of the trip.

23. StartTOD is the time of day of start of trip (in the hhmm format, for example, 1527 refers to 3:27 PM)

24. EndTOD is the time of day of end of trip (in the hhmm format)

25. NRecRatio is an accuracy measure computed as the ratio of the number of valid GPS records corresponding to this trip to the total number of records for this trip.

Format of the ".csv" file

The comma-separated file (or the flat file) has all the information as the trip records in the ".trp" file. The first field of a record in this file has the entry "TR" indicating a trip record. The next two fields are respectively the vehicle and driver identifiers (i.e., the HHVehID and HHPersID variables in the ".trp" file). The next twenty four columns correspond respectively to the fields 2-25 in the trip records of the ".trp" file.

5. Basic Operations

This chapter of the report describes the three basic operations that an analyst performs using the software. These are (1) Pre-processing GPS streams, (2) Running the trip detection algorithm, and (3) Querying the outputs. Instructions for performing each of these operations are presented in this chapter. Further, these discussions are in the context of the sample data files provided along with the software.

The rest of this chapter is organized as follows. Section 5.1 describes the sample input files provided along with the software. Section 5.2 describes pre-processing the GPS streams. Section 5.3 describes running the software in both basic and enhanced modes. Finally, Section 5.4 describes the procedures for querying the results and interpreting the results.

5.1. SAMPLE DATA DESCRIPTION

Sample electronic data files are provided in the folder C:\GPSTDG\Laredo. This folder contains the following <u>input</u> data files and folders:

1. Laredo.demo is the demographics file. This file has data for 13 drivers from 12 households. The structure of this file is consistent with the description provided in Section 4.1.2.
2. Laredo.link is the relationships file. This file "links" each vehicle (represented by files in the GPS folder) with a primary driver (represented as person records in the Laredo.demo file). Again, the structure of this file is consistent with the description provided in Section 4.1.2.
3. The folder GPS contains the raw GPS streams from 53 vehicles. Each file in this folder represents streams in the Geologger format from a single vehicle. These file names have .dlv extensions and are consistent in format with the description provided in Section 4.1.1
4. The folder SpatialData has GIS files for (1) the TAZ boundaries for the Laredo area, and (2) Parcel-level land use. The specifications of these files have been discussed in Sections 4.1.3 and 4.1.4 respectively.

In addition to the above input data items, the default input parameters are also provided as files. These are the GPSTDG.options file containing the algorithm parameters and the ModelParameters.csv file containing the MNL model specification. Both these files are available in the C:\GPSTDG folder.

The outputs corresponding to the above-described sample data are also provided with the software.
1. The pre-processed GPS streams files are provided in the folder C:\GPSTDG\Laredo\GPSClean. There is one processed GPS file with a ".cln" extension corresponding to every input file and these files have a structure as described in Section 4.3.1.

2. The derived trip diaries are provided in the C:\GPSTDG\Trips folder. The structures of these two files have been described in Section 4.3.2.

5.2. PREPROCESSING GPS STREAMS

The raw GPS streams can be preprocessed for subsequent input to the software in the following way:

1. Create a folder called GPSClean in folder C:\GPSTDG\Laredo.
2. Double click on the GPSTDG.bat file to run the GPS-TDG software.
3. Click on *Tools* in the Menu Bar and select *Preprocess GEO* from the drop-down list.
4. The *Select Source Directory* window pops. Navigate to C:\GPSTDG\Laredo\GPS and click the *Select Source Directory* button.
5. The *Select Target Directory* window pops. Navigate to C:\GPSTDG\Laredo\GPSClean and click the *Select Target Directory* button. This triggers the pre-processor to run.
6. The completion of the run is indicated by a pop-up window. The pre-processed GPS stream files can be found in the C:\GPSTDG\Laredo\GPSClean folder.

5.3. RUNNING TRIP DETECTION

Once the GPS streams have been pre-processed, the trip detection component of the software can be run. The procedure to run the software in the Enhanced Mode is described in Section 5.3.1. The simplified procedure for running the software is run in the Basic Mode is discussed next in Section 5.3.2.

5.3.1 Enhanced Trip Detection

The following are the steps for running GPS-TDG software in the Enhanced Mode

1. Click on the Enhanced Pane in the Command Area
2. Click on the *Select GPS Directory* button. Navigate to C:\GPSTDG\Laredo\GPSClean and select this directory.
3. Click on the *Select TAZ File* button. Navigate to C:\GPSTDG\Laredo\SpatialData\LaredoTAZ.shp and select this file.
4. Click on the *Select Demographics File* button. Navigate to C:\GPSTDG\Laredo\Laredo.demo and select this file.
5. Click on the *Select Relation File* button. Navigate to C:\GPSTDG\Laredo\Laredo.link and select this file.
6. Click on the *Select Output File* button. Enter "LaredoE" as the file name and click OK.
7. Click on the *Select LandUse File* button. Navigate to C:\GPSTDG\Laredo\SpatialData\LUTrimmed.shp and select this file.
8. Click on the *Select Model Parameters File* button. Navigate to C:\GPSTDG\ModelParameters.csv and select this file.

9. This run will be performed using the default algorithm parameters (as indicated in Section 4.2.2 and provided in the C:\GPSTDG\GPSTDG.options file). Hence the *Options* text box is to be left blank.
10. Click on the *Detect Trips* button. This triggers the software run. A progress bar pops up to indicate processing progress. Files LaredoE.trp and LaredoE.csv are created in the directory C:\GPSTDG\Trips. When the run is complete, the output is loaded into the Data Area as shown below:
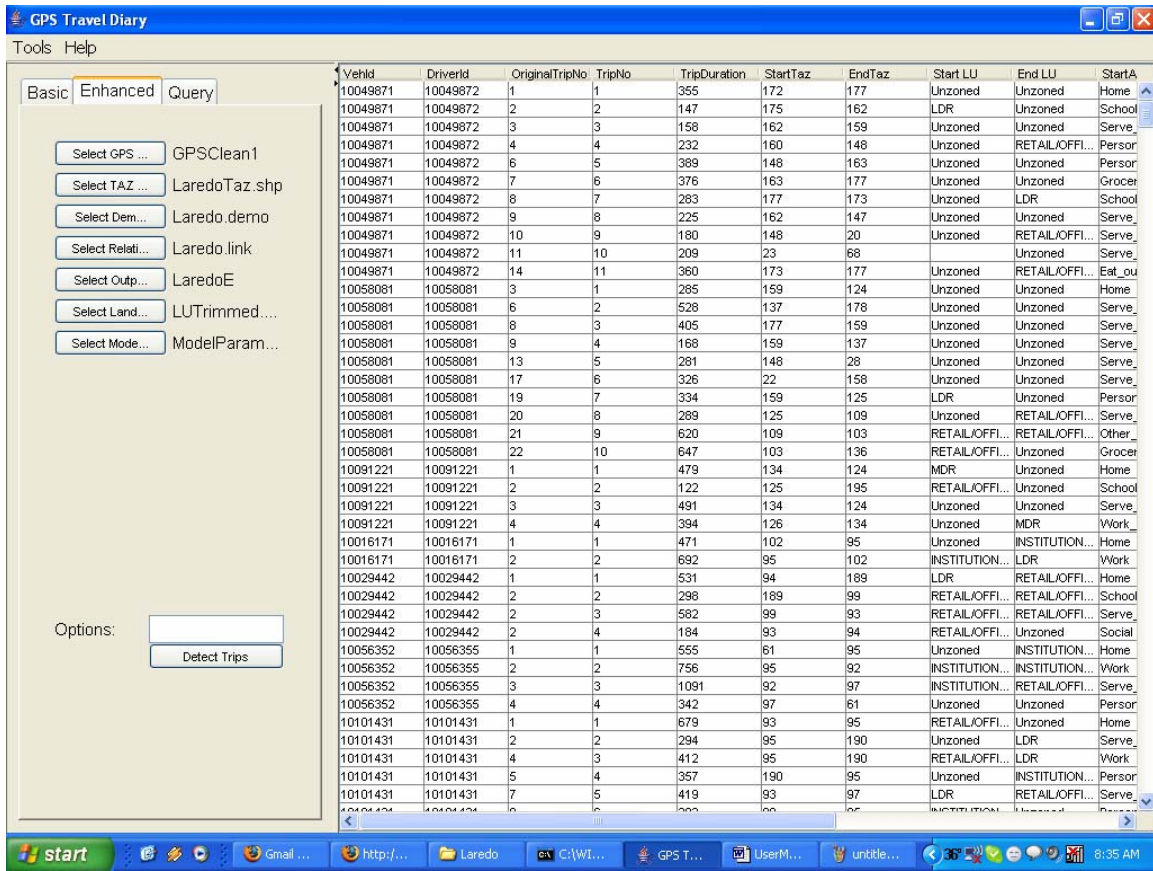

5.3.2 Basic Trip Detection

The following are the steps for running GPS-TDG software in the Basic Mode

1. Click on the Basic Pane in the Command Area
2. Follow steps 2-5 as described in Section 5.3.2.
3. Click on the *Select Output File* button. Enter "LaredoB" as the file name and click OK.
4. This run will be performed using the default algorithm parameters (as indicated in Section 4.2.2 and provided in the C:\GPSTDG\GPSTDG.options file). Hence the *Options* text box is to be left blank.

5. Click on the *Detect Trips* button. This triggers the software run. A progress bar pops up to indicate processing progress. Files LaredoB.trp and LaredoB.csv are created in the directory C:\GPSTDG\Trips. When the run is complete, the output is loaded into the Data Area as shown below
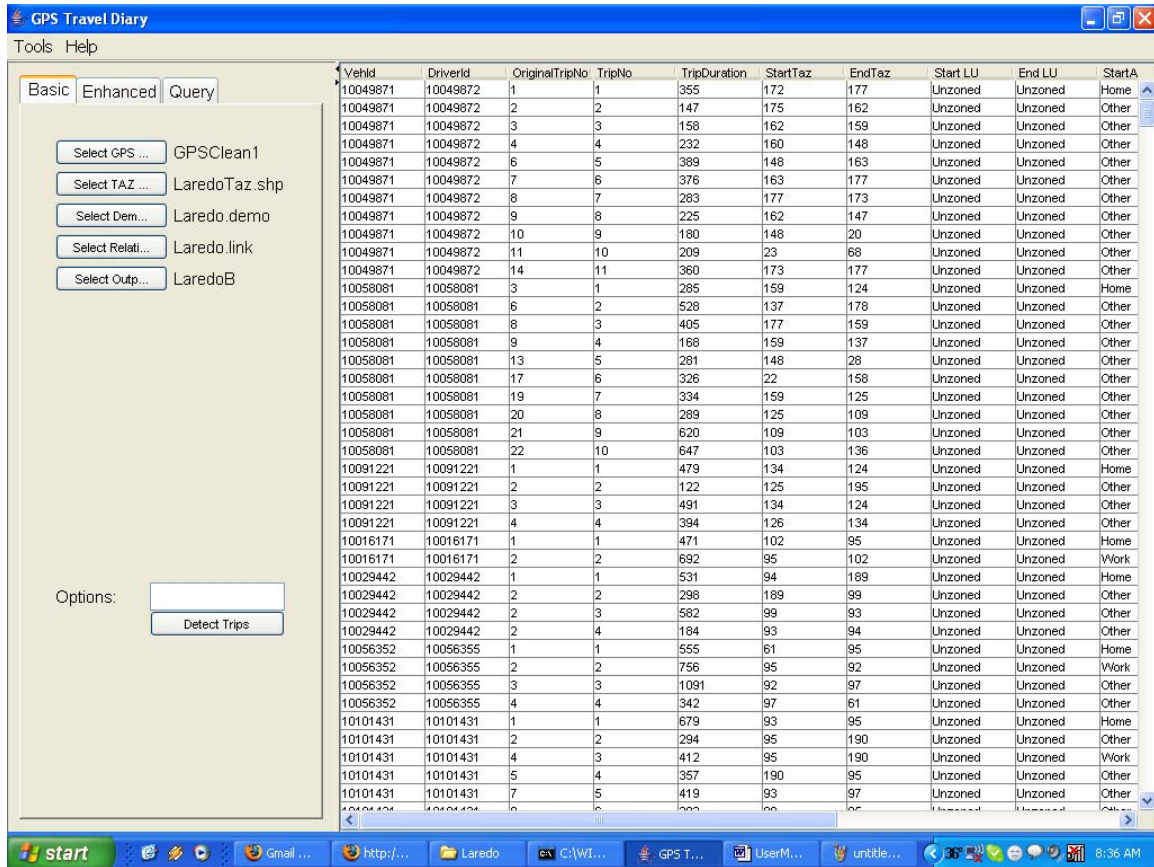
The reader will note that (1) the StartLU and EndLU columns have the entry "Unzoned" reflecting the absence of land use data and (2) the StartActivity and EndActivity columns classify trip ends into one of three aggregate categories, viz., Home, Work, and Other.

5.4. RUNNING QUERIES AND INTERPRETING RESULTS

In this section, we discuss some of the aggregation-analysis that can be performed using GPSTDG. We will use the output file generated from running the software in the Enhanced Mode (i.e., LaredoE.csv). To load this data into the software for analysis, first click on the Query Tab in the Command Area. Click on the *Select Trip File* button, navigate to C:\GPSTDG\Trips\LaredoE.csv and select this file. The rest of this section describes the procedures for performing some aggregation queries and also discusses the output results.

Overall Average Measures

To compute the averages over the entire derived trip file, simply click the *Execute Query* button without selecting any of the check boxed under the "Aggregate By" and "Select By" Sections of the Query Pane. The output is one row of measures as shown in figure below.
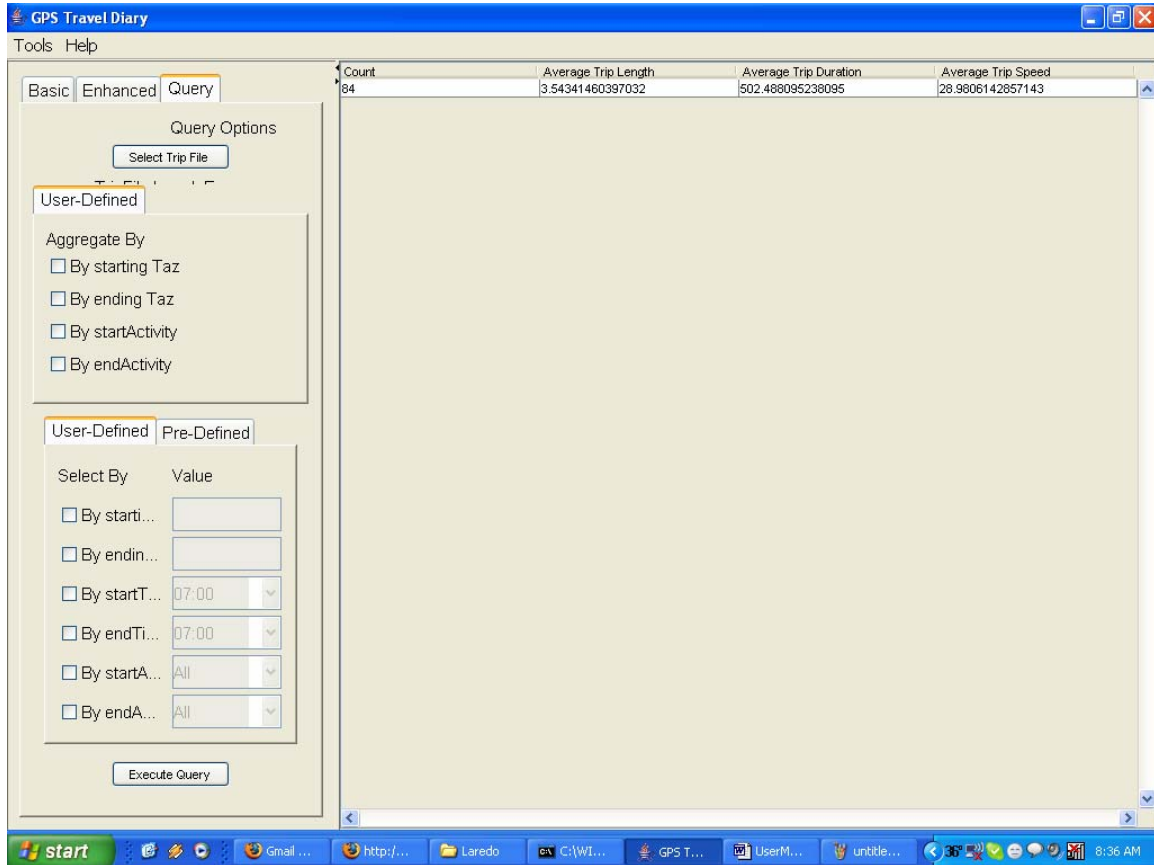
The reader will note that there are 84 trips in the overall. The average trip length is 3.54 miles, the average trip duration is 502 seconds, and the average trip speed is 29 miles per hour.

Trip table and inter-zonal performance measures

To obtain the inter-zonal performance measures and the trip table, check the By Starting TAZ and By ending TAZ boxes under the "Aggregate By" section of the Query Pane and click the *Execute Query* button. The output comprises one row of data for each zonal pair represented in the data as shown below.
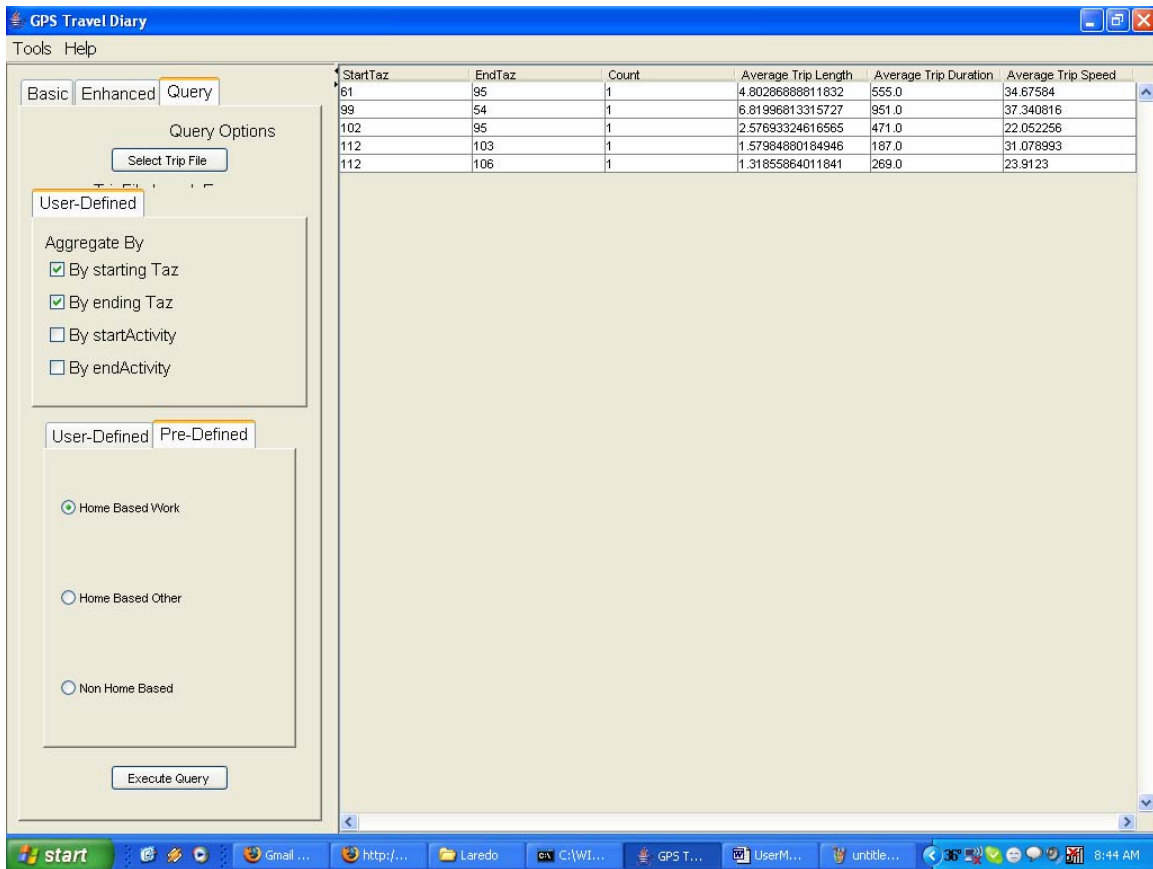
For example, the reader will note that there are 3 trips between zones 95 and 190. The average trip length for these trips is 2.04 miles, the average trip duration is 286 seconds, and the average trip speed is 33 miles per hour.

Trip table and inter-zonal performance measures by time of day

The previous analysis determined the trip tables and highway performance measures over the entire day. To determine the same measures, but for a specific time period of the day (say the PM peak of 3:30 PM to 6:30 PM), first check the *By start time* box and enter the value (i.e., 15:30) in the editable list adjacent to the check box. Next, check the *By end time* box and enter the value (i.e., 18:30) in the editable list adjacent to the check box. Next, check the By Starting TAZ and By ending TAZ boxes under the "Aggregate By" section of the Query Pane and click the *Execute Query* button. The output comprises one row of data for each zonal pair represented in the data as shown below.

The reader will note that there is 1 trip between zones 96 and 103 that originates after 3:30 PM and ends before 6:30 PM. The average trip length for this trip is 2.54 miles, the average trip duration is 439 seconds, and the average trip speed is 24 miles per hour.

Trip table and inter-zonal performance measures by trip purpose

The second analysis determined the trip tables and highway performance measures over all trip purposes. To determine the same measures, but for a specific trip purpose (say the home-based work), first click on the Pre-defined tab in the "Select By" area of the Query pane and check the box marked *Home Based Work.* Next, check the By Starting TAZ and By ending TAZ boxes under the "Aggregate By" section of the Query Pane and click the *Execute Query* button. The output comprises one row of data for each zonal pair represented in the data as shown below.

The reader will note that there is 1 home-based work trip between zones 99 and 54. The average trip length for this trip is 6.82 miles, the average trip duration is 951 seconds, and the average trip speed is 37 miles per hour.

Appendix A

Contents of the ModelParameters.csv file, the default specification for activity type classification estimated using data from Laredo

RecordType,Activity_Type,Industrial_or_warehouse,Institutional,Retail/office_space,LDR,MDR_or_HDR,No_land_use,Trip_duration,Activity_duration,7_AM_9_AM,9_AM_2_PM,2_PM_6_PM,after_6_PM,Trip_origin_is_home,Trip_origin_is_work,Employed,Student,Male,Number_of_children_in_HH,Constant
MREC,Work_related,1.5104,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.4378,0.0,0.0,0.0,0.0,0.0,0.0,3.0294,0.0,0.0,0.0,0.0,0
MREC,School,0.0,0.7022,0.0,0.0,0.0,0.0,0.0,0.0,0.00.0051,1.1321,0.0,0.0,0.0,0.0,0.977,0.0,0.0,0.0,3.265,0.0,0.0,1.4004
MREC,Grocery_shopping,0.0,0.0,1.9651,0.0,0.0,0.0,1.2656,0.0,0.00.0165,0.0,0.501,0.0,0.6922,0.0,0.0,0.0,0.0,0.0,0.0,0.0,3.9153
MREC,Other_shopping,0.0,0.0,2.0343,0.0,0.0,0.0,0.0,0.0,0.00.0115,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,3.7517
MREC,Personal_services,0.0,0.0,0.3002,0.0,0.0,0.0,0.0,0.0,0.00.0123,0.0,0.0,0.7788,0.4191,0.0,0.0,0.0,0.0,0.0,0.0,0.0,5.0575
MREC,Social,0.0,0.0,0.0,0.0,1.4129,1.1106,0.0,0.0,0.0,0.00.0083,0.0,0.0,0.0,0.0,0.9557,0.0,0.0,0.0,0.0,0.0,0.0,0.0,4.4005
MREC,Eat_out,0.0,0.0,0.8477,0.0,0.0,0.0,0.0,0.00.0157,0.00.0129,0.0,1.2642,0.0,1.1763,0.0,1.2323,0.0,0.0,0.0,0.2752,0.0,4.0095
MREC,Serve_passenger_at_school,0.0,1.8989,0.0,1.7472,1.9238,0.9117,0.0,0.00.0294,2.6581,0.0,1.6574,0.0,0.8614,0.0,0.0,0.0,0.00.3425,0.4816,3.3162
MREC,Serve_passenger_at_other location,0.0,0.0,0.0,0.9835,0.0,0.0,0.0,0.00.0384,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.1509,5.8298

Appendix B

Contents of the GPSTDG.options file with the default flags and parameter values

DebugFlag (-debug/-nodebug)# -nodebug
EnhancedFlag (-enhanced/-standard)# -standard
DwellTime Minimum time to detect an engine power off stop(in secs)# -t 30
NEPO DwellTime Minimum time to detect a non engine power off stop(in secs)# -t2 60
Non Engine Power Off Flag (-nepo/-nonepo)# -nepo
Minimum Trip Duration(in seconds)# -mt 120
Minimum Trip Length(in miles)# -ml 1
Work Threshold Maximum distance from work of a trip end point to be considered Work(miles)# -w 0.5
Home Threshold Maximum distance from home of a trip end point to be considered Home(miles)# -h 0.3
Minimum Interval-Minimum time difference between records to be used for distance computation(secs)# -i 1
Speed Threshold-Minimum speed of a record to be used for distance computation(m/sec)# -s 0.5
Speed Threshold2-Minimum speed of a record to be used for detecting NEPO stops(m/sec)# -s2 0.1
Minimum work duration(secs)# -d 3600
Update Time- Time between GPS records (secs)# -u 1