Using GPS Data Loggers
To Replace Travel Diaries
In the Collection of Travel Data

A Thesis
Presented to
The Academic Faculty

by

Jean Wolf

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Civil Engineering

Georgia Institute of Technology
July 2000

Copyright © 2000 by Jean Wolf
DEDICATION

This dissertation is dedicated to those who told me it did not matter if I finished (i.e., my family and my dogs), those who said I had to finish (i.e., my friends), and those who thought I would never finish (i.e., my colleagues). I would also like to recognize my educational clock as a motivational force in trying to get this finished before I turn 40. Finally, I dedicate this dissertation to my Grandma, who has always believed in me and inspired me.
# TABLE OF CONTENTS

DEDICATION.................................................................................................................................................. II

TABLE OF CONTENTS................................................................................................................................... III

LIST OF TABLES .......................................................................................................................................... V

LIST OF FIGURES ..................................................................................................................................... VI

LIST OF ABBREVIATIONS........................................................................................................................... VII

SUMMARY.................................................................................................................................................. X

INTRODUCTION.......................................................................................................................................... 1

BACKGROUND......................................................................................................................................... 6

TRAVEL SURVEYS .................................................................................................................................... 8
  Travel Diary Data Collection .................................................................................................................. 8
  Nonresponse and Accuracy Issues ........................................................................................................ 14
  Route Choice Data Collection .............................................................................................................. 18

GLOBAL POSITIONING SYSTEMS............................................................................................................ 20
  Overview of Global Positioning Systems ............................................................................................... 21
  Uses of GPS in transportation .............................................................................................................. 26
  Uses of GPS in Travel Behavior Surveys............................................................................................... 33

GPS TRAVEL BEHAVIOR DATA COLLECTION.................................................................................... 36

SELECTED GPS TRAVEL BEHAVIOR STUDIES.................................................................................... 36
  Lexington, Kentucky.................................................................................................................................. 37
  The Netherlands......................................................................................................................................... 39
  Austin, Texas............................................................................................................................................... 41
  Quebec City, Canada................................................................................................................................. 42
  Findings .................................................................................................................................................. 44

THE GEORGIA TECH GPS STUDIES ....................................................................................................... 46
  CETMS Overview .................................................................................................................................. 47
  SMARTRAQ Overview ............................................................................................................................ 50
  Findings and Recommendations .......................................................................................................... 52
  Summary ................................................................................................................................................. 58

USING GPS DATA TO DERIVE TRAVEL DIARY DATA ........................................................................ 60

THIS STUDY............................................................................................................................................ 61
  SCOPE.................................................................................................................................................... 64
  DATA ELEMENT ANALYSIS .................................................................................................................. 65
  Travel Survey Trip Data .......................................................................................................................... 65
### Table of Contents

**DATA PROCESSING** ........................................................................................................... 94

**CONCLUSIONS AND RECOMMENDATIONS** ....................................................................... 155

**CONTRIBUTIONS** ............................................................................................................. 153

**FINDINGS** .......................................................................................................................... 134

**OVERALL STUDY FINDINGS** ............................................................................................ 137

**MATCHING OF ELEMENTS** ............................................................................................... 73

**DATA COLLECTION** ........................................................................................................... 79

**PARTICIPANT FEEDBACK** ................................................................................................... 91

**SUMMARY** .......................................................................................................................... 77

**APPENDICES** ..................................................................................................................... 164

**GIS Components** ............................................................................................................. 72

**MATCHING OF ELEMENTS** ............................................................................................... 73

**Trip Origin Address and Start Time** ................................................................................... 74

**Trip Destination Address and Finish Time** ......................................................................... 75

**Travel Distance (and Route)** ............................................................................................... 75

**Trip Purpose** ...................................................................................................................... 76

**Travel Mode** ....................................................................................................................... 76

**Mode Details** ...................................................................................................................... 77

**SUMMARY** .......................................................................................................................... 77

**DATA COLLECTION PROCEDURES** .................................................................................. 83

**DATA COLLECTION RESULTS** ........................................................................................... 88

**PARTICIPANT FEEDBACK** ................................................................................................... 91

**TRIP DETECTION** ................................................................................................................ 97

**Trip Detection Process** ....................................................................................................... 97

**Trip Detection Results** ....................................................................................................... 102

**LAND USE AND ADDRESS ASSIGNMENT** ....................................................................... 113

**Atlanta Land Use Inventory Development** ......................................................................... 114

**Assignment of Land Use and Address** .............................................................................. 115

**Analysis of Consecutive Land Use and Address Assignment Results** ............................... 122

**TRIP PURPOSE DERIVATION** ........................................................................................... 125

**Trip Purpose Derivation Process** ....................................................................................... 128

**Trip Purpose Derivation Results** ....................................................................................... 129

**TRIP DISTANCE** ................................................................................................................ 131

**FINDINGS** .......................................................................................................................... 134

**COMPARISONS BY PARTICIPANT** ..................................................................................... 134

**OVERALL STUDY FINDINGS** ............................................................................................ 137

**KEY ISSUES** ....................................................................................................................... 142

**KEY OPPORTUNITY: MISSED TRIP IDENTIFICATION** ......................................................... 145

**Trip Chains** ....................................................................................................................... 146

**Passenger Pick Up or Drop Off** ......................................................................................... 149

**Round Trips** ....................................................................................................................... 150

**Short Distance Trips** ......................................................................................................... 151

**CONTRIBUTIONS** ............................................................................................................. 153

**CONCLUSIONS AND RECOMMENDATIONS** ....................................................................... 155

**CONCLUSIONS** ................................................................................................................ 156

**RECOMMENDATIONS** ....................................................................................................... 157

**A LOOK FORWARD** ........................................................................................................... 161

**APPENDICES** ..................................................................................................................... 164

**REFERENCES** ..................................................................................................................... 217

**VITA** ..................................................................................................................................... 224
LIST OF TABLES

Table 2-1: Typical Activity Data Elements ................................................................. 11
Table 3-1: Summary of Four Recent GPS Travel Behavior Studies ......................... 43
Table 3-2: Comparison of Two GPS-Enhanced Equipment Packages .................... 45
Table 4-1: Summary of GPS-based Data Collection Analyses ................................. 61
Table 4-2: Final Master Trip File Elements (TRPUNL91.DAT) ................................. 67
Table 4-3: Final Combined Work File -- Trip Elements Only (COMB92.DAT) ......... 67
Table 4-4: Place Codes (one-digit land-use codes) .................................................... 68
Table 4-5: Trip Purpose Codes .................................................................................. 69
Table 4-6: Key GPS Fields For Use in Travel Studies ............................................... 73
Table 4-7: Matching Travel Diary Elements with Available GPS or GIS Data ......... 74
Table 4-8: Inputs and Outputs of GPS Data Processing Analysis ............................ 78
Table 5-1: Equipment Used in Passive In-Vehicle GPS Data Loggers ..................... 81
Table 5-2: Equipment Package Components ............................................................ 82
Table 5-3: Example of GPS Data File Logged on Palm IIIx ................................... 86
Table 5-4: Summary of Participants by Home County and Equipment Package .... 90
Table 5-5: List of Participants .................................................................................... 90
Table 6-1: Preview of Reported versus Derived Trip Results .................................. 96
Table 6-2: Example of Trip Ends File ...................................................................... 99
Table 6-3: GPS Tests – Trips Detected versus Trips Reported ................................. 101
Table 6-4: Summary of Differences Between Reported and Detected Stops .......... 104
Table 6-5: Trip Destination to Next Trip Origin Analysis Results ............................ 123
Table 6-6: List of Standardized Land Uses and Corresponding Trip Purposes ....... 126
Table 6-7: Summary of Trip Purpose Derivation Analysis ...................................... 130
LIST OF FIGURES

Figure 2-1: Evolution of Travel Survey Instruments and Data Collection Methods...... 14
Figure 2-2: The Navstar GPS Operational Constellation ........................................... 22
Figure 2-3: GPS Position Determination Technique .................................................. 23
Figure 5-1: Completed Memory Jogger ...................................................................... 87
Figure 6-1: Examples of Trips Detected for Complete Match (Participant #15) .......... 105
Figure 6-2: Detected Missed Trips for Participants # 12 (left) and #20 (right) ........... 106
Figure 6-3: Examples of Undetected Passenger Pick Ups (Participant #17) ............. 108
Figure 6-4: Examples of Incomplete Trip Data Due to Loose Cabling ....................... 112
Figure 6-5: Example of Land Use Determination and Address Assignment (Home) ... 118
Figure 6-6: Example of Land Use Determination and Address Assignment (Retail) .... 120
Figure 6-7: Trip Ends for 151 Trips Remaining in the Analysis .................................. 121
Figure 6-8: Example of Complete Trip Data (left) vs. Missing Trip End Data (right) .. 124
Figure 7-1: Comparison Log for Participant #12 ....................................................... 135
Figure 7-2: Derived Trip Elements ............................................................................. 136
Figure 7-3: Example of Trip Reported but Not Detected .......................................... 137
Figure 7-4: Example of Four Missed Trips (Origin at Georgia Tech) ......................... 147
Figure 7-5: Maps of Four Missed Trip Destinations .................................................. 148
Figure 7-6: Examples of Trip Chains with Imbedded Passenger Pick Ups ............... 150
Figure 7-7: Example of Round Trip (Two Trips) Without Vehicle Stop ................... 151
Figure 7-8: Example of Short Trip Detected Not Reported ........................................ 152
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARC</td>
<td>Atlanta Regional Commission</td>
</tr>
<tr>
<td>ASCII</td>
<td>American Standard Code Information Interchange</td>
</tr>
<tr>
<td>ATIS</td>
<td>Advanced Traveler Information Systems</td>
</tr>
<tr>
<td>AVLN</td>
<td>Automatic Vehicle Location and Navigation</td>
</tr>
<tr>
<td>CAAA</td>
<td>Clean Air Act Amendments</td>
</tr>
<tr>
<td>CADAC</td>
<td>Computer-Assisted Data Collection</td>
</tr>
<tr>
<td>CAPI</td>
<td>Computer-Assisted Personal Interview</td>
</tr>
<tr>
<td>CASI</td>
<td>Computer-Assisted Self Interview</td>
</tr>
<tr>
<td>CATI</td>
<td>Computer-Assisted Telephone Interview</td>
</tr>
<tr>
<td>CCT</td>
<td>Cobb Country Transit</td>
</tr>
<tr>
<td>CDC</td>
<td>Centers for Disease Control</td>
</tr>
<tr>
<td>CETMS</td>
<td>Comprehensive Electronic Travel Monitoring System</td>
</tr>
<tr>
<td>CHASE</td>
<td>Computerized Household Activity Scheduling Elicitor</td>
</tr>
<tr>
<td>CPU</td>
<td>Central Processing Unit</td>
</tr>
<tr>
<td>CR</td>
<td>Carriage Return</td>
</tr>
<tr>
<td>CSI</td>
<td>Communication Systems International, Inc.</td>
</tr>
<tr>
<td>DCI</td>
<td>Differential Corrections, Inc.</td>
</tr>
<tr>
<td>DGPS</td>
<td>Differential Global Positioning System</td>
</tr>
<tr>
<td>DOQQ</td>
<td>Digital Orthophoto Quarter-Quadrangle</td>
</tr>
<tr>
<td>DOT</td>
<td>Department of Transportation</td>
</tr>
<tr>
<td>EDT</td>
<td>Eastern Daylight Time</td>
</tr>
<tr>
<td>EST</td>
<td>Eastern Standard Time</td>
</tr>
<tr>
<td>ETD</td>
<td>Electronic Travel Diary</td>
</tr>
<tr>
<td>FHWA</td>
<td>Federal Highway Administration</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>FOC</td>
<td>Full Operational Capacity</td>
</tr>
<tr>
<td>GGA</td>
<td>Global Positioning System Fix Data</td>
</tr>
<tr>
<td>GPS</td>
<td>Global Positioning System</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographic Information System</td>
</tr>
<tr>
<td>GNSS</td>
<td>Global Navigation Satellite System</td>
</tr>
<tr>
<td>GSA</td>
<td>GPS DOP and Active Satellites</td>
</tr>
<tr>
<td>GSV</td>
<td>GPS Satellites in View</td>
</tr>
<tr>
<td>ISTEA</td>
<td>Intermodal Transportation Efficiency Act</td>
</tr>
<tr>
<td>ITS</td>
<td>Intelligent Transportation Systems</td>
</tr>
<tr>
<td>IVIS</td>
<td>In-Vehicle Information Systems</td>
</tr>
<tr>
<td>LCD</td>
<td>Liquid Crystal Display</td>
</tr>
<tr>
<td>LF</td>
<td>Line Feed</td>
</tr>
<tr>
<td>MARTA</td>
<td>Metropolitan Atlanta Regional Transit Authority</td>
</tr>
<tr>
<td>MPO</td>
<td>Metropolitan Planning Organization</td>
</tr>
<tr>
<td>NCHRP</td>
<td>National Cooperative Highway Research Program</td>
</tr>
<tr>
<td>NDGPS</td>
<td>Nationwide Differential Global Positioning System</td>
</tr>
<tr>
<td>NMEA</td>
<td>National Marine Electronics Association</td>
</tr>
<tr>
<td>NPTS</td>
<td>National Personal Travel Survey</td>
</tr>
<tr>
<td>OEM</td>
<td>Original Equipment Manufacturer</td>
</tr>
<tr>
<td>PAPI</td>
<td>Paper and Pencil Interview</td>
</tr>
<tr>
<td>PC</td>
<td>Personal Computer</td>
</tr>
<tr>
<td>PCMCIA</td>
<td>Personal Computer Memory Card International Association</td>
</tr>
<tr>
<td>PDA</td>
<td>Personal Digital Assistant</td>
</tr>
<tr>
<td>RAM</td>
<td>Random-Access Memory</td>
</tr>
<tr>
<td>RMC</td>
<td>Recommended Minimum Specific GPS / Transit Data</td>
</tr>
<tr>
<td>RTCM</td>
<td>Radio Technical Commission Marine</td>
</tr>
<tr>
<td>SA</td>
<td>Selective Availability</td>
</tr>
<tr>
<td>SMARTRAQ</td>
<td>Strategies for Metropolitan Atlanta’s Regional Transportation and Air Quality</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>TAZ</td>
<td>Traffic Analysis Zone</td>
</tr>
<tr>
<td>UTC</td>
<td>Coordinated Universal Time</td>
</tr>
<tr>
<td>VMT</td>
<td>Vehicle Miles Traveled</td>
</tr>
</tbody>
</table>
SUMMARY

Transportation planners conduct travel surveys to collect data needed as input to travel demand models, which are used to forecast the need for road and transit enhancements and expenditures. Travel surveys are conducted using a combination of paper trip diaries for data collection and computer-assisted telephone interviews (CATI) for data retrieval. Recently, several pilot studies have combined Global Positioning System (GPS) technology with travel survey data collection to evaluate opportunities for improving the quantity and accuracy of travel data. These studies used GPS to supplement the traditional data elements collected in paper or electronic travel diaries.

This dissertation presents the results of a proof-of-concept study that examined a new opportunity for using GPS technology in travel surveys – that is, the use of GPS data collection to completely replace, rather than supplement, traditional travel diaries. The challenge in this new approach is to process the GPS data in a manner such that most of the traditional trip data elements can be derived. If this processing is done correctly and quickly, then the CATI retrieval call, which is used to collect household, person, vehicle and travel information, could be modified so that the only travel questions asked would be those necessary to clarify any uncertainties in the processed GPS travel data and to collect any missing elements. This process would not only improve the quality and robustness of the data, but would also greatly reduce respondent burden and telephone interview times.
This study was conducted using GPS data loggers to collect travel data in personal vehicles. The GPS data were then processed within a Geographic Information System (GIS) to derive most of the traditional travel diary elements. These derived data were compared with the data recorded on paper diaries by the same survey participants and were found to match or exceed the reporting quality of the participants. In fact, the GPS data revealed trips that were not recorded by the participants. Most importantly, this research demonstrates that it is feasible to derive trip purpose, the one element that has been considered “underivable”, from GPS data by using accurate GIS databases.
Urban and regional planners use travel demand models to estimate changes in transportation activity over time. These models predict the number of trips generated by households as a function of various demographic and socioeconomic considerations and also predict the number of trips attracted to various employment and commercial centers. These models also produce estimates for mode choice, distribution of trip destinations across the metropolitan region, and traffic volumes on various roads. Regional travel surveys, or travel diary studies, are used to collect the input and calibration data used to derive and validate travel demand models. Consequently, data collected from thousands of households across the region are analyzed to estimate current travel demand and to predict future travel demand. These regional travel estimates are also used to predict emissions from motor vehicles and serve as primary input data for air regional quality analyses. The accuracy and completeness of the household travel data obviously have a critical impact on model results.

In these regional travel diary surveys, randomly selected survey participants from the metropolitan region record information about each trip that they make during a one-day to one-week period. Participants log trip origin, destination, time, mode, and other characteristics into their paper travel diaries. Early travel studies used mail-out / mail-back survey methods in which the members of the population received survey documents
in the mail and were asked to fill out the survey and to return it by mail. Data entry personnel would then key in the data from the survey forms. One variation of this method was the use of telephone interviewers who would retrieve the diary data by calling the respondents and then recording the information on paper. These manual methods, often referred to as paper-and-pencil interviews (PAPI), were fraught with errors due to the omission of trips or trip elements, language issues, illegible handwriting, key entry errors, and a variety of other reasons.

Within the past several decades, these PAPI methods for travel diary data collection have been supplemented or replaced with computer-assisted-telephone interviews (CATI). CATI involves telephone interviewers using specialized software who call survey respondents and ask for details of all trips taken on a prior day. (The respondents often have previously been given a paper diary when recruited to assist in this recall process.) The software used by the CATI operators has imbedded logic checks that assist with the detection of missing or faulty data, but are still prone to recall errors and intentional omissions.

Most recently, computer-assisted-self-interview (CASI) methods, in which respondents are provided with an electronic data recorder, are being evaluated. Full automation of PAPI and CATI travel survey data collection processes will produce more data and more accurate data for use in transportation planning model development. Several recent studies have evaluated the potential of automated travel diaries using handheld computers to administer computer-assisted self-interview (CASI) travel-time collection of expanded travel data, and more accurate data with lower survey costs. Data
elements such as trip purpose, origin and destination names, and driver and passenger
names can all be selected from predefined lists, and trip start times and finish times can
be automatically collected with the push of a button. Other studies are beginning to
examine the use of the Internet for travel surveys [Polak 1999, RSG 1999].

Global Positioning System (GPS) technologies, which can provide second-by-
second kinematic position data with accuracies of three to five meters, as well as highly
accurate velocity and time data, introduce a whole new level of comprehensiveness and
accuracy when combined with electronic travel diaries. The potential advantages of using
GPS to supplement travel survey data collection are numerous: 1) trip origin, destination,
and route data are automatically collected without burdening the respondent for the data;
2) routes are recorded for all trips, allowing for the post-processing recovery of
unreported or misreported trips (including linked trips); 3) accurate trip start and end
times are automatically determined, as well as trip lengths; and 4) the GPS data can be
used to verify self-reported data.

An electronic travel diary (ETD) with GPS can more accurately capture all driver
and passenger travel activities than are typically captured using manual travel diary or
telephone survey methods. For each trip, this information includes trip mode, vehicle
identification, driver identification, passenger identification, driver and passenger trip
purposes, trip start time, finish time (or duration), origin location, destination location,
and distance traveled. In addition to these traditional elements, route choice and travel
speed can be captured, and functional classification of each link (with traffic conditions)
can be determined by tying the GPS data to a GIS database, all of which greatly
enhancing the original data collected. Two recent studies in Lexington, Kentucky and in the Netherlands have examined the use of electronic travel diaries with GPS as a more comprehensive travel data collection system [Wagner 1997, Draijer 2000], and the year 2000 U.S. NPTS (National Personal Travel Survey) may offer an electronic travel diary with GPS as part of the add-on study. In the recent Austin, Texas household travel survey, passive in-vehicle GPS systems were deployed to detect unreported trips [Casas 1999]. Similar systems are now being considered for supplementing traditional travel surveys to be conducted later this year in California and Washington. In Atlanta, the year 2000 household travel survey will include both an ETD with GPS as well as a passive in-vehicle GPS system to capture enhanced travel data.

All of the studies mentioned have used or will use GPS with either paper travel diaries or electronic travel diaries to supplement traditional data elements collected in the diary itself. This dissertation presents a new opportunity for using GPS technology in travel surveys – that is, to completely replace the travel diary with GPS data collection. The challenge is to process the GPS data in a manner in which most of the traditional trip data elements can be derived. If this processing is done correctly and quickly, the computer-assisted telephone interview, which is used to collect the household, person, vehicle and travel information, can be modified. Specifically, the collection of detailed travel data could be replaced by the clarification of any uncertainties in the processed GPS travel data and with the collection of any missing data elements. This process would not only improve the quality and robustness of the data, thus improving travel
demand modeling, but also would greatly reduce respondent burden and telephone interview times.

This dissertation first provides background information on traditional travel survey data collection methods, an overview of the Global Positioning System, and a summary of various applications of this technology in the transportation field. Next, several recent studies evaluating the use of GPS in travel surveys are reviewed, and GPS data collection study results and recommendations from research conducted at the Georgia Institute of Technology (Georgia Tech) are summarized. Then, an approach for collecting travel data using GPS technology and for processing the data to derive most, if not all, of the traditional travel diary elements is presented.

To demonstrate this approach, a proof-of-concept study was conducted at Georgia Tech in which GPS data loggers were used to collect travel data in personal vehicles of Georgia Tech staff and associates. The data have been processed within a Geographic Information System (GIS) to derive most of the traditional travel diary elements. The derived travel data were then compared with the reported travel data, which were recorded on paper diaries, by the same survey participants. This approach and the results of the study are reviewed in depth in this dissertation. In addition, recommendations for validating and completing these derived travel data within a CATI environment are offered.
CHAPTER II

BACKGROUND

During the 1990’s, the United States experienced a renewed interest in urban transportation planning and policy analysis as a result of the Intermodal Surface Transportation Efficiency Act (ISTEA) and the Clean Air Act Amendments (CAAA). New analytical tools and supporting data were needed to support required planning and policy analyses. The household travel survey has traditionally served as the primary source of personal travel behavior data. Other types of travel surveys are also used to obtain personal travel behavior information for infrequent behaviors and to supplement the household travel survey. Examples of these other surveys include on-board surveys (or transit ridership surveys), employer-based surveys, and roadside origin-destination surveys [Lawton 1996].

Data collected in these surveys are then used by metropolitan planners as input to travel demand models that are, in turn, used to estimate changes in transportation activity over time. These models predict the number of trips generated by households as a function of various demographic and socioeconomic considerations and also predict the number of trips attracted to various employment and commercial centers. Estimates for travel mode choice, distribution of trip destinations across the metropolitan region, and traffic volumes on various roads also come from these travel demand models. These regional travel estimates are then used to evaluate alternative transportation plans and
policies, and to consequently substantiate high-cost transportation infrastructure investments. In addition, the travel estimates are also used to predict emissions from motor vehicles and serve as primary input data for air regional quality analyses.

Traditional data collection methods such as paper-and-pencil interviews, which have been fraught with data recording and transcription errors, are being replaced by automated methods such as computer-assisted telephone interviews and computer assisted personal interviews, and computer-assisted self interviews. The availability and affordability of GPS technology has introduced another travel data collection automation tool to supplement other methods in the collection of the basic travel elements: trip origin and departure time, trip destination and arrival time, trip purpose, and travel mode. GPS technology provides the capability to capture accurate spatial and temporal details of travel for every second of every trip. This capability offers great opportunities for expanded data sets at higher accuracy levels while simultaneously decreasing respondent burden.

One data element that has not been collected in various travel surveys is route choice. The burden of collecting such detailed information from survey participants for each trip made using traditional data retrieval methods has always been considered prohibitive. However, with advances made in the GPS over the last two decades, including improved functionality and accessibility, along with user-friendly interfaces and reasonable costs, planners can now consider GPS as an effective means to collect detailed route data without increasing user burden.
Although the use of GPS in travel survey studies is relatively new, transportation engineers and planners have already found many other spatial and temporal data collection tasks made easier through the application of GPS solutions. Transportation applications with GPS include numerous surveying applications, development or improvements of Geographic Information Systems (GIS), maintenance and inventory applications, automatic vehicle location and navigation applications, and travel time / travel speed studies.

Travel Surveys

Travel surveys are used worldwide to collect personal travel behavior data. This overview of travel surveys will begin with a brief review of travel survey data collection, will continue with an examination of nonresponse issues and “solutions”, and will conclude with a summary of route choice data collection studies and benefits.

Travel Diary Data Collection

Travel diaries are the standard method for collecting household travel activity for studies conducted at metropolitan, state, and national levels. In these studies, households are selected to provide a representative sample of the study area. The primary purpose of the diaries is to capture a “snapshot” of travel for a day or range of days for each member of a recruited household. These travel details are collected in tandem with information about the household, each person in the household, and each vehicle available to the household. These sampled household characteristics and travel activities are then expanded to represent travel patterns for the study area. Simply stated, these expanded
patterns are used as input (in the form of trip productions and attractions for each traffic analysis zone) in trip generation models.

Travel surveys are conducted at the household level because total travel demand is better understood at this level rather than at the person level, given the interdependencies of residents within a home with respect to trip needs and scheduling. There are three general categories of data collected in travel surveys: household information, person information, and activity or travel information for a given day or multi-day period. The household data includes information about the physical household (such as location, type of dwelling, and length of residence), the household vehicles (including the number of vehicles available to the household and vehicle details), and the household occupants (including the number of people in the household and total household income). The personal information data elements include socio-economic and personal details for each household member (including relationship to other household members, gender, ethnicity, and birth date), employment details (such as employment status and occupation, weekly work schedule, and transportation taken to work), and education.

The design of travel diaries has evolved over the past several decades as researchers and practitioners have come to realize that trip-oriented diaries do not match well with how people think about the daily activities. It is this incongruity that has led to a significant underreporting of trips. Therefore, where early travel surveys focused on the trips between activities, more recent surveys have focused on the activities at the trip ends. Surveys that focus on out-of-home activities are commonly referred to as activity surveys, as opposed to surveys that collect details on activities that occur both in and out
of the home, which are more commonly called time-use surveys. Of course, activity and time-use surveys require much more detail from the respondents and therefore often result in lower response rates and higher data collection costs.

The most recent trend in travel survey instrument design are “place-based” surveys, which blend the trip-based survey and activity-based survey to focus the respondent on their movements from one place to the next throughout the survey period. This method has been found to improve the recollection of incidental trips. Another method to reduce respondent burden is the use of a reduced diary, more commonly known as a memory jogger. The memory jogger is used instead of a detailed diary when CATI retrieval calls can be made within a day or two after the study day. It is thought that the few elements recorded should provide sufficient information to assist the respondent to remember all trip details given the short recall period. (Examples of these four types of survey instruments are contained in Appendix A.) The evolution of travel diary formats reflects the need for better information regarding how people schedule their activities within a day or multi-day period so that modelers can better understand travel behavior.

The actual travel diary or memory jogger is used by each respondent to record the details of travel for the given survey period. Both activity and place-based surveys are attempts to orient the survey form itself towards a more intuitive way of thinking about daily travel. The activity information collected includes activity details, travel modes used, and travel mode details such as personal vehicle and parking information, transit details, walk or bike information, and shopping activities. Table 2-1 provides a summary of the standard activity data elements [Stecher 1996].
<table>
<thead>
<tr>
<th>Category</th>
<th>Data Element</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity</td>
<td>Type of activity</td>
<td>The level of detailed categories has been expanding.</td>
</tr>
<tr>
<td></td>
<td>Start and end time of activity</td>
<td>Travel time is calculated as the time between the end of one activity and the start of the next</td>
</tr>
<tr>
<td></td>
<td>Name of place where activity took place</td>
<td>Important for geocoding, especially when precise address information is not recorded</td>
</tr>
<tr>
<td></td>
<td>Address where activity took place</td>
<td>Cross-streets are often all that respondents can report</td>
</tr>
<tr>
<td></td>
<td>Frequency activity takes place</td>
<td>This can be an awkward data element to collect, particularly for routine activities</td>
</tr>
<tr>
<td>Mode</td>
<td>Mode</td>
<td>Include walk and bicycle prominently in presenting options</td>
</tr>
<tr>
<td>Personal Vehicle</td>
<td>Use of household vehicle or other vehicle</td>
<td>Linking the vehicle used for trips to the detailed information collected about each household vehicle permits valuable air quality analyses</td>
</tr>
<tr>
<td></td>
<td>Driver or passenger</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vehicle occupancy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Type of parking</td>
<td>Refers to parking in a lot, on the street, etc.</td>
</tr>
<tr>
<td></td>
<td>Parking cost</td>
<td>Collects the amount and the interval</td>
</tr>
<tr>
<td></td>
<td>Parking payment method</td>
<td>Captures parking costs paid for by validation, by the employer, out of pocket</td>
</tr>
<tr>
<td>Transit</td>
<td>Transit fare</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transit payment method</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Location of access/egress</td>
<td>May be captured in several ways</td>
</tr>
<tr>
<td></td>
<td>Mode of access/egress</td>
<td>May be captured in several ways</td>
</tr>
<tr>
<td></td>
<td>Wait time</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Number of transfers/other transfer data</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Availability of personal vehicle</td>
<td>Must define “availability”</td>
</tr>
<tr>
<td>Walk</td>
<td>Distance</td>
<td>Can be captured in blocks or miles</td>
</tr>
<tr>
<td></td>
<td>Availability of personal vehicle</td>
<td>Must define “availability”</td>
</tr>
<tr>
<td>Bicycle</td>
<td>Use of bicycle lanes or paths</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Means of securing bicycle at destination</td>
<td></td>
</tr>
<tr>
<td>Shopping</td>
<td>Mall or shopping center</td>
<td>Captures if shopping occurred in mall or shopping center</td>
</tr>
<tr>
<td></td>
<td>If in a mall or shopping center, number of stores visited</td>
<td></td>
</tr>
<tr>
<td></td>
<td>If in a mall or shopping center, did respondent eat</td>
<td></td>
</tr>
<tr>
<td>Commercial Trips</td>
<td>Commercial vehicles stopping at the household</td>
<td>Captures trip attractions to the household</td>
</tr>
</tbody>
</table>
For decades, the household travel survey was administered using mail-out/mail-back travel diaries. This paper-and-pencil interview method (PAPI) typically involves a series of mail-out reminders and/or phone calls made to hundreds or thousands of households within a given region. Each recruited household is asked to record details for each trip made by each household member on a given day or range of days. For each trip, a respondent is asked to record detailed information, including origin and destination names and addresses, trip purpose, type of location (i.e., land use), mode of transport, driver and passenger names (if the mode is personal vehicle), trip start and finish times, and trip length. With this high respondent burden, it has been common for participants to misreport or omit trips, to significantly round off trip times and lengths, and to omit origin and destination addresses.

To overcome manual recording, reporting, and recollection errors, automation of travel diary processes is continually under evaluation to the extent it is feasible, practical, and cost-effective. Manual methods for data collection have given way to computer-assisted methods (i.e., CADAC: computer-assisted data collection), including computer-assisted telephone interviews, personal interviews and self-interviews. Starting in the 1980’s, computer-assisted telephone interviews (CATI) were combined with PAPI methods in an attempt to improve data accuracy and completeness. CATI involves telephone interviewers using specialized software who call survey respondents and ask for details of all trips taken on a prior day. The respondents have typically been given a paper diary when recruited to assist in this recall process.) However, these combined
methods are also subject to participant recording errors and omissions, and to operator data entry errors. Furthermore, CATI methods are subject to respondent recall problems.

Most recently, computer-assisted personal interviews (CAPI) and computer-assisted-self-interview (CASI) methods are being evaluated as a means to capture more travel data and more accurate travel data. Personal interviews are normally cost-prohibitive for large-scale studies and are therefore not typically used in large travel surveys. With self-administered electronic travel diaries, data elements such as trip purpose, origin and destination names, and driver and passenger names can all be selected from predefined lists, and trip and finish times can be automatically collected with the push of a button. The complex nesting of lists necessary to support detailed activity descriptions, which is impossible to support in paper diaries, can be easily handled by a user-friendly interface.

There are numerous references that review the traditional diary instruments and the transition to automated data collection methods [Purvis 1999, Kalfs 1997, Sarasua 1996]. Two recent European research projects examined the use of handheld PC's and Internet-based web sites for long distance travel data collection [Haubold 1998, Polak 1999]. Additional U.S. studies have examined the opportunities presented by using GIS technologies in travel studies [Abdel-Aty 1995, Greaves 1997]. Figure 2-1 shows the evolution of travel survey instruments and data collection methods.
Nonresponse and Accuracy Issues

Given the critical role of household travel data in policy decision-making and transportation infrastructure development, data accuracy and completeness are essential. Yet, due to sampling errors, nonresponse errors, and inaccurate reporting due to manual data collection and recall retrieval processes (such as paper diaries and telephone interviews), travel data quality has long been suspect. Although sampling errors are typically addressed through better sampling designs and larger sample sizes, survey biases associated with unit (or household) nonresponse and item nonresponse are much more difficult to reduce [Richardson 1996].

Unit nonresponse occurs when households decline to participate in the study. Often, these nonrespondents come from segments of the population that are significantly different from the sampled population both in terms of demographics and in travel behavior. Some researchers theorize that households that do not answer their telephone
may not be home because they make many trips and are not home often. Exclusion of these households would result in an underestimation of travel. Conversely, if households that travel very little, such as households without phones, are excluded from the study, then an overestimation of travel will occur if forecasts are based on sampled households only.

Several methods are used in an attempt to alleviate unit nonresponse. The use of incentives and/or follow-up procedures during the recruitment process can improve response rates slightly. Infill sampling may also be used and involves the targeted recruitment of participants in traditionally underrepresented strata. Finally, nonresponse interview surveys, often conducted as personal interviews, may be used to gain insight with respect to the reasons for nonresponse, which can in turn be used to improve response rates for the next wave or survey on the same target population.

Item nonresponse occurs when households agree to participate, but then fail to report all details, which could include critical household socioeconomic information, such as household income, travel information for one or more household members, or entire trips or trip details for individual household members. A problem related to item nonresponse is item misreporting, which occurs when the respondents mistakenly or intentionally report inaccurate household or travel data. Zmud and Arce [1997] report that item nonresponse is the result of five major phenomena: knowledge gaps or memory lapses, comprehension failure, perceived and/or real burden, privacy concerns, and deliberate misreporting. A typical knowledge gap is the lack of detailed address information for locations other than home and work. Other examples of inaccurate
reporting occur when respondents round or estimate trip start and departure times to the nearest 5-minute, 15-minute, or half hour mark; and round trip distances to the nearest 5 or 10-mile increment.

Perhaps the most examined problem caused by item nonresponse in travel demand modeling is lower trip rates due to unreported trips, which results in an underestimation of travel demand when aggregated to the regional level. Paper survey participants may not keep accurate records, which results in misreported or underreported travel, including the omission of entire trips. In addition, the need for more detailed activity data over longer periods of time has contributed greatly to increased respondent burden, which often results in decreased response rates as the respondent learns what is involved in reporting each trip. Respondents may decide that some short trips or trip types are not important to the study or are not really trips at all. For example, passenger (or article) pickup and drop-off are frequently not considered as trips.

There are numerous approaches used to prevent, or at least reduce, item nonresponse and misreporting. Much work has focused on improving the design of travel survey instruments (recall the evolution of trip-based to activity-based to place-based diary orientation) and on improving travel data collection methods (such as seen in the migration from paper to telephone to computerized methods). CATI and CASI methods of data collection allow for probing for missing data and validity checks during the data collection process. In addition, computer coding of responses and editing checks further improve the accuracy of the collected data.
Once the data are collected, there are several methods used to correct the data. One simple method is to simply ignore the nonresponse, which reduces the sample size and could therefore increase the nonresponse bias [TSM 1996]. Within the U.S., the use of weighting factors, such as socioeconomic expansion factors, non-reporting correction factors, and nonresponse correction factors, are commonly used. Imputation, which is a process whereby the value of a missing data item is somehow inferred based on observed data using statistical relationships, is finding greater acceptance among international travel behavior researchers [Polak 1997]. Nonresponse follow-ups may also be conducted; research suggests that socioeconomic and travel characteristics of respondents that reply to each wave of follow-ups are successively closer to the characteristics of those who never respond [Richardson 1995, 2000].

Finally, while much attention has been appropriately focused on improving item nonresponse and misreporting, it is important to note that these problems pertain to a somewhat constant set of travel data elements. The fact remains that, even as travel behavior data needs continue to increase, the level of burden placed on respondents to capture and provide these data is still a limiting constraint. Until automated methods are fully introduced that capture the desired data with high levels of accuracy and completeness while containing or reducing respondent burden, additional information such as attitudinal information and trip/activity scheduling thought processes cannot be captured from travel survey participants.
Route Choice Data Collection

Route choice data can be used to improve the imbedded traffic assignment algorithms and to validate and/or calibrate travel demand models. These data can also be used to model route choice in other transportation applications, such as those interested in the impact of advanced traveler information systems (ATIS) or in-vehicle information systems (IVIS) on driver route choice behavior. In mobile source emissions modeling, route choice plays a significant role in allocating emissions to the appropriate road segments. Finally, route choice data can be used to determine route distance, which in the case of vehicle trips, is the equivalent of Vehicle Miles Traveled (VMT), an important input in several transportation models.

Although route choice data are useful for a variety of purposes, traditional travel diary data collection methods, including both PAPI and CATI methods, have not typically collected route data due to the additional respondent burden. Instead, survey respondents are usually asked to record or report the street address of each trip origin and destination. Within the transportation planning process, these addresses are then associated with their corresponding traffic analysis zones, from which travel routes are generated using shortest path (or lowest cost) algorithms.

Some travel diary studies have collected route choice data. Mahmassani et al. [1993] reported success in route choice collection by implementing a two-staged survey in which the first stage screened for respondents willing to provide additional information in the second stage. Abdel-Aty et al. [1995] summarized numerous studies collecting route choice data for the purpose of analyzing the impact of traffic information on route
choice behavior. However, all of these studies used self-reported manual procedures for capturing actual route choice. Trip distance is more commonly collected in travel surveys, although even this element is not consistently collected or reported.

Another method used to capture route choice decisions employs stated preference surveys, in which respondents are asked how they would behave (i.e., what route would they take) under a specific set of conditions [Lee-Gosselin 1996, Polak 1995]. These surveys can be administered via manual or automated methods. One obvious advantage of this method is that it does not depend on recall; however, it may also not reflect real-world route choice decisions.

Recently, a few surveys have collected route choice data using GPS technology. A 1996 Michigan study used in-vehicle GPS receivers to compare route choice behavior and perceptions of three different in-vehicle navigation-assistance systems [Eby 1997]. Other studies conducted in Lexington, Austin, Quebec City, and the Netherlands have tested the use of GPS in collecting route choice data as a supplement to traditional travel behavior data collection methods [Wagner 1997, Casas 1999, Doherty 1999, Draijer 2000]. Finally, researchers at the Georgia Tech, in Atlanta, Georgia, have developed a comprehensive vehicle instrumentation package for monitoring individual trip-making behavior [Wolf 1999a]. The Georgia Tech instrumentation package includes a handheld computer for logging general trip information, a GPS device to capture all spatial and temporal coordinates, and an onboard engine monitoring system to capture a variety of vehicle and engine parameters on a second by second basis. This package, along with two other GPS-enhanced equipment packages, will be used for various sub-samples in
the Year 2000 Atlanta Metropolitan Household Travel Survey. The Lexington, Austin, Quebec City, Netherlands, and Georgia Tech projects are reviewed in detail in Chapter 3.

**Global Positioning Systems**

The Global Positioning System (GPS) is a satellite-based positional system initiated by the U.S. Military in the 1970’s that achieved full operational capacity (FOC) in 1995. This system allows receivers located anywhere on the earth’s surface to compute their position, velocity, and time. Both civilians and the military use this positioning system for a variety of land, sea, and air applications. Technology has improved rapidly to meet increasing demands for GPS receivers and antennas for a large range of uses.

The field of transportation has many obvious needs for an accurate positioning system, and transportation researchers and practitioners have quickly adopted GPS technology as a relatively low-cost, high accuracy solution for meeting various positioning requirements. These applications begin with traditional transportation areas such as point surveying; GIS development; and Department of Transportation (DOT) maintenance and inventory applications; and have rapidly moved into newer areas such as automatic vehicle location and navigation (AVLN) applications; travel time studies; enhanced mobile source emissions modeling; and travel behavior surveys. Much attention has been given in the last few years to the possibilities offered by GPS with respect to improving traditional travel surveys, an area which has been subject to both nonresponse and accuracy problems.
Overview of Global Positioning Systems

The GPS consists of three segments: the space segment with 24 (or more) satellites providing worldwide coverage, the operational control segment that monitors and controls the space segment, and the user segment [McDonald 1999]. The space segment contains the 24 satellites that orbit the globe every 12 hours and are non-symmetrically distributed in six 55° orbit planes of four satellites each; this coverage pattern ensures that six to eleven satellites are always in view at any point on the globe’s surface (see Figure 2-2). The operational control segment consists of a master control station located in Colorado Springs, Colorado and five monitor and four upload stations that are geographically distributed. This control segment accurately tracks the GPS satellites, updates each satellite’s ephemeris data and clocks, and monitors the health and status of the satellites.

The user segment receives the data contained in the satellites signals and uses the data to compute position, velocity, and time. Users of GPS today include both the military and civilians who require positional information. Military applications include target acquisition, missile guidance, sensor emplacement, coordinate bombing, and remotely piloted vehicle operations. Examples of civil applications include en-route navigation, flight guidance, fleet management, search and rescue, recreation, theft deterrence, and mapping.
The GPS computes ground position by first measuring the signal travel times between a group of satellites and a ground-based receiver. Since radio signals travel at the speed of light, these travel times can be used to calculate the distances from receiver to the satellites. Finally, the position of the ground receiver’s antenna is calculated using trilateration to solve four unknowns: the x, y and z coordinates and the difference between the satellites’ clocks and the receiver’s internal clock [Czerniak 1997]. Figure 2-3 shows the variables used in GPS-based position determination using four satellites.
The calculations involved for solving the user’s position coordinates \((X_u, Y_u, Z_u)\) and clock bias \((\Delta T_B)\) are contained in the following four equations:

\[
\begin{align*}
(X_1 - X_u)^2 + (Y_1 - Y_u)^2 + (Z_1 - Z_u)^2 &= (\tilde{R}_1 - c\Delta T_B)^2 \\
(X_2 - X_u)^2 + (Y_2 - Y_u)^2 + (Z_2 - Z_u)^2 &= (\tilde{R}_2 - c\Delta T_B)^2 \\
(X_3 - X_u)^2 + (Y_3 - Y_u)^2 + (Z_3 - Z_u)^2 &= (\tilde{R}_3 - c\Delta T_B)^2 \\
(X_4 - X_u)^2 + (Y_4 - Y_u)^2 + (Z_4 - Z_u)^2 &= (\tilde{R}_4 - c\Delta T_B)^2
\end{align*}
\]
Using a GPS receiver, the user measures four pseudorange \( (\tilde{R}_i) \) values:

\[
R_i = \tilde{R}_i - c\Delta T_B
\]

in which \( R_i \) is the true range, \( \tilde{R}_i \) is the pseudorange, \( c\Delta T_B \) represents the clock bias range error, \( \Delta T_B \) represents the user clock bias or timing error, and \( c \) is speed of light.

An alternate method for visualizing the position calculation process is to consider the sphere around each satellite that describes the calculated distance from the satellite to the receiver. The receiver must lie somewhere on the surface of this sphere. Three satellites in view provide three spheres. The intersection of these three spheres yields two points, one on the surface of the earth and another one in outer space, which is automatically discarded. To determine and eliminate any clock drift, four satellites must be in view to compute a three-dimensional position. A two-dimensional \((x, y)\) position can be calculated by using the third satellite to resolve the clock differential rather than for determining the \(z\) coordinate.

GPS data are subject to numerous sources of error, including satellite orbit errors, satellite clock errors, receiver errors, tropospheric and ionospheric errors, multi-path errors, and Selective Availability (SA). SA is the intentional introduction of satellite ephemeris errors by the U.S. Department of Defense. SA degrades the precision of GPS accuracy up to 100 meters for real-time, non-military users. SA is by far the biggest source of errors in GPS positioning [Zito 1995]. Some GPS receivers support
differential correction, a feature that eliminates both SA-introduced errors and atmospheric errors in either a real-time or post-processing mode and results in accuracy levels of 3 to 10 meters typical. Uncorrected GPS data are referred to as autonomous or raw data.

Differential correction is the technique of reducing GPS errors by collecting data with two units simultaneously. A base station receiver and antenna are placed at a known position. The satellite data collected at this known position allow for the computation of corrections for the GPS signals. These corrections are then applied to the data collected by another receiver, eliminating the effects of SA and other global errors. Differential corrections can either be sent by the base station to the receiver in real time through a satellite or radio link (real-time DGPS) or can be stored on the base station and applied later to the receiver’s data using a process known as post-processing differential correction (post processing DGPS). Issues with accuracy will be covered in more detail in the GPS studies presented in Chapter 3.

There are hundreds of commercially available GPS receivers and antennas on the market today. These products have prices ranging from the low $100s to the tens of thousands of dollars; prices typically reflect the level of precision available, where the cheapest receivers produce simple uncorrected position information (30 to 100 meter level accuracy with SA on and 5 to 20 meters with SA off) and the most expensive survey-quality receivers can provide sub-centimeter level accuracy. These receivers support a variety of uses (e.g., marine, aircraft, vehicle, person) and correction modes (e.g., autonomous, real-time DGPS, post-processing DGPS, inverse DGPS), come in a
range of forms (e.g., PCMCIA cards, OEM boards, sensors, handheld computers), and have a wide selection of antennas that also are available in an assortment of forms, features, and configurations. This vast array of options has been developed to meet a maturing GPS marketplace as user groups increase their acceptance of this relatively new positioning technology.

As a final note on the Global Positioning System, on May 1, 2000, U.S. President Clinton, without any advanced notification, announced the immediate termination of Selective Availability. The end of this intentional degradation of GPS spatial accuracy offers incredible opportunities for the use of low-cost GPS technology to obtain fairly accurate (10 to 15 meters typical) spatial data. Since this event occurred only a few weeks prior to the presentation and defense of this dissertation, with no advanced notice and after all data had been collected, the content of the dissertation has not been modified to reflect this significant event. However, the author does offer a few paragraphs in the last chapter “Looking Forward” that discuss how this event would impact a similar study conducted in a non-SA world.

Uses of GPS in transportation

As new technologies are introduced throughout the transportation field, researchers and practitioners are automating numerous manual data collection processes. These advances generally reduce labor costs and manual recording errors. Automation of data processing and data analysis reduces data transcription errors and further reduces overall research costs when employed for large datasets in ongoing studies. The detailed data now available through new technologies are enabling innovative modeling and analysis
realms. Areas in transportation that have been swept up in the automation wave include traditional areas such as traffic signal timing and synchronization, travel speed studies, and traffic counts. In addition, emerging areas such as Intelligent Transportation Systems (ITS) and complex computer modeling of transportation systems and vehicle emissions, which allow for policy decision analysis, are founded in automation. The use of such models places high demand on both the quantity and quality of transportation data needed for model calibration [Williams 1997]. GPS is one such technology that is revolutionizing how and what transportation data are collected.

When GPS was first introduced in the 1980’s, various state Departments of Transportation (DOTs), commercial businesses, and federal agencies have explored the use of GPS in traditional transportation spatial data collection activities, specifically those involving high-precision geodetic survey tasks. However, it soon became apparent that GPS could also be used for other spatial data collection tasks with varying accuracy requirements. Over the past decade, GPS has found its way into a variety of transportation data collection tasks, including numerous surveying applications, development or improvements of Geographic Information Systems (GIS), maintenance and inventory applications, automatic vehicle location and navigation applications, travel time / travel speed studies, and travel behavior surveys. GPS is now such an integral tool in DOT data collection activities that Transportation Secretary Rodney Slater has identified GPS as one of the department’s Flagship Initiatives. Here are a few key points in the department’s Two Year Action Plan for the DOT Flagship Initiative [DOT 1999]:

27
• The Presidential Decision Directive on U.S. GPS Policy (PDD NSTC-6) of March 28, 1996 directs the Department of Defense (DoD) to acquire, operate, and maintain the basic GPS.

• PDD NSTC-6 directs the Department of Transportation (DOT) to serve as the lead agency within the U.S. Government for all Federal civil GPS matters.

• The DOT’s plan for this initiative is to develop and implement reliable GPS augmentation systems for navigation, and to promote GPS signal improvements and international acceptance of GPS in navigation.

• A properly protected, GPS-based transportation infrastructure will improve the measured safety and efficiency of all modes of transportation.

It is beyond the scope of this dissertation to report on all existing applications within the transportation arena; however, a few examples are provided in the following sections to illustrate the breadth of transportation applications that are currently using GPS. Additional examples can be found in NCHRP Synthesis 258, which is dedicated to the topic of applications of GPS for DOTs [Czerniak 1998].

Surveying Applications

Control points are needed in numerous transportation projects. The surveying of such points traditionally required direct referencing into a national or state geodetic referencing system. With GPS referencing, line-of-sight between adjacent points is not required, which greatly reduces labor and infrastructure costs while improving accuracies significantly. Examples of control point applications include corridor and project control, mapping control, construction reference control, structural control, cadastral surveys (for
right-of-way and property boundaries), and airborne GPS / photogrammetry control [Czerniak 1998]. One high-profile example of GPS use for structural control occurred during the construction of the Chunnel – the tunnel under the English Channel that connects England and France. GPS was used in this project to establish a unified control reference network for each end of the tunnel.

**GIS Applications**

The underlying accuracy of any existing GIS is highly dependent on the types of materials used in its creation. Source maps often have ambiguous spatial accuracy, and techniques such as “rubber sheeting,” which attempt to merge different source maps, further distort the data. Given widely-available and inexpensive GIS base maps such as the Census Bureau TIGER files, which have a reported accuracy of 30 to 50 meters, it is easy to see that GPS data with accuracies of 3-5 meters offer great improvements. Numerous agencies and businesses are using GPS technology to improve the spatial integrity of their existing databases or to develop new GIS foundations.

**Maintenance and Inventory Applications**

As state DOTs realized the opportunities to associate asset information with location data, several studies were conducted to prove how easily maintenance and inventory applications using GPS combined with a GIS. NCHRP Reports 334 and 361 were released in the early 1990’s; these reports defined and demonstrated the uses of advanced data acquisition technologies, such as GPS, for maintenance management systems. In one field test in Arizona, a GPS manufacturer demonstrated how a GPS receiver with DGPS corrections could be used to collect digital roadway alignment data
and selected features including signs with spatial accuracies of 2 to 5 meters [Hyman 1993]. The Virginia DOT recently initiated a 10-year GIS strategic plan that includes the collection of highway inventory data using DGPS; preliminary tests of representative data collection efforts for point, line, and polygon data resulted in point-level accuracies of approximately 2 meters [Brich 1997]. Another recent application of DGPS in-vehicle data capture has been for the collection of road grade data; one research project conducted in 1996 in Australia used GPS techniques to collect road slope data for a 3010-kilometer highway upon which the World Solar Challenge road race is held every three years [Han 1999].

**Automatic Vehicle Location and Navigation Applications (AVLN)**

The ability to remotely locate a vehicle on any road network is dependent on accurate spatial data, a communications link, and a reasonably accurate base map. As the cost of GPS technology continues to decrease, there has been an explosion of AVLN applications that use GPS to locate lost or stolen vehicles, provide navigational assistance and/or traffic information to a driver, dispatch emergency personnel in the case of an accident, and monitor and manage commercial fleets more efficiently. General Motor’s OnStar in-vehicle computer system is considered to be state-of-the-art technology. OnStar’s menu of services includes Air Bag Deployment Notification, Convenience Services, Concierge Services, Emergency Services, Remote Door Unlock, Roadside Assistance, Route Support, Theft Protection, Med-Net, Accident Assist, and Ride Assist – all of which are enabled by an in-vehicle CPU, a cellular phone and modem, a speaker phone, a GPS receiver, and a set of automatic controls and sensors [Forman 2000]. Two
projects currently operating in Minnesota and upstate New York have implemented GPS-equipped systems with a cellular link and an accelerometer that automatically detect accidents and transmit a request for assistance [Albert 2000]. In 1992 Avis participated in the TravTek ITS study in Orlando, where 75 of its vehicles were tracked on a minute-by-minute basis at the Traffic Management Center to determine patterns of vehicle use [Peters 1993].

AVLN applications can be easily extended to personal tracking and navigation systems. Arkenstone, a non-profit organization that offers information access through technology for people with vision and reading disabilities, began research on a personal navigation system for the visually impaired. The resulting product, named Strider, consists of GPS and DGPS receiver and antennas, a laptop with a GIS, a small handheld keypad, and a voice synthesizer – all contained within a backpack. As a person walks around with the system, it provides information regarding the person’s location and their surroundings [LaPierre 1998]. GPS-enhanced personal monitoring and tracking systems are now commonly used for people on parole or house arrest. Even cellular phone manufacturers are considering GPS positioning solutions. In 1996 the U.S. Federal Communication Commission issued its E911 requirement, mandating that all cellular phones must convey the caller’s position to within 125 meters 67% of the time by October 31, 2001 [Zagami 1998]. Although the original regulation required the builders of cellular telephone networks to use signals from nearby cell-phone towers to triangulate to find the caller’s approximate position, a recent modification to the regulation
accelerated the timetable for implementation and provided the technology alternative of using GPS technology to determine caller position [Schwartz 1999].

**Travel Time Studies / Traffic System Performance**

Travel-time surveys are a common method of collecting performance data for the assessment of traffic systems. Prior to GPS availability, data collection activities were primarily conducted either as floating car / chase car studies in which observers recorded travel data using paper, laptop computers or other data recorders, or as stationery studies in which teams of observers located at fixed distances on the network conducted license plate, origin-destination, or path-trace surveys. These methods were very labor intensive, often resulted in averaged speed data on spot-speed data only, and provided an assessment of the system at a time in the past. However, in-vehicle GPS data collection systems are being evaluated as an efficient and economical tool for gathering comprehensive travel and traffic data [Quiroga 1999, Choi 1998]. Taken one step further, when this GPS data is transmitted in real-time to a central location (such as a traffic management center) and integrated into a GIS, real-time congestion monitoring of a given transportation network and ATIS applications could be realized [D’Este 1999].

**Enhanced Mobile Source Emissions Modeling**

Current mobile source emissions models are based on a set of emission factors and vehicle activities; these activities are high-level averages of travel speed, vehicle characteristics, and vehicle operating modes. Studies in California have used GPS to collect second-by-second vehicle speed so that functional relationships could be developed between these velocity statistics and the network traffic data collected by road
sensors [Barth 1996]. The goal of these studies is to obtain real-time localized emissions estimates by improving these functional relationships and combining them with given speed-flow-density measures of freeway traffic. Researchers in Atlanta are working on another project whose goal is to improve existing emissions models. The Comprehensive Electronic Travel Monitoring System (CETMS) will simultaneously collect GPS data with engine and vehicle activity data recorded by an on-board engine monitoring device so that insights can be gained with respect to the driver-vehicle interactions that impact vehicle emissions [Wolf 1999a]. Data collected by the CETMS will help model developers understand the linkages between travel decisions and vehicle operating conditions. This, in turn, will lead to the development of improved algorithms for predicting emissions as a function of vehicle characteristics, on-road traffic conditions, and household/driver demographics.

**Uses of GPS in Travel Behavior Surveys**

As paper-based methods for travel behavior data collection such as paper-and-pencil interviews are being replaced by automated methods such as computer-assisted telephone interviews and computer assisted self-interviews (which can be administered via handheld, laptop, or desktop PC with local or web-based applications), researchers and practitioners continue searching for the means to collect more and more accurate travel data. Global positioning systems can add significantly improved temporal and spatial details to travel diaries by capturing exact origin and destination coordinates, exact trip start and finish times, and actual route choice. From these data elements, trip durations and lengths can be easily derived. Functional classifications of traversed
network links and land-use classifications of origin and destination locations can be determined by importing the GPS data into a GIS. Second-by-second speed data is also available from GPS, further describing actual travel behavior. Given the spatial and temporal nature of travel and transportation, GPS appears to offer great hope in the cost-effective collection of detailed travel data.

As mentioned previously, there have been several research projects investigating automated diaries with GPS, including the 1996 FHWA-sponsored Lexington study [Wagner 1997] and more recent travel surveys conducted in the Netherlands [Draijer 1999]. These two projects are the first to combine electronic travel diaries with GPS receivers to gain exact temporal and spatial details of each trip. Austin’s 1998 household travel survey included a passive GPS component to supplement their traditional survey data [Casas 1999], and a 1998 California truck study also used a passive GPS system to capture travel information [Wagner 1998]. Studies underway in Quebec City, Canada are using passive GPS receivers combined with computerized activity scheduling surveys to record detailed spatial-temporal activity patterns and the underlying decision-making processes of individuals with a household [Doherty 1999]. Researchers at Louisiana State University have been testing a process in which GPS receivers are sent to recruited survey participants and installed in each vehicle of the household [Stopher 2000]. The Lexington, Austin, Netherlands, and Quebec City studies will be reviewed detail in the following chapter.

Numerous travel surveys to be conducted in the U.S. within the next year are also planning a GPS data collection component. The year 2000 U.S. NPTS may offer an
electronic travel diary with GPS as part of the add-on study data collection set. Passive in-vehicle GPS systems are also planned as supplements to traditional PAPI and CATI surveys for several travel surveys in California and Washington. In Atlanta, the year 2000 household travel survey will include both an ETD with GPS as well as the passive in-vehicle GPS system to capture enhanced travel data [Wolf 1999b, Wolf 2000b].
CHAPTER III

GPS TRAVEL BEHAVIOR DATA COLLECTION

Projects using GPS-enhanced travel behavior data collection fall into two general categories: 1) electronic travel diaries with GPS, as tested in the 1996 Lexington, Kentucky study with 100 participants and 1999 Netherlands study with 150 respondents; and 2) passive in-vehicle GPS systems, as implemented in the 1998 Austin household travel survey. The primary use of GPS in the electronic travel diary has been to augment the electronic trip data that are entered by the study participants. The intent of passive in-vehicle GPS systems has been to conduct a passive audit of in-vehicle travel that can be compared in a post-processing step to the recorded travel diary of the respondent to validate the reported data and/or to determine trip under-reporting rates. Research conducted at Georgia Tech has focused on GPS data accuracy and equipment functionality necessary to support a variety of travel data collection needs, including travel surveys and vehicle emissions studies.

Selected GPS Travel Behavior Studies

Four important studies have used GPS to supplement travel behavior data collection. Two of these studies, conducted in Lexington and the Netherlands, explored the use of handheld data electronic collectors with GPS to capture trip details as well as GPS data. The other two studies, conducted in Austin and Quebec City, took a closer
look at the use of passive in-vehicle GPS systems for detailed vehicle travel data collection over one-day and multi-day periods, respectively. Table 3-1, which follows the individual project reviews, contains a summary of the primary characteristics of each study. To conclude this section, a summary of findings with respect to equipment objectives, design, and performance is presented.

**Lexington, Kentucky**

In the fall of 1996, a field test was conducted by FHWA in Lexington, KY, in which 100 households were recruited to test the use of GPS in personal travel surveys [Wagner 1997, Murakami, 1997]. Off-the-shelf GPS and PDA (personal digital assistant) equipment were obtained and programmed to record all vehicle trips taken by the recruited participant / vehicle. The primary objectives of this field test were to test the equipment performance, to evaluate respondents’ willingness to participate, and to compare PDA/GPS measured trip data for a single vehicle in the household with self-reported data by the primary driver acquired via an NPTS-style recall interview for one day of the six-day survey period.

Out of the 100 households recruited in this study, 14 of the equipment packages contained PDA-logged trip data with no GPS data; this loss of data was attributed to a faulty external power supply for the GPS receiver. For the remaining 86 households, 23% of the PDA-recorded trips had no valid GPS positional points. Factoring in all data losses, a total of 63% of all PDA-recorded trips contained good GPS data sufficient for further analysis. In addition, although the software application was designed to log GPS data at one-second intervals, the actual recording occurred at 2 to 3 second intervals due
to a software communications problem. Missing GPS data was also an issue during GPS start-up periods and other trip segments; some of these missing segments could be attributed to the respondent not activating the device during a trip.

After the recall travel data were collected, there were several challenges encountered while attempting to match the PDA, GPS, and recall data for comparison. First, there were differences in the PDA and GPS data due to the design of the system; the PDA trip initiation occurs before GPS signal acquisition and the PDA trip-termination could occur after the last valid GPS position fix. Next, the GPS-derived trip distance is dependent on measured GPS points. If the trip start is not recorded, then that distance is not obtained. Other losses of GPS data within a trip can also result in lost distance calculations if the travel path cannot be derived from the other GPS data. Trip matching between the GPS data and recall data was also problematic since participants typically do not know the exact address of their destinations, interviewers may be unfamiliar with the travel area, and even matched GIS addresses may have shortcomings associated with the last GPS position fix, which may or may not be close to the actual destination. Finally, it is likely that some trips were missing from the PDA and GPS data because the respondent either forgot or refused to turn on the equipment.

Once the researchers were able to match 61% of the recall trips to machine-recorded trips successfully, the analyses revealed significantly different distributions in trip start times and trip distances between the two data sources. As suspected, people tend to report trip start times rounded to the nearest quarter or half hour. In addition, trip distances are often rounded by respondents to the nearest 5-mile increment for trips over
5 miles in distance. Comparing actual travel times and distances to the recall data revealed that this tendency towards rounding most often results in an overstatement of travel distance and time in self-reported studies. These findings were consistent when compared with the 1995 NPTS 6-month interim data set [Wagner 1997].

With respect to future potential of using GPS technology with regional and national travel behavior surveys, the study concluded that the project was successful in demonstrating that GPS can be used to improve travel behavior data collection. The researchers stated that standardization of handheld computer operating systems and GPS PCMCIA units was essential for allowing low cost options in the near future. Map-matching routines to process GPS data for transportation applications were identified as a critical software need. Finally, given the continued improvements in equipment size, weight, functionality, and cost, the researchers predicted that handheld PCs or PDAs with GPS could soon be used for travel data capture for all modes of travel.

The Netherlands

In 1997, the Dutch Ministry of Transport, Public Works, and Water Management initiated a pilot study on the use of GPS in collecting travel data based on positive results of a pre-pilot study [Draijer 2000]. Twelve handheld data logging devices were equipped with a combined GPS / DGPS receiver and battery pack. This equipment was stored in a video camera bag with the GPS and DGPS antennas woven into the shoulder strap itself. These 2-kilogram devices were given to 102 randomly selected participants and 49 personally recruited contacts for use in a four-day travel survey. (Delays in equipment development and testing caused the random recruitment goal to be replaced with personal
recruiting for the last 49 participants.) Participants were also given paper diaries with instructions to record trips on the paper diary whenever the GPS equipment was not used. It was felt that this back-up procedure was necessary to gain insight on missing trips and equipment package non-use. Once participation was complete, respondents filled out an evaluation form about using the packages and were given a $12 participation reward.

This study revealed both problems and promise. Technical difficulties with the packages resulted in the loss of data for 22 of the 151 respondents. Data quality issues caused the loss of travel data for another 8 people. Of the total 151 original respondents, 80 used only the GPS equipment, 17 used only the paper diary, and the remaining 54 used a combination of both. The follow-up evaluations revealed that the GPS equipment package weight and size played an important role in reasons why the GPS device was not used, especially for bicycle, walk, and public transport trips, as well as for specific trip purposes such as shopping and visiting. Overall, the researchers felt that respondents are willing to use the equipment and that further reductions in size and weight, as well as improvements in user interface, would increase participation rates.

With respect to GPS reception, the study found that nearly all of the vehicle trips (90%) were captured when the traveler was the driver, as compared to lower capture rates for bus (85%), bicycle (79%), walk (78%), vehicle passenger (70%), and train (53%) trips. DGPS signal reception occurred only 23% of the time; this low rate was attributed to poor antenna access and design. DGPS signal reception also varied greatly by mode; bus trips with GPS reception obtained the correction signal 71% of the time whereas vehicle trips experienced only 20% DGPS signal reception. In summary, the researchers
felt that it is possible to monitor all travel modes; however, they also noted that data completeness and quality varies among them.

**Austin, Texas**

During the 1997-1998 Austin one-day household travel survey, which included 1750 households, a passive GPS data collection activity also occurred with a subset of 200 households [Casas 1999]. For each household in the GPS study group, up to three vehicles were instrumented with a passive GPS data logging system so that a total of 356 vehicles were instrumented. The passive GPS systems were powered either continuously through a connection to the vehicle’s main power supply or only during times in which the vehicle was actually powered (note that the researchers refer to this as switched power). Differential corrections were collected at a centralized base station so that the GPS data could be post-processed to improve spatial accuracy to three to ten meters.

There were several problems with this effort. First, 78 of the 356 vehicle files were not recovered, which meant 22% of the sample was lost [Pearson 1999]. Next, due to problems with the vehicle GPS data files, only 104 households had complete, useable data (for 164 vehicles). In addition, numerous base station correction files were lost, which resulted in only 128 corrected vehicle files (36% of the original sample). Preliminary analysis of the uncorrected data files revealed that it was extremely difficult to identify trip ends for those systems that received continuous power (57% of the sample) because the effects of selective availability will cause speed measurements of two to four miles per hour even when the vehicle is stationary. For the switched power vehicles, numerous trip starts were lost due to GPS acquisition delays caused by power
supply resumption; typical satellite acquisition times ranged from 14 seconds up to 3 minutes. Analysis results of this data set will be released in late 2000.

**Quebec City, Canada**

In late 1998, researchers in the Interdisciplinary Research Group on Mobility, Environment, and Safety at Laval University in Canada began a multi-day GPS data collection test in which four researchers had GPS equipment packages installed in their vehicles [Doherty 1999]. The first objective of the study was to determine the capabilities of GPS technology in providing passive traces of vehicle movements over one to two week periods. Forty-nine total days of travel were recorded, in which 913 kilometers of travel and 164 stops occurred.

The tests revealed several problems with the equipment, including GPS acquisition times, power supply stability, data storage limits, and cold temperature tolerance issues with LCD screens. The researchers determined that even though there were a few technical problems, it was definitely possible to collect multi-day data for vehicle travel. The research team then developed several GIS-based algorithms to reduce the GPS data into route data and to identify any potential stops that occurred on the route.

The final objective of this research is to explore ways that this route and stop data can be presented back to travelers so that a greater understanding of activity scheduling processes can be gained. To achieve this objective, the researchers plan to integrate GPS data with CHASE (Computerized Household Activity Scheduling Elicitor), a software program developed at the University of Toronto in 1997 that allows survey participants to record and modify their scheduled activities throughout the survey period.
<table>
<thead>
<tr>
<th>Study location</th>
<th>Electronic Travel Diary with GPS</th>
<th>Passive In-vehicle GPS System</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Survey period</strong></td>
<td>Lexington, KY 6 days</td>
<td>Austin, Texas 1 day</td>
</tr>
<tr>
<td><strong>Time frame</strong></td>
<td>Fall 1996</td>
<td>1997-1998</td>
</tr>
<tr>
<td><strong>Purpose</strong></td>
<td>To test integration of GPS</td>
<td>To compare GPS data with</td>
</tr>
<tr>
<td></td>
<td>with self-reported travel</td>
<td>reported trips and start/end</td>
</tr>
<tr>
<td></td>
<td>behavior; to compare differences</td>
<td>times; to compare route/time</td>
</tr>
<tr>
<td></td>
<td>between two methods</td>
<td>with models; to examine</td>
</tr>
<tr>
<td></td>
<td></td>
<td>speed profiles</td>
</tr>
<tr>
<td><strong>Sample size</strong></td>
<td>100 households * 1 vehicle</td>
<td>203 households * 1 vehicle</td>
</tr>
<tr>
<td></td>
<td>per household</td>
<td>per household</td>
</tr>
<tr>
<td></td>
<td></td>
<td>up to 3 vehicles / household*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>total of 356 vehicles</td>
</tr>
<tr>
<td><strong>Modes collected</strong></td>
<td>Vehicle only</td>
<td>Vehicle only</td>
</tr>
<tr>
<td><strong>Additional data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>collection methods</strong></td>
<td>Recall interview for 1 day</td>
<td>Paper diaries were provide</td>
</tr>
<tr>
<td></td>
<td>of survey</td>
<td>to record trips not captured</td>
</tr>
<tr>
<td></td>
<td></td>
<td>by GPS equipment</td>
</tr>
<tr>
<td><strong>GPS interval</strong></td>
<td>3 seconds</td>
<td>1 second</td>
</tr>
<tr>
<td><strong>Differentially</strong></td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>corrected</strong></td>
<td></td>
<td>Real-time, FM</td>
</tr>
<tr>
<td><strong>Routing analyses</strong></td>
<td>Yes, Univ. of WI has compared</td>
<td>Limited</td>
</tr>
<tr>
<td></td>
<td>path variations between</td>
<td>Not yet</td>
</tr>
<tr>
<td></td>
<td>matched trips</td>
<td></td>
</tr>
<tr>
<td><strong>Positive findings</strong></td>
<td>CASI with GPS can improve travel</td>
<td>It is possible to capture GPS</td>
</tr>
<tr>
<td></td>
<td>survey data quality</td>
<td>data for all travel modes</td>
</tr>
<tr>
<td><strong>Findings of concern</strong></td>
<td>User activation issue</td>
<td>GPS can be used to identify</td>
</tr>
<tr>
<td></td>
<td>GPS power supply loss</td>
<td>missed trips</td>
</tr>
<tr>
<td></td>
<td>Data logging frequency</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Difficulty in matching recall,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PDA, GPS trips</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GPS acquisition time</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Up to 3 vehicles per household * Total of 356 vehicles

**Notes:**
- CASI: Computed Assisted Survey Instrument
- GPS: Global Positioning System
- DGPS: Differential GPS
- PP: Post-processing
- FM: Frequency Modulation
- CATI: Computer-Assisted Telephone Interview
Findings

There are two GPS applications that are evolving for use in travel survey data collection – a handheld PC with GPS capabilities for capture of all travel modes and a passive in-vehicle GPS system for capture of all vehicle-specific trips. Both systems show great promise in improving the quality of travel data collected and increasing the amounts and types of data that can be used to improve transportation and air quality models, which in turn should result in better policy-making decisions.

In the three field studies above, significant data losses occurred in all applications. These losses were attributed to several sources, included faulty power supplies, loose cabling, improper transfer procedures, and loss of satellite signals. The first three sources of data loss can be corrected with better equipment design and data transfer controls. Loss of satellite signals cannot be overcome when the sky view is completely blocked, such as when the vehicle is in a parking garage, in a tunnel, or in an urban canyon. However, some receivers do perform better than others when under heavy tree canopies. The systems also experienced GPS data loss during GPS signal acquisition periods when the equipment was first powered up for each trip. This issue may be mitigated in the passive GPS device if the power supply is continuous; however, limited survey duration due to power supply constraints then becomes an issue. There are also some receivers available that report superior cold start acquisition times as a result of new technologies. Finally, if the GPS receiver does not have a clear sky view during periods of vehicle inactivity, GPS signal reacquisition time is still an issue regardless of the power management approach.
While the handheld PC with GPS does have the capability to capture all trips, it is still very much dependent on the respondent to carry the package and to enter basic trip information. The passive in-vehicle system, on the other hand, is designed to have no user interface and therefore no user dependency; every vehicle trip should be recorded. Finally, a DGPS component is much easier to implement in the passive in-vehicle GPS system since size and weight are not a concern in this package. A summary of attributes and trade-offs between the two packages can be seen in Table 3-2.

<table>
<thead>
<tr>
<th>ETD with GPS</th>
<th>Passive In-Vehicle GPS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Travel Modes</strong></td>
<td>All modes can be captured</td>
</tr>
<tr>
<td><strong>User Dependence</strong></td>
<td>Highly dependent upon user</td>
</tr>
<tr>
<td><strong>Design Constraints</strong></td>
<td>Must be small, lightweight, user-friendly, with strong connections</td>
</tr>
<tr>
<td><strong>Power Supply</strong></td>
<td>Must be lightweight, thus limiting the length of the survey period</td>
</tr>
<tr>
<td><strong>GPS Accuracy</strong></td>
<td>Due to size and weight issues, DGPS is difficult to include; therefore 30 - 100 meters with SA</td>
</tr>
<tr>
<td><strong>Primary Use</strong></td>
<td>To supplement ETD data</td>
</tr>
</tbody>
</table>

The choice of which package to use depends upon the desired use of the data and the nature of travel mode use in the targeted survey area. Given the trade-offs in coverage and accuracy between the two applications, it seems practical to use the passive in-vehicle GPS system combined with paper and CATI methods, especially in a highly...
vehicle-dependent society such as the U.S. However, in locations where typical travel modes other than the personal automobile are more common, the electronic travel diary with GPS may be the only option if the study is intended to collect complete trip routes for all trips.

The handheld ETD with GPS may be a much more attractive solution in the near future. As PDA and GPS technologies continue to improve rapidly, it appears that low-cost DGPS integrated solutions are within a year or two away. Handheld PCs and palmtops prices are dropping steadily, along with the costs and availability of simple handheld GPS devices. Therefore, it is entirely feasible that handheld travel diaries with GPS capabilities will be available soon for less than $600. However, the challenge to get the respondent to carry and use the device will remain.

The Georgia Tech GPS Studies

Two studies conducted at Georgia Tech during the past four years have provided great insight into the use of GPS technology in travel studies – for both on-person and in-vehicle applications. The first project, sponsored by FHWA, developed a prototype system called the Comprehensive Electronic Travel Monitoring System to collect travel and vehicle activity data concurrently. The second project, Strategies for Metropolitan Atlanta’s Regional Transportation and Air Quality (SMARTRAQ), was initiated in 1998 to evaluate the relationships between travel, land use, and air quality impacts. One component of this region-wide study is the development and deployment of several GPS-enhanced equipment packages to augment the Year 2000 Atlanta Regional Household
Travel Survey with automated methods to capture more and more accurate detail regarding individual and vehicle travel patterns. The ability to continue evaluating available GPS and data logging technologies as part of these research efforts has enabled the research team to develop functional and robust systems for each application area.

**CETMS Overview**

From 1997 through 1999, Georgia Tech researchers undertook a FHWA-sponsored study to develop a prototype system, the Comprehensive Electronic Travel Monitoring System or CETMS, that would capture travel and vehicle activities concurrently. At the time, vehicle activity studies and travel behavior studies had never been coordinated. This research effort focused on the development of electronic monitoring equipment that would allow such studies to be undertaken jointly. The prototype system developed in this research effort allowed for simultaneous monitoring of trip characteristics and vehicle and engine operating conditions by capturing all traditional travel diary information as well as second-by-second GPS coordinates and engine / vehicle activity data. Hence, data collected with this equipment could help model developers understand the linkages between travel decisions and vehicle operating conditions. This, in turn, could lead to the development of improved algorithms for predicting emissions as a function of vehicle characteristics, on-road traffic conditions, and household/driver demographics.

To collect all required data streams for use in travel model development and validation, the integrated instrumentation package designed for this project included a handheld electronic travel diary, a global positioning system receiver and antenna, an
onboard engine computer monitor, a rugged laptop computer, and related cabling and power supplies. Because many manufacturers and models within each equipment class had the potential to meet minimum design and performance criteria, the research team first developed a set of system requirements in the form of functional requirements [Wolf 1999c]. These functional specifications served to narrow the field of potential equipment that would meet project requirements. Given a smaller set of potential equipment solutions, the research team reviewed the detailed technical specifications for the most promising equipment options. The technical specifications analyses are contained in Wolf et al., 1999d. The most promising components identified were then selected for purchase and testing.

Each class of equipment was first tested separately to determine superior performers within each class. To facilitate competitive testing, the research team identified those factors that could influence the accuracy and ease of use of each equipment type. With this knowledge in mind, the team then developed standardized test scripts and testing procedures that challenged the ability of equipment to perform under real-world test conditions. The descriptions of the test metrics and the performance of each individual component are summarized in the test plans and results document [Wolf 1999e]. The selection of individual components, final assembly of the prototype system, and software development were based upon these test results. Once the final package was assembled, field tests were conducted on the prototype system. The final report on this project is available in Wolf et al., 1999a.
The system was designed to capture or compute the following data elements:

- All driver and passenger travel activities that would normally be captured using manual travel diary or telephone survey methods. For each trip, this information includes: vehicle (or other mode) identification, driver identification, passenger identification, driver and passenger trip purposes, trip start time, finish time (or duration), origin location, destination location, and distance traveled. In addition to these traditional elements, route choice, travel speed, and functional classification of each link (with traffic conditions) can be determined by tying GPS data to a GIS database, greatly enhancing the original data collected.

- All vehicle and engine operating conditions affecting emissions that could feasibly be captured via an onboard engine computer monitor. These data include such variables as vehicle speed, acceleration, engine revolutions per minute (rpm), manifold absolute pressure, throttle position, catalyst temperature, gear selection, air/fuel ratios, and coolant temperature.

The role of the GPS data stream was critical in this project. First, it served to improve and supplement traditional travel diary data elements such as trip start and finish times, trip lengths and durations, and origin and destination locations. In addition, the second-by-second trip route information (i.e., position, speed, and time) that was captured was matched to the second-by-second engine activities (such as rpm, throttle position, and catalyst temperature) so that vehicle and engine activities could be mapped spatially and temporally onto the transportation network. In fact, it is the GPS element that allows this matching and mapping to occur.
SMARTRAQ Overview

Building upon the findings of the CETMS project, researchers at Georgia Tech have developed three instrumentation packages that will automate the capture of personal travel data for various sub-samples in the Year 2000 Atlanta Metropolitan Regional Travel Survey. These three packages include: 1) a passive in-vehicle GPS system, which has a GPS receiver, antenna and data logger to capture vehicle trips only; 2) a handheld electronic travel diary (ETD) with GPS capabilities to capture trip information for all modes of travel; and 3) a comprehensive electronic travel monitoring system (CETMS), which includes an ETD, a rugged laptop computer, a GPS receiver and antenna, and an onboard engine monitoring system, to capture all trip and vehicle performance information. All three systems have been designed to capture data for survey durations up to four days.

The Year 2000 Atlanta Metropolitan Regional Travel Survey is a component of Strategies for Metropolitan Atlanta’s Regional Transportation and Air Quality (SMARTRAQ). This major research endeavor will collect and analyze sufficient data to develop an enhanced suite of travel demand models capable of addressing the land use, travel behavior, and air quality issues critical to the Atlanta metropolitan region. To support these data needs, a regional household travel survey will be administered to 8000 randomly selected households that are stratified by household income, land-use type and household size. In addition, automated data capture was identified as the primary method for gaining significantly greater detail and insight into the relationships between land-use, emissions, and travel behavior within the region. Three specific survey efforts were
identified as needing automation support: the General Purpose Survey, the Physical Activity Study, and the Summer Ozone Study.

The General Purpose Survey will use the passive in-vehicle GPS system for a subset of the respondents employing traditional travel survey data collection methods. The GPS data will be compared with reported trip data, allowing for the development of correction factors for trip under-reporting and mis-reporting, as well as providing detailed insight on trip making and trip chaining behaviors within the region. The Physical Activity Study entails an add-on survey for a subset of the General Purpose Survey participants and will use the electronic travel diary with GPS to more accurately capture all traditional travel survey elements (traveler and companion identities, trip purpose, modes, origin location, start time, destination location, finish time, and travel distance) as well as additional information on travel routes and speeds. These data will support the primary objective of the Physical Activity Study, which will examine factors affecting the level of physical exercise that travelers obtain as a function of their trip-making patterns and the affects that land use may have on net physical activity. The Atlanta-based Centers for Disease Control (CDC) is a primary sponsor of this sub-survey.

The third system, the CETMS, will support the Summer Ozone Study. It includes the ETD as described above and a comprehensive in-vehicle data collection system with both GPS technology and an engine-monitoring device. In addition to collecting all traditional travel data elements, as well as exact travel routes and second-by-second vehicle positions, the CETMS will also collect second-by-second vehicle and engine operating conditions from the vehicle’s onboard engine computer monitor. Engine and
vehicle data available for capture include such variables as speed, acceleration, engine rpm, manifold absolute pressure, throttle position, catalyst temperature, gear selection, air/fuel ratios, and engine coolant temperature. As mentioned previously, this system will enable modelers to examine how network driving patterns (via speed and acceleration profiles and engine operation data) leads to elevated emissions and to differentiate between driver-induced emissions and traffic stream-related emissions.

Implementation of the first two systems will coincide with the General Purpose Survey and Physical Activity Study, which will commence in the fall of 2000 and continue through the spring of 2001. Although funding for the Summer Ozone Study is not guaranteed at this time, plans for CETMS deployment include the spring and summer of 2001. Preliminary estimates on sample sizes for the instrumentation studies are 400 to 675 households with the passive in-vehicle GPS system, 510 respondents with the ETD and GPS device, and 160 vehicles with the CETMS.

Findings and Recommendations

The CETMS project that completed at Georgia Tech in the spring of 1999 developed the prototypes of both the CETMS and the ETD. This development effort involved the evaluation of numerous GPS and DGPS receivers and antennas, handheld and laptop computers, batteries, and engine monitoring devices. Since then, Georgia Tech researchers have continued evaluating the latest GPS equipment, rugged laptops, handheld PC’s, and batteries for possible use in the SMARTRAQ instrumentation packages. A summary of the key findings from these research efforts follows. For a
more thorough review, readers are encouraged to reference several papers and reports by Wolf, J. et al. (1999a, 1999b, 1999c, 1999d, 1999e, 1999f, 2000a, 2000b).

**GPS / Differential GPS (DGPS)**

- To obtain accurate origin location, destination location, and route choice data, differential correction of GPS data is required. Data accuracy errors associated with Selective Availability – the U.S. government’s intentional degradation of signal accuracy – and from atmospheric delays can range from 40 to 140 meters. The Atlanta study has set a 10-meter positional accuracy requirement to ensure complete map matching in the Atlanta network. Comparisons made of corrected and uncorrected GPS route data collected in Atlanta neighborhoods revealed that incorrect streets would be assigned if uncorrected data were used [Wolf 1999e]. Both the Lexington and Austin GPS-enhanced travel surveys [Wagner 1997 and Pearson 1999], which collected raw GPS data, spent an unexpected amount of time processing this uncorrected GPS data. Even with various map matching software applications, there is a need for extensive quality control of the routes if the data are not differentially corrected.

- Vehicle speeds are estimated through a GPS phase differential calculation (change in signal phase with time) this calculation is also affected significantly by Selective Availability; hence, differential GPS should be employed in any GPS system requiring collection of accurate speed and acceleration data. DGPS is required to meet the Atlanta study speed accuracy requirement of plus or minus two percent.
• Post-processing differential correction of GPS signals involves setting up two GPS receivers with portable computers to record all pertinent satellite information at both the base station and field unit. The field data are then post-processed to correct the x, y, and velocity readings given that the known base station location was stationary. Real-time differential GPS involves the real-time transmission of a base station correction signal from a broadcast station or satellite to a DGPS receiver connected to a GPS receiver in the field.

• To avoid the data storage and labor costs, as well as the potential for base station data losses associated with post processing differential correction methods, real-time differential GPS should be employed in any passive GPS systems requiring route choice data. For example, satellite signals collected and processed by the GPS carry a great deal of data associated with satellite and signal performance. Approximately 1 megabyte per hour of raw data needs to be logged to a data logger if post-processing DGPS is used. Using real-time differential correction methods, however, will reduce the data storage requirements for second-by-second storage of 10 GPS data elements to approximately 80 kilobytes.

• Two feasible real-time DGPS solutions are available in the Atlanta metropolitan area – an FM sub-carrier based commercial transmission and the radio beacon signal transmitted free-of-charge by the U.S. Coast Guard. Preliminary analyses of the FM-based differential signal conducted by Georgia Tech researchers have revealed that coverage and reliability are a serious concern for the 13-county metropolitan area.
• Evaluation of radio beacon signal reception, coverage, and reliability in the Atlanta region are currently being conducted on the radio beacon broadcast from the Macon, Georgia station, which came online March 20, 2000. Field tests conducted on Atlanta metropolitan highways have shown full NDGPS coverage for the area.

• Extensive research into small, economical, and integrated DGPS solutions needed for the ETD package found that there are none currently available. To integrate DGPS into the ETD with existing technology would require deployment of a much heavier and larger system since most DGPS technology improvements are currently focused on vehicle and marine applications. A pilot travel diary study conducted in the Netherlands in 1999 used an electronic travel diary with GPS and DGPS components. The total system weighed approximately 4.5 pounds (2 kilograms) and 41% of the participants reported negative feedback on the size and weight of the system [Draijer 2000]. In fact, size and weight were found to play an important role for non-use of the system in trips made by bicycle, walking, and public transport, and for trip purposes such as shopping and visits. Consequently, the SMARTRAQ research team has decided to recommend not using differential GPS for the ETD package, at least until there are smaller, integrated solutions available.

Other GPS-Specific Findings

• GPS satellite signals are often lost when a vehicle passes beneath a tree canopy or under an overpass. The response of the individual GPS system to these conditions
(which can only be determined through field-testing because manufacturer specifications are typically for “optimal” conditions) affects the amount and accuracy of the data collected. A GPS system resistant to loss-of-lock under tree canopies is critical for Atlanta and other urban areas with heavy forestation.

- Multi-path errors, which result when satellite signals reflect off of an object (e.g., buildings, vehicles, and ground surfaces) prior to their reception at the GPS antenna, can also cause degradation of position accuracy. GPS receivers and antennas are available that claim to minimize these effects and should be considered. In addition, antenna placement can be optimized to reduce these effects.

- When the GPS system loses satellite lock, it is important for the system to reacquire signals as quickly as possible. Most systems experience an extended cold start signal acquisition delay when starting from a power-off or loss-of-signal condition of more than several hours. To ensure data capture for accurate origin data during this emissions-sensitive period, initial GPS positions should be determined within a maximum of 45 seconds after a cold start. Reacquisition time after an intermittent loss of satellite lock should be no more than 2 seconds.

- Equipment located in the trunk or cabin of a vehicle must be capable of withstanding sustained temperatures of 150°F. Even with high-efficiency insulated packaging, the equipment must be capable of operating at the average daily temperature. Low temperatures below freezing are also a major issue in the
winter months. Most GPS devices without displays are designed to withstand such temperature extremes.

**Geographic Information Systems (GIS)**

- In addition to addressing issues with GPS position accuracy (as mentioned above), it is also critical that a spatially accurate GIS base map is used for GPS data map matching. It would be a wasted effort to obtain GPS spatial accuracies of 5 meters if the base map used for travel route reconstruction is only accurate to 50 or 100 meters.

**Handheld and Portable Computers**

- Handheld and portable computers typically cannot withstand high or low temperatures, since both the screen and the CPU will lock up at such extremes. Therefore, rugged handheld and laptop solutions for the passive in-vehicle GPS system and CETMS system are needed since these systems that will remain in the vehicle throughout the survey period.

- The Palm IIIx handheld organizer cannot withstand extreme trunk temperatures. However, the electronic travel dairy will be carried with the traveler (as opposed to being left in the vehicle), so temperature extremes should not be an issue with the use of the Palm IIIx in the ETD.

**Power Supply**

- For liability and installation cost reasons, it was initially decided that there would be no connection made to the vehicle’s power system (i.e., all systems would be powered by their own power supply). Therefore, power demand of system
components is a significant concern for extended (3-7 day) studies. All systems must be internally-powered for the entire period or capable of running off of a 12V deep cycle marine battery placed in the trunk of the vehicle.

Summary

The studies conducted in Lexington, Austin, Quebec City, and the Netherlands examined the use of GPS to supplement travel diary data collection, in either paper or electronic diary form. The research at Georgia Tech has also examined both types of configurations, as well as adding a new configuration that augments the system with engine monitoring capabilities. Research efforts for all three equipment types will move into an implementation phase in the year 2000 and beyond as metropolitan planning organizations (MPOs) begin to explore and adopt the use of GPS in travel studies. In fact, it is anticipated that several products will be available in a production mode within the next year for use in GPS-enhanced travel studies. In addition, the FHWA has been sponsoring a GPS map-matching project that should also make software available for processing the GPS data in a GIS environment.

Preliminary processing of GPS data collected in the completed studies has revealed both opportunities and challenges. The opportunities are essentially as predicted: more accurate and complete travel data and new travel elements (e.g., route and speed). The challenges in using GPS and DGPS technologies are due to the learning curve associated with understanding these relatively new technologies and with keeping up with the rapid advancements being made in the field itself. These challenges include gaining a better understanding of equipment functionality, integrating these components into the overall
package, minimizing receiver and antenna size and weight, obtaining power and data storage capabilities for the study duration, determining the data accuracy required and obtained, processing the collected data, and map matching the data to obtain GIS integration.

Of course, great gains have been in each of these areas as researchers continue to investigate the GPS equipment and data, and as manufacturers continue to improve the equipment itself. Overall, technology in the GPS field and the handheld / portable computer field is changing rapidly, as functionality improves while size is reduced and costs continue to fall. Of particular promise is the trend towards more integrated packages as manufacturers compete in this technology-hungry marketplace. Finally, the removal of Selective Availability on May 1, 2000 promises to make GPS technology ubiquitous.
CHAPTER IV

USING GPS DATA TO DERIVE TRAVEL DIARY DATA

Data quality problems such as item nonresponse and inaccuracy have led researchers and practitioners to seek better methods for personal travel behavior data collection, most recently in the area of automated methods, including the use of mobile computers (laptop or handheld) and GPS devices. The research projects described in Chapter 3 investigated the use of GPS for collecting personal travel. These data were collected primarily as a supplement to travel diary data that was collected concurrently, either by paper or electronic means. The overviews of these studies presented in Chapter 3 focused on the data collection aspect of the studies.

Some analyses have been conducted on these data sets with respect to their ability to match or exceed traditional travel survey data quality. Table 4-1 below lists the five studies presented previously and the scope of the analysis work performed to date. For example, the Battelle report on the Lexington study includes detailed comparisons of GPS-recorded trips details with travel reported by the same participants during a recall interview conducted a short time after the study period [Wagner 1997]. This same GPS data set was used by researchers at the University of South Florida to examine GPS – based data collection methods for capturing multi-stop trip-chaining behavior [Yalamanchili 1999] and by researchers at the University of Wisconsin for using GPS data to evaluate path choice assumptions in trip assignment models [Jan 2000].
Table 4-1: Summary of GPS-based Data Collection Analyses

<table>
<thead>
<tr>
<th>Study / Analyst</th>
<th>Data Analyses</th>
</tr>
</thead>
</table>
| Lexington / Battelle            | 1) Compared ETD & GPS-captured trips versus trips reported by CATI recall interviews for same trips; elements analyzed include individual trips, travel times, and trip lengths.  
2) Evaluated GPS data accuracy and completeness; map matching software was considered to identify trip links |
| Lexington / UWi                 | 1) Compared derived travel paths with shortest paths                          |
| Lexington / USF                 | 1) Evaluated use of GPS-based data collection methods in capturing multi-stop trip chains |
| Austin / NuStats TTI            | 1) Compared passive GPS-collected data with travel diary / telephone interview reported travel data |
| Netherlands / Ministry of Transport | 1) Compared GPS performance across different travel modes                     
2) Assessed GPS system usage by survey respondents |
| Quebec City / Laval University  | 1) Tested feasibility of recording multi-week vehicle travel using GPS        
2) Developed GIS algorithms to identify trip ends, travel times, road usage, and speed |
| Atlanta / GA Tech               | 1) Tested accuracy levels and performance characteristics of different GPS equipment with respect to ability to successfully map match GPS data to underlying GIS network database |

This Study

Automated data collection studies to date have examined the benefits of electronic travel diaries and/or GPS technologies for improving the accuracy and completeness of travel data. However, the ETD projects have not reduced respondent burden because of the new equipment operating responsibilities placed on the respondent. In addition, passive GPS data loggers have created additional data processing work for survey researchers as they attempt to reconcile reported and recorded data streams. One research topic in the area of GPS travel survey data collection that has not been explored is the
possibility of using GPS data to completely replace the travel diary recording and retrieval process.

If a vehicle could be instrumented with a GPS data logger to collect all trip data, and if the data were processed in such a way that most traditional trip data could be derived, then traditional travel diaries could be eliminated from the travel survey process. The transformed GPS data could be easily integrated into CATI operations so that household data retrieval phone calls could validate derived travel data and could collect any missing elements. To assess the feasibility of this approach, this dissertation will evaluate the use of GPS data loggers to collect travel data and the application of software and GIS databases to derive travel data elements that are traditionally recorded on paper diaries by respondents and retrieved by CATI operators. The ultimate goal of this process is to allow travel survey administrators to use GPS data loggers instead of paper travel diaries in the collection of personal travel behavior.

In this scenario, survey recruitment calls would be conducted to identify survey respondents and preliminary household data (in the same manner as they are conducted in other household travel surveys). Household, person, and vehicle data sheets would still be sent to each recruited household just as in a normal survey; however, instead of the respondents receiving a paper or electronic diary, they would receive a passive in-vehicle GPS data logger to record all vehicle trips. The GPS data logger would be installed for the targeted study day(s) and then returned to the survey control center. The data would then be processed quickly (i.e., in one to two days) and then be made available to CATI operators who, when conducting the normal retrieval call for household, person, and
vehicle data, would use the processed GPS trip data to verify any questionable elements and/or to collect any missing elements (such as non-vehicle trips). This processed data would be presented to the CATI operator in a user-friendly format to expedite the confirmation or collection of travel data.

This study will assess the ability of this new process to replicate reported travel behavior. Success will be based on comparisons made between the reported and derived travel data. Since previous GPS travel data collection studies have already investigated the use of GPS data to identify trips, along with trip start times, trip durations, and trip routes and distances, this research will focus on the one trip characteristic that has not been evaluated in previous GPS travel data collection studies – trip purpose. It is the derivation of this trip element that would make this new process feasible.

To test this approach, a proof-of-concept study was proposed to examine the collection of second-by-second GPS data during in-vehicle trips and the subsequent processing of such data to generate all traditional trip elements associated with personal vehicle trips. As such, thirty survey participants were recruited to use GPS data loggers in their personal vehicles for three-day periods. Twenty-four of the participants were also asked to keep a paper trip diary (in the form of a memory jogger) for the same survey days. The data collected were processed and translated into trip logs; these derived trip logs were then compared with manual trip logs recorded by the same participants. Chapters 5, 6, and 7 contain a review of the data collection and processing steps, along with the subsequent analyses and results.
Scope

Due to the types and quantities of equipment available for this data collection effort, only vehicle-based trips were targeted for analysis; this simplification allowed for the development of four comparable in-vehicle GPS data loggers. In addition, developing a robust process for in-vehicle trips is a logical first step in developing the appropriate logic for processing GPS collected during all modes of travel. To capture all modes of travel, a personal GPS data logger would be required. (The trips captured in this study were at the vehicle level, not the person or household level, which is the traditional level used in household travel surveys.) To obtain a household-level solution using GPS equipment, a household-level approach would be needed; all persons and/or vehicles would need to be equipped with personal GPS data loggers. Therefore, for vehicle-dependent households (which is very much the case in Atlanta), a vehicle approach to household travel data collection is reasonable.

Also, for the purpose of this analysis, the 1990 Atlanta Household Travel Study was used as the source of trip data requirements. Although the format of travel diaries have evolved over the past 10 years, the core trip elements have not changed given that most MPOs are still using the traditional four-step travel demand models that have been in place for the last decade. In addition, the Atlanta Regional Commission (ARC) is just now preparing to conduct their next regional travel survey, which will begin in late 2000, and is still in the processing of defining their exact travel data needs. A meeting conducted in April 2000 between the Georgia Tech research team and the ARC planners to discuss data needs led to a review of the 1990 Atlanta Household Travel Study final
report and travel demand user manuals. Finally, although the GPS data are analyzed with respect to the ability to identify trips, trip start and finish times, and trip distances, the analysis will focus on the feasibility of deriving trip purpose from the GPS data.

Data Element Analysis

Before examining the collection and processing of GPS data to develop travel data, a review was conducted of the primary data elements collected in the travel diary component of the 1990 Atlanta Household Travel Study, as well as of the data available in the standard GPS messages. The GPS data standards and communications protocols, as defined by the National Marine Electronics Association, are the source of the GPS data element analysis. GIS databases available at the GIS Center at Georgia Tech were inventoried for their applicability to the GPS data processing tasks required for this study. Finally, the essential trip elements were matched to their potential GPS and/or GIS data sources.

Travel Survey Trip Data

In the fall of 1991, the Atlanta Regional Commission conducted the 1990 Household Travel Survey with 2433 households in the eleven-county region (this included the nine ARC counties plus Cherokee and Bartow counties) [ARC 1993]. These 2433 households contained an average of 2.61 persons greater than four years of age for whom travel information was collected, for a total of 6269 persons. Eighty-two visitors to the sampled households also made trips with characteristics that were collected during the study. The households averaged 9.3 trips per day and the average person made 3.8 trips
per day; the total number of trips reported for the study was 23,308. Of the total 6351 respondents, 5173 (81.5%) used the travel diary provided to record the trips made on their assigned travel day. This determination was based on the interviewer’s assessment of the confidence and clarity of responses provided by the respondent (i.e., the question was not asked directly to the respondent). If the interviewer could not make this assessment, a “no diary use” was coded.

The survey data was collected by telephone interviews conducted by NuStats, a consulting firm specializing in travel survey data collection. The data collected was recorded on paper (this was not a CATI retrieval process), and then the forms were edited, coded, and entered into a computer. This data was then provided to ARC in the form of three primary file types: the household file, the person file, and the trip files. They also provided a combined work file that had elements of all three files. Table 4-2 lists the key fields in the trip file and Table 4-3 shows the trip-related fields contained in the combined work file. Note that the travel diary trip origin and destination addresses contained in the trip file are used to determine the traffic analysis zone (TAZ) for use in the trip generation step of the travel demand modeling process. The trip files were created in two forms, unlinked and linked. Linked trips combined mode change trips into a single transit trip and combined drop off / pick-up passenger trips if they were incidental to the trip purpose and if the stop at the location was less than six minutes in duration.
### Table 4-2: Final Master Trip File Elements (TRPUNL91.DAT)

<table>
<thead>
<tr>
<th>Type of Data</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>File and record information</td>
<td>Record Type, Sample Number, Person Number, Trip Number</td>
</tr>
<tr>
<td>Address Type and Details</td>
<td>Home, Exact Street Address, Intersecting Streets, Landmark, Place Name w/o Address, or Out-of Area Trip</td>
</tr>
<tr>
<td>Jurisdiction</td>
<td>Jurisdiction of trip destination</td>
</tr>
<tr>
<td>Place Code</td>
<td>9 Land Use Codes</td>
</tr>
<tr>
<td>Trip Purpose</td>
<td>10 Trip Purposes</td>
</tr>
<tr>
<td>Beginning Time of Trip</td>
<td>Within time range of 0400 to 2800</td>
</tr>
<tr>
<td>Ending Time of Trip</td>
<td>Within time range of 0400 to 2800</td>
</tr>
<tr>
<td>Mode of Travel</td>
<td>Driver or passenger in auto, van, pick-up, or motorcycle; vanpool/carpool, taxi, MARTA train, bus (include CCT), school bus, social service / special bus, walk or bike to work, other</td>
</tr>
<tr>
<td>Number in Vehicle</td>
<td>Only if mode is driver of auto, van, pick-up, or motorcycle</td>
</tr>
</tbody>
</table>

### Table 4-3: Final Combined Work File -- Trip Elements Only (COMB92.DAT)

<table>
<thead>
<tr>
<th>Type of Data</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>File and record identification</td>
<td>Record Type, Sample Number, Person Number, Trip Number</td>
</tr>
<tr>
<td>Origin information</td>
<td>Place Code, Trip Purpose, TAZ</td>
</tr>
<tr>
<td>Destination information</td>
<td>Place Code, Trip Purpose, TAZ</td>
</tr>
<tr>
<td>Beginning Time of Trip</td>
<td>Within time range of 0400 to 2800</td>
</tr>
<tr>
<td>Ending Time of Trip</td>
<td>Within time range of 0400 to 2800</td>
</tr>
<tr>
<td>Mode of Travel</td>
<td>Driver or passenger in auto, van, pick-up, or motorcycle; vanpool/carpool, taxi, MARTA train, bus (include CCT), school bus, social service / special bus, walk or bike to work, other</td>
</tr>
<tr>
<td>Number in Vehicle</td>
<td>Only if mode is driver of auto, van, pick-up, or motorcycle</td>
</tr>
<tr>
<td>General Trip Purpose</td>
<td>Generalized trip purpose based on origin and destination trip purposes</td>
</tr>
<tr>
<td>Reported Travel Time (minutes)</td>
<td>Based on reported beginning and ending trip times</td>
</tr>
<tr>
<td>Elapsed Activity Time (minutes)</td>
<td>Estimated activity time at destination, calculated based on ending time of one trip and beginning time of next trip</td>
</tr>
</tbody>
</table>
Finally, there were no detailed vehicle data collected in the 1990 survey; consequently there was no vehicle file. The only variables collected that were related to vehicles were the number of autos available to the household (contained in the household file) and whether or not each person had a driver’s license (contained in the person file).

One key objective of this dissertation is to determine if land use codes and trip purposes can be derived from GPS data. The codes for both appear as defined and used in the 1990 study are contained in Tables 4-4 and 4-5. For each trip record, the individual trip purposes as reported by the respondents were coded as one of the 10 options below.

Table 4-4: Place Codes (one-digit land-use codes)

<table>
<thead>
<tr>
<th>Code</th>
<th>Land-Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Residential</td>
</tr>
<tr>
<td>1</td>
<td>Agriculture, Forestry, Fishing</td>
</tr>
<tr>
<td>2</td>
<td>Manufacturing – Durable Items</td>
</tr>
<tr>
<td>3</td>
<td>Manufacturing – Non-Durable Items</td>
</tr>
<tr>
<td>4</td>
<td>Transportation, Communications, and other Non-manufacturing industrial</td>
</tr>
<tr>
<td>5</td>
<td>Commercial Retail</td>
</tr>
<tr>
<td>6</td>
<td>Commercial Services</td>
</tr>
<tr>
<td>7</td>
<td>Wholesale Trade and Contracting</td>
</tr>
<tr>
<td>8</td>
<td>Public and Quasi-Public Buildings</td>
</tr>
<tr>
<td>9</td>
<td>Public and Quasi-Public Open Spaces</td>
</tr>
</tbody>
</table>
Table 4-5: Trip Purpose Codes

<table>
<thead>
<tr>
<th>Code</th>
<th>Trip Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Return Home</td>
</tr>
<tr>
<td>2</td>
<td>Go to Work</td>
</tr>
<tr>
<td>3</td>
<td>Personal Business</td>
</tr>
<tr>
<td>4</td>
<td>Shop</td>
</tr>
<tr>
<td>5</td>
<td>School</td>
</tr>
<tr>
<td>6</td>
<td>Social / Recreational</td>
</tr>
<tr>
<td>7</td>
<td>Work Related</td>
</tr>
<tr>
<td>8</td>
<td>Drop off / Pick up Passenger</td>
</tr>
<tr>
<td>9</td>
<td>Change Mode</td>
</tr>
<tr>
<td>0</td>
<td>Other</td>
</tr>
</tbody>
</table>

To convert these purposes into useful input for the trip generation travel demand modeling step, the origin and destination purposes for each trip were then translated into a generalized trip purpose that fell into two broad categories: home-based and non-home based. The home-based trips were defined as a trip with the person’s home as one of the trip ends. If neither trip end was at the person’s home, then the trip was considered to be non-home based. Although it will be sufficient in this study to determine the individual trip purpose only (since the generalized trip purposes are simply a table-lookup of the individual trip-end pairs), the generalized trip purposes are listed here for completeness:

- Work-related non-home based
- Other non-home based (two codes for work to non-work, non-work to work)
- Home-based work (different codes for first or last trip, and mid-day trip)
- Home-based school
- Home-based shop
• Home-based personal
• Home-based social/recreational
• Home-based work related
• Home-based eat meal
• Home-based serve passenger
• Home-based other

**GPS Receiver Output**

The ultimate goal of collecting GPS data in travel surveys is to automatically obtain highly accurate temporal and spatial details of each trip taken by each participant. By collecting GPS information at each trip start and finish, travel diary data elements such as trip origin and destination, and trip start time and finish time are acquired. If the GPS data is collected periodically throughout the trip itself, then travel path and travel distance can be derived.

Most GPS receivers output information using a subset of available NMEA 0183 GPS message formats. NMEA 0183 is the National Marine Electronics Association’s 0183 ASCII interface standard for marine electronic devices. This standard was designed to define electrical signal requirements, data transmission protocol and timing, and specific sentence formats for a 4800-baud serial data bus [NMEA 1998]. According to NMEA, the standard has been developed “to permit ready and satisfactory data communication between electronic marine instruments, navigation equipment, and communications equipment when interconnected via an appropriate interface.” Examples
of such equipment include weather instruments, timekeepers, velocity sensors, radar, Loran C, heading sensors, GPS and Global Navigation Satellite System (GNSS), electronic chart systems, and communication modes such as satellite, radio-telephone, and scanning receiver.

The NMEA 0183 standard calls for data communication in the form of coded "sentences." Each sentence begins with the character "$" and ends with a check sum (optional), carriage return and line feed (<CR><LF>). These last two characters are "control" characters and are not evident in the data. Between the beginning and end of each sentence a number of data fields separated by commas. The first field in any sentence (field 0) begins with the two-letter talker mnemonic code ("talkers" are devices that send out information; "listeners" take it in) followed by the three-letter sentence ID. The talker ID for the GPS sentences is “GP.” Data then follows in the standard format for that sentence. If any data for a given field is not available, the field is simply omitted, but the commas that delimit it are still sent, with no space between them. In addition, all leading zeros in non-zero values are transmitted.

For example, the Garmin GPS 35 combination receiver / antenna has a 1-second factory-set default rate for message generation at 4800 baud, but this can be modified if the user wants to vary the baud rate or output sentences [Garmin 1996]. All data are generated in sentence form and the sentences are transmitted contiguously. The Garmin GPS 35 outputs coordinated universal time (UTC, also referred to as Greenwich mean time, or GMT) date and time of day in its sentences. Prior to initial position fix, the
receiver’s on-board clock provides the date and time; after the first position fix, GPS satellite data is used to determine date and time.

Sentence formats for the four most important NMEA 0183 GPS sentences can be found in Appendix B. These sentences are: 1) Recommended Minimum Specific GPS / Transit Data (RMC); 2) Global Positioning System Fix Data (GGA); 3) GPS DOP and Active Satellites (GSA); and 4) GPS Satellites in View (GSV). Table 4-6 contains the fields from the RMC and GGA sentences that have relevance in travel diary studies [NMEA 1998, Garmin 1996]. Whereas the first eight fields are directly related to position, time, and velocity, the last four fields provide estimates of accuracy. Some GPS receivers also transmit proprietary messages in addition to NMEA sentences.

GIS Components

The key research focus of this study is to determine if second-by-second GPS data collected during personal vehicle travel be used to derive the essential data elements typically collected in the paper travel diary. The use of accurate GIS databases is a critical component for success in this process. These databases must be spatially accurate with all roadway centerlines for the study area, and must contain land use characteristics at the parcel level. In addition, integrated aerial photography is recommended for visual examination of questionable data.
Table 4-6: Key GPS Fields For Use in Travel Studies

<table>
<thead>
<tr>
<th>Field / Format</th>
<th>Use</th>
<th>NMEA sentence ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>UTC time of position fix, hhmmss format</td>
<td>time</td>
<td>RMC, field &lt;1&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGA, field &lt;1&gt;</td>
</tr>
<tr>
<td>UTC date of position fix, ddmmyy format</td>
<td>date</td>
<td>RMC, field &lt;9&gt;</td>
</tr>
<tr>
<td>Latitude, ddmm.mmmm format</td>
<td>y position</td>
<td>RMC, field &lt;3&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGA, field &lt;2&gt;</td>
</tr>
<tr>
<td>Latitude hemisphere, N or S</td>
<td>y position</td>
<td>RMC, field &lt;4&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGA, field &lt;3&gt;</td>
</tr>
<tr>
<td>Longitude, ddmm.mmmm format</td>
<td>x position</td>
<td>RMC, field &lt;5&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGA, field &lt;4&gt;</td>
</tr>
<tr>
<td>Longitude hemisphere, E or W</td>
<td>x position</td>
<td>RMC, field &lt;6&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGA, field &lt;5&gt;</td>
</tr>
<tr>
<td>Speed over ground, 0.0 to 999.9 knots</td>
<td>velocity</td>
<td>RMC, field &lt;7&gt;</td>
</tr>
<tr>
<td>Course over ground, 0.0 to 359.9 degrees, true</td>
<td>heading</td>
<td>RMC, field &lt;8&gt;</td>
</tr>
<tr>
<td>GPS quality: 0 = GPS fix not available</td>
<td></td>
<td>RMC, field &lt;8&gt;</td>
</tr>
<tr>
<td>1 = GPS only fix available</td>
<td></td>
<td>GGA, field &lt;6&gt;</td>
</tr>
<tr>
<td>2 = GPS and DGPS fix available</td>
<td>GPS/DGPS status</td>
<td>GGA, field &lt;6&gt;</td>
</tr>
<tr>
<td>Number of satellites in use, 00 to 12</td>
<td>accuracy</td>
<td>GGA, field &lt;7&gt;</td>
</tr>
<tr>
<td>Horizontal dilution of precision, 1.0 to 99.9</td>
<td>(x,y) accuracy</td>
<td>GGA, field &lt;8&gt;</td>
</tr>
<tr>
<td>Status: A= Valid Position, V= NAV receiver warning</td>
<td></td>
<td>RMC, field &lt;2&gt;</td>
</tr>
</tbody>
</table>

Matching of Elements

Table 4-7 shows the preliminary matching of the key travel diary elements, as defined in the 1993 Atlanta Final Report, to the GPS elements that are most likely to provide the desired travel diary field, along with the GIS components that are also expected to greatly assist the travel data derivation process. The remainder of this section of the dissertation contains expectations held prior to the study with respect to the collection and derivation of each travel diary element.
### Table 4-7: Matching Travel Diary Elements with Available GPS or GIS Data

<table>
<thead>
<tr>
<th>Key Travel Diary Elements</th>
<th>Derived from GPS Data</th>
<th>Derived from GIS data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trip Origin Address</td>
<td>Latitude and Longitude</td>
<td></td>
</tr>
<tr>
<td>Trip Start Time</td>
<td>First second of movement</td>
<td></td>
</tr>
<tr>
<td>Trip Destination Address</td>
<td>Latitude and Longitude</td>
<td></td>
</tr>
<tr>
<td>Trip Finish Time</td>
<td>Last second of movement</td>
<td></td>
</tr>
<tr>
<td>Travel Distance</td>
<td>GPS route (points)</td>
<td>GIS links</td>
</tr>
<tr>
<td>Trip Purpose</td>
<td>Coordinate match of pre-coded common trip ends</td>
<td>Origin and destination land uses (GIS)</td>
</tr>
<tr>
<td>Travel Mode(s)</td>
<td>** Vehicle only **</td>
<td>** Vehicle only **</td>
</tr>
<tr>
<td>Mode Details:</td>
<td>Driver or Passenger Vehicle Occupancy</td>
<td></td>
</tr>
</tbody>
</table>

**Trip Origin Address and Start Time**

The GPS data logger will start recording data as soon as the vehicle is powered. At this moment, the GPS data that are logged will contain the exact date, time, latitude and longitude of the vehicle. There is one key issue that may impact the GPS receiver’s ability to log this initial data point – the receiver’s signal acquisition time upon start up (which could range from 2 to 15 minutes) and which could possibly result in the loss of initial trip data. However, each trip’s destination coordinates can be used as the starting coordinates for the next trip. Examination of the trip data will be necessary whenever these two sets of coordinates are not close in distance. Finally, if continuous power-supply equipment configurations are used, start-up acquisition times are only an issue when the vehicle is parked in a location that does not have a clear sky view (e.g., garage, urban canyon, or dense tree canopy).
Assuming the GPS receiver is logging coordinates within a few seconds after the vehicle is started, the initial coordinates logged could still be inaccurate based on the type of GPS service used. Standard GPS positional accuracy with Selective Availability activated is 30 to 100 meters, whereas many DGPS solutions available for in-vehicle applications can provide accuracies within 3 to 10 meters. Of course, the accuracy of the underlying GIS database is also critical for accurate matching to parcel-level details, which will be necessary to derive the land use to be utilized in the trip purpose derivation.

Trip Destination Address and Finish Time

The GPS data logger will continue recording GPS data points for each second of each trip until the engine is turned off. Therefore, the exact date, time, latitude, and longitude of the vehicle upon reaching its destination will be captured. The only times in which this information will not be recorded are when there is a complete satellite signal obstruction, such as in an enclosed parking or private garage. However, the last point recorded prior to complete signal loss will be available, which should be sufficient for determining land-use for most trips. Of course, it may not be possible to determine exact trip purpose when the vehicle is parked in common areas, such as on-street parking and in parking lots and garages used by more than one business type.

Travel Distance (and Route)

One data element that is offered by GPS data logging that has not been available in traditional travel surveys is the actual trip route. The GPS points can be matched to a GIS base map, from which a simple algorithm can be applied to sum the lengths of the individual route links to obtain the total travel distance of each trip. Other than the trip
start and finish issues identified previously, it is possible that the GPS signal could
experience some periodic signal losses during travel through urban canyons or dense tree
canopies. However, it is likely that there will be some points captured through these
areas that will enable recovery of the complete trip route and travel distance.

Trip Purpose

Given the spatial data that are available for each trip end, it should be possible to
access a land-use database within a GIS to determine the land-use codes for most origins
and destinations. It is also feasible that these land-use pairs could be used to predict trip
purposes. This element should prove to be the most challenging in the analysis process.

One possible method for improving the success of this derivation process would be
to collect address information during the recruitment call for the most common
destinations. Home address, along with work addresses and/or school addresses for each
household member, could be geocoded prior to the GPS data processing, which would
greatly improve the derivation process. For example, the 1990 Atlanta survey found that
35.6% of trips originated at home, 20% originated at work, 6% originated at school, and
4.5% originated as a drop off or pick up passenger. If home, work, and school addresses
are collected in advance, then it is possible that, based on the 1990 Atlanta trip data, 66%
of the trip ends could be known in advance and simply confirmed through the process –
leaving only one third to be derived from scratch.

Travel Mode

For the purpose of this study, travel mode is limited to personal vehicle only.
Mode Details

For vehicle trips, the only detail typically captured is whether the study person is the driver or a passenger in the vehicle. If the person is the driver, then typically the number of passengers other than the driver is also collected. These details could possibly be determined based on trip purposes and destinations, but most likely are the one category that, if required, will need to be collected during the telephone interview.

Summary

This dissertation presents a case study in which the use of GPS data loggers is tested as a replacement to traditional travel diary data collection. Study participants are given GPS data loggers and paper diaries on which they are instructed to collect travel data for a 3-day period. The data are then collected, processed, and analyzed to determine the feasibility of eliminating the paper diary from future household travel surveys. Table 4-8 shows the inventory of GPS and GIS inputs that will be used to test the derivation of trip details.

In a real-world application of this approach, the households would still need to provide the traditional household, person, and vehicle data that is currently being collected by CATI operators. However, in this scenario, there is no travel diary (paper or electronic). Instead of travel diaries, GPS data loggers are installed in the respondent’s vehicle and collected after the study period. The GPS data is then processed to derive all traditional travel data elements possible, and then verified or completed during the CATI household, person, and vehicle data retrieval process.
Table 4-8: Inputs and Outputs of GPS Data Processing Analysis

<table>
<thead>
<tr>
<th>GPS elements (second-by-second)</th>
<th>GIS elements</th>
<th>Vehicle trip details</th>
</tr>
</thead>
<tbody>
<tr>
<td>date</td>
<td>spatially accurate, comprehensive GIS database</td>
<td>trip origin address</td>
</tr>
<tr>
<td>time</td>
<td>roadway centerlines</td>
<td>trip start time</td>
</tr>
<tr>
<td>latitude</td>
<td>land use characteristics at the parcel level</td>
<td>trip destination address</td>
</tr>
<tr>
<td>longitude</td>
<td>integrated aerial photography</td>
<td>trip finish time</td>
</tr>
<tr>
<td>speed</td>
<td></td>
<td>travel distance</td>
</tr>
<tr>
<td>heading</td>
<td></td>
<td>trip purpose / activity</td>
</tr>
<tr>
<td>accuracy measures</td>
<td></td>
<td>driver or passenger</td>
</tr>
<tr>
<td></td>
<td></td>
<td>if driver, number of passengers</td>
</tr>
</tbody>
</table>
CHAPTER V

DATA COLLECTION

This proof-of-concept study was designed to evaluate if GPS data logged during travel in personal vehicles could be used to derive trip details typically recorded by travel survey respondents in traditional paper diaries. To meet this objective, thirty survey participants were initially recruited; all recruits were affiliated with the Georgia Tech School of Civil and Environmental Engineering. The first six participants were used to test the GPS data logger equipment installation and use procedures. The remaining 24 participants were given both the GPS data logging equipment and paper diaries, and were asked to use both to record and capture all in-vehicle travel for the duration of their assigned survey period.

The equipment package installed was one of three possible GPS data logging packages (as described in the next section). The different packages contained different GPS and DGPS receivers so that analyses could be made regarding the ability to accurately derive travel data based on varying accuracy levels available in the GPS data. Each respondent was given a memory jogger trip sheet on which they were instructed to record each trip and the corresponding trip details – date, start time, finish time, destination, address or cross roads, purpose, and distance traveled (see Appendix A, Example 4). The survey periods were approximately three days, implemented as two waves per week, in either a Friday through Sunday or Tuesday through Friday cycle.
Equipment Used

To support the in-vehicle GPS data collection activity, four GPS data loggers were developed, each using a Palm IIIx handheld PC for recording the GPS data. With the Palm IIIx as the data logger, there were four possible GPS configurations:

P1: Garmin 35LP with FM sub-carrier differential corrections (DCI’s RDS 3000 receiver with a Radio Shack whip antenna)

P2: Garmin 35LP with radio beacon differential corrections (CSI’s ABX receiver with loop antenna)

P3: Garmin 35LP without DGPS

P4: Garmin II Plus without DGPS

Table 5-1 contains a description of the components of all deployed packages.

In addition to maintaining the same Palm product for each GPS data logger, the same power supplies were also used for each package – the vehicle’s power supply was used to power the GPS equipment via a cigarette lighter adapter and two AAA batteries were used to power each Palm. Finally, a camera bag was provided with each package in which the cabling, Palm IIIx, and DGPS receiver, if applicable, was stored.

The equipment packages were developed using off-the-shelf hardware. For a real world survey deployment, the packages would be fully customized and integrated to minimize size and power demand while providing maximum durability. The reports on the two Atlanta studies (CETMS and SMARTRAQ) contain recommendations for deployment system integration and durability (Wolf et al. 1999a, 2000b).
<table>
<thead>
<tr>
<th><strong>Table 5-1: Equipment Used in Passive In-Vehicle GPS Data Loggers</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Palm IIIx" /></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><img src="image" alt="Garmin GPS35LP" /></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><img src="image" alt="Garmin GPS II Plus" /></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><img src="image" alt="DCI's RDS-3000" /></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
The differences in the GPS equipment were designed to allow for comparisons of data processing performance using corrected (DGPS) versus uncorrected (GPS) data. Packages P1 and P2 both included a DGPS receiver so that the GPS data stream could be corrected to within 3 to 10 meters. There was one serious problem, however, with the cabling provided to connect the CSI antenna with the CSI receiver; the cables had a faulty connection at the receiver end and even the replacement cables had the same defect. As a result, the research team was forced to remove the second package (P2) from the study until the vendor could provide improved cabling. Unfortunately, the enhanced cabling did not arrive in time for use in the study and therefore only one DGPS package was deployed.

Table 5-2: Equipment Package Components

<table>
<thead>
<tr>
<th>Equipment Package</th>
<th>Palm IIIx</th>
<th>Data Cable</th>
<th>Garmin 35LP GPS Receiver/ Antenna</th>
<th>Garmin II Plus GPS Receiver</th>
<th>Sigem GPS Patch Antenna</th>
<th>DCI DGPS Receiver</th>
<th>DGPS Antenna</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>P3</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5-2 shows the exact equipment contained in each package. Both packages P1 and P3 used the Garmin GPS 35LP combined receiver and antenna. Package P4 contained the Garmin GPS II Plus as an alternative form factor. The trade-offs between the two forms include the 35LP as an all-in-one solution that is approximately the size of a computer mouse versus a handheld GPS receiver with either a small, connected antenna or a small patch antenna that can be mounted remotely with the II Plus. Finally, it should
be noted that, although both receivers capture GPS signals at 1 second interval, the
Garmin II Plus is not able to log at that frequency due to the processing requirements of
updating the display screen every second. Therefore, Garmin states that the II Plus can,
at best, log at two-second frequencies. Two second logging frequencies were determined
to be acceptable for the requirements of this study and the II Plus was included in the test plan.

Data Collection Procedures

Respondents were given simple instructions for installing and operating the GPS
equipment package. (See Appendix C for a copy of the instruction sheet for equipment
package P1.) A member of the research team performed the installation while the
participant observed and reviewed the instructions. The directions for installation were
provided in case the respondent planned to use a different vehicle for the next several
days and wanted to move the equipment. Also, the operating instructions also allowed
the respondent to bring the antenna(s) inside the vehicle in case of rain so that water
would not enter the interior of the vehicle through the window opening. In this case,
respondents were asked to return the antenna(s) back on the roof as soon as the rain
ceased. Finally, since the Palm IIIx is self-powered, respondents were instructed to
quickly swap out the two AAA batteries with two new batteries supplied in the camera
bag in the event of a “low power” message displayed on the Palm’s screen.

Since the GPS equipment was powered by the vehicle (recall that the Austin study
referred to this as “switched power” as compared to “continuous power”), the GPS
receiver would attempt to acquire satellite signals as soon as the vehicle was started and would automatically stop receiving data as soon as the vehicle was powered down. Each Palm IIIx device had Data Logger, the GPS data logging application, installed. The respondents were instructed to power on the Palm each time they started their vehicle, tap on a Start icon to begin logging, and then to tap again on a Stop icon when the trip was complete.

Data Logger was programmed to log second-by-second GPS data in a moving vehicle. The application also contains logic to reduce storage requirements; for example, the software would stop logging if 60 seconds of non-movement were detected and would restart as soon as movement was detected again. To reduce processing burden on the Palm, it was decided to use zero or near-zero speeds as non-movement indicators rather than to calculate distance between consecutive points. The exact GPS data logging rules programmed in Data Logger are:

1) If the GPS receiver is not getting a valid GPS signal, do not log the record.
2) If the speed has fallen below 3 mph (in GPS mode) or 1 mph (in DGPS mode) for more than 60 consecutive records, stop logging records beginning with the 61st record.
3) Start recording again as soon as the speed exceeds the appropriate threshold.

If the first two conditions were not met, then Data Logger takes all GPS messages received and transferred by the GPS equipment, selects the desired fields, and writes these fields in a record for each second of the trip. The second-by-second data elements are date, time, latitude, longitude, speed, heading, DGPS flag, number of satellites used, HDOP, and DGPS age. Data Logger also has a few user-defined parameters, including
the non-movement speed thresholds for GPS and DGPS modes, the countdown period before the data logging would stop once the non-movement condition was detected, and the number of hours offset between UTC (i.e., universal time) and local time. For this study, the time offset was set to –5 from the beginning of the study in March through April 1 and then to –4 for the remaining survey period (due to the switch to Daylight Savings Time, also known as Eastern Daylight Time or EDT, on April 1).

Upon completion of the survey period, a research team member accompanied the participants to their vehicle to remove the equipment and to collect both the paper trip diary and any feedback from each participant. The equipment was returned to the research office, the data were downloaded to a PC via the Palm HotSynch Cradle, the transfer was confirmed via visual inspection of the file contents, and then the data were cleared from the Palm device. Each equipment package was then reassembled and all connections confirmed as secure, and each Palm’s batteries were checked for sufficient power for another survey.

A copy of a completed trip diary and a segment of the GPS data file corresponding to the first trip made by that respondent can be seen in Figure 5-1 and Table 5-3.
Table 5-3: Example of GPS Data File Logged on Palm IIIx

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Heading</th>
<th>Speed</th>
<th>DGPS Flag</th>
<th># SV</th>
<th>HDOP</th>
<th>DGPS Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>20000331</td>
<td>194153</td>
<td>33.78149</td>
<td>-84.3901</td>
<td>90.5</td>
<td>14.9</td>
<td>2</td>
<td>7</td>
<td>1.3</td>
<td>1</td>
</tr>
<tr>
<td>20000331</td>
<td>194154</td>
<td>33.78149</td>
<td>-84.3900</td>
<td>91.1</td>
<td>13.2</td>
<td>2</td>
<td>7</td>
<td>1.3</td>
<td>0</td>
</tr>
<tr>
<td>20000331</td>
<td>194155</td>
<td>33.78149</td>
<td>-84.3900</td>
<td>91.2</td>
<td>11.3</td>
<td>2</td>
<td>7</td>
<td>1.3</td>
<td>1</td>
</tr>
<tr>
<td>20000331</td>
<td>194156</td>
<td>33.78149</td>
<td>-84.3899</td>
<td>90.6</td>
<td>9.4</td>
<td>2</td>
<td>7</td>
<td>1.3</td>
<td>0</td>
</tr>
<tr>
<td>20000331</td>
<td>194157</td>
<td>33.78149</td>
<td>-84.3899</td>
<td>91.1</td>
<td>7.3</td>
<td>2</td>
<td>7</td>
<td>1.3</td>
<td>1</td>
</tr>
<tr>
<td>20000331</td>
<td>194158</td>
<td>33.78149</td>
<td>-84.3899</td>
<td>92.1</td>
<td>5.5</td>
<td>2</td>
<td>6</td>
<td>1.3</td>
<td>0</td>
</tr>
<tr>
<td>20000331</td>
<td>194159</td>
<td>33.78149</td>
<td>-84.3898</td>
<td>92</td>
<td>3.7</td>
<td>2</td>
<td>6</td>
<td>1.3</td>
<td>1</td>
</tr>
<tr>
<td>20000331</td>
<td>194200</td>
<td>33.78149</td>
<td>-84.3898</td>
<td>92</td>
<td>3.5</td>
<td>2</td>
<td>6</td>
<td>1.3</td>
<td>0</td>
</tr>
<tr>
<td>20000331</td>
<td>194201</td>
<td>33.78149</td>
<td>-84.3898</td>
<td>91.9</td>
<td>3.7</td>
<td>2</td>
<td>6</td>
<td>1.3</td>
<td>1</td>
</tr>
<tr>
<td>20000331</td>
<td>194203</td>
<td>33.78149</td>
<td>-84.3898</td>
<td>91.3</td>
<td>3.3</td>
<td>2</td>
<td>6</td>
<td>1.3</td>
<td>1</td>
</tr>
<tr>
<td>20000331</td>
<td>194204</td>
<td>33.78149</td>
<td>-84.3898</td>
<td>91.3</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>1.3</td>
<td>0</td>
</tr>
<tr>
<td>20000331</td>
<td>194205</td>
<td>33.78149</td>
<td>-84.3898</td>
<td>91.3</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>1.3</td>
<td>1</td>
</tr>
</tbody>
</table>
### Figure 5-1: Completed Memory Jogger

<table>
<thead>
<tr>
<th>Trip #</th>
<th>Day of Week</th>
<th>Trip Start Time</th>
<th>Trip Stop Time</th>
<th>Destination / Place Name</th>
<th>Address / Street &amp; Cross Street</th>
<th>Activity at Destination</th>
<th>Travel Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FRIDAY</td>
<td>7:40 PM</td>
<td>8:18</td>
<td>HOME</td>
<td>2872 CATHEDRAL PARK CIR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>WHITE</td>
<td>8:45</td>
<td>8:57</td>
<td>IN-LAWS</td>
<td>5416 MISSION RD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>WHITE</td>
<td>10:18</td>
<td>10:30</td>
<td>HOME</td>
<td>3432 DA WOODS LN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>SATURDAY</td>
<td>2:35</td>
<td>3:05</td>
<td>FIRST PURPOSE</td>
<td>5631 FAIRFAX AVE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>SATURDAY</td>
<td>3:15</td>
<td>3:25</td>
<td>PUBLIC LIBRARY</td>
<td>5638 FAIRFAX AVE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>SATURDAY</td>
<td>3:30</td>
<td>3:45</td>
<td>NEXT PAPER</td>
<td>5638 FAIRFAX AVE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>SATURDAY</td>
<td>3:35</td>
<td>4:00</td>
<td>TEXACO</td>
<td>5638 FAIRFAX AVE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>SATURDAY</td>
<td>4:08</td>
<td>4:12</td>
<td>HOME</td>
<td>3720 FAIRFAX AVE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>SATURDAY</td>
<td>7:50</td>
<td>8:05</td>
<td>GMA FILM</td>
<td>2891 CATHEDRAL PARK CIR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>SATURDAY</td>
<td>8:10</td>
<td>8:30</td>
<td>WAL-MART</td>
<td>2891 CATHEDRAL PARK CIR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>SATURDAY</td>
<td>9:30</td>
<td>9:50</td>
<td>HOME</td>
<td>2891 CATHEDRAL PARK CIR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>SATURDAY</td>
<td>11:45</td>
<td>12:05</td>
<td>GMA FILM</td>
<td>2891 CATHEDRAL PARK CIR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>SATURDAY</td>
<td>1:10</td>
<td>1:30</td>
<td>GMA FILM</td>
<td>2891 CATHEDRAL PARK CIR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>SATURDAY</td>
<td>2:30</td>
<td>2:45</td>
<td>CEE-GE</td>
<td>2891 CATHEDRAL PARK CIR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>SATURDAY</td>
<td>3:30</td>
<td>4:00</td>
<td>CEE-GE</td>
<td>2891 CATHEDRAL PARK CIR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>SATURDAY</td>
<td>4:30</td>
<td>5:00</td>
<td>CEE-GE</td>
<td>2891 CATHEDRAL PARK CIR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>SATURDAY</td>
<td>7:15</td>
<td>7:30</td>
<td>HOME</td>
<td>2891 CATHEDRAL PARK CIR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>SATURDAY</td>
<td>7:45</td>
<td>8:00</td>
<td>GMA FILM</td>
<td>2891 CATHEDRAL PARK CIR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>SATURDAY</td>
<td>8:15</td>
<td>8:30</td>
<td>GMA FILM</td>
<td>2891 CATHEDRAL PARK CIR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>SATURDAY</td>
<td>10:05</td>
<td>10:30</td>
<td>GMA FILM</td>
<td>2891 CATHEDRAL PARK CIR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>SATURDAY</td>
<td>10:20</td>
<td>10:40</td>
<td>GMA FILM</td>
<td>2891 CATHEDRAL PARK CIR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>SATURDAY</td>
<td>10:45</td>
<td>11:00</td>
<td>GMA FILM</td>
<td>2891 CATHEDRAL PARK CIR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>SATURDAY</td>
<td>11:00</td>
<td>11:30</td>
<td>GMA FILM</td>
<td>2891 CATHEDRAL PARK CIR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>SATURDAY</td>
<td>11:30</td>
<td>11:45</td>
<td>CEE-GE</td>
<td>2891 CATHEDRAL PARK CIR</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data Collection Results

From March 23 through April 28, 2000, thirty survey participants collected GPS data with three types of GPS data loggers installed in their vehicles for periods of three days. The first six participants were used to pretest the installation and usage procedures for the equipment packages and therefore were not given paper diaries. These pretest participants are not included in the trip derivation analyses and statistics. However, their qualitative observations with respect to equipment installation, functionality, and concerns are included in the corresponding summaries. In addition, the trips logged by these six packages were used to develop and calibrate the trip detection macros using manual (i.e., visual inspection) processes.

Of the 24 respondents who were given both a GPS equipment package and a paper diary, four experienced equipment problems and therefore did not complete the data collection effort. Three of these four failures appeared to be related to a faulty Palm IIIx (used in package P3) that did not power down correctly. Consequently, these four participants were removed from the analysis. Finally, one respondent did not return his paper diary despite repeated reminders. As a result, there were 19 respondents who both successfully collected travel data with the GPS data logger and who returned a completed a paper trip diary. These are the samples that are the basis of the GPS data to trip details analysis.

Table 5-4 provides a summary of the participants by home county and equipment package. There was a fairly even distribution of participants by home county across Fulton (7), Dekalb (7), and Cobb (5). Six of the participants used the DGPS package.
with 10-meter accuracy and the remaining 13 used a GPS-only package (with 30 to 100 meter accuracy). Nine participated on the weekend schedule (Friday afternoon through Monday morning) and the remaining 10 participated on the weekday schedule (Tuesday afternoon through Friday morning). A more detailed list of participants is provided in Table 5-5. Note that most participants worked in the School of Civil and Environmental Engineering at Georgia Tech; these participants have been given work place abbreviations of Mason (e.g., the Mason Building) or SEB (the Sustainable Education Building) in Table 5-5.

Sixteen of the participants work at Georgia Tech in the School of CEE or College of Computing; the three others were recruited wives or friends of previous participants, or of current faculty or researchers. An attempt was made to recruit participants who did not work at Georgia Tech; however, it was difficult to obtain final commitments from these participants and to get access to them and their vehicles within the short time period in which this study was conducted. It is recommended that the next phase of this research focus on a more random distribution of work locations among the recruited participants in order to perform a better test of work location and mid-day trips at some distance away from the Georgia Tech campus.
Table 5-4: Summary of Participants by Home County and Equipment Package

<table>
<thead>
<tr>
<th>Home County</th>
<th>P1 -- DGPS</th>
<th>P3 -- GPS only</th>
<th>P4 -- GPS only</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fulton</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Dekalb</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Cobb</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>6</td>
<td>5</td>
<td>8</td>
<td>19</td>
</tr>
</tbody>
</table>

Table 5-5: List of Participants

<table>
<thead>
<tr>
<th>No.</th>
<th>Study ID</th>
<th>Home County</th>
<th>Work Place</th>
<th>Survey Date</th>
<th>Pkg</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>Dekalb</td>
<td>Mason</td>
<td>3/31- 4/3</td>
<td>P1</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>Cobb</td>
<td>SEB</td>
<td>3/31- 4/3</td>
<td>P4</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>Cobb</td>
<td>SEB</td>
<td>4/4 – 4/7</td>
<td>P3</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>Cobb</td>
<td>SEB</td>
<td>4/4 – 4/7</td>
<td>P4</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>Dekalb</td>
<td>SEB</td>
<td>4/7 – 4/10</td>
<td>P1</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>Dekalb</td>
<td>SEB</td>
<td>4/7 – 4/10</td>
<td>P4</td>
</tr>
<tr>
<td>7</td>
<td>16</td>
<td>Fulton</td>
<td>Mason</td>
<td>4/11 –4/14</td>
<td>P1</td>
</tr>
<tr>
<td>8</td>
<td>17</td>
<td>Fulton</td>
<td>SEB</td>
<td>4/11 –4/14</td>
<td>P3</td>
</tr>
<tr>
<td>9</td>
<td>18</td>
<td>Fulton</td>
<td>Mason</td>
<td>4/11 –4/14</td>
<td>P4</td>
</tr>
<tr>
<td>10</td>
<td>19</td>
<td>Fulton</td>
<td>SEB</td>
<td>4/14 –4/17</td>
<td>P1</td>
</tr>
<tr>
<td>11</td>
<td>20</td>
<td>Dekalb</td>
<td>SEB</td>
<td>4/14 –4/17</td>
<td>P3</td>
</tr>
<tr>
<td>12</td>
<td>21</td>
<td>Dekalb</td>
<td>College of Computing</td>
<td>4/14 –4/17</td>
<td>P4</td>
</tr>
<tr>
<td>13</td>
<td>23</td>
<td>Dekalb</td>
<td>Home</td>
<td>4/18 –4/21</td>
<td>P3</td>
</tr>
<tr>
<td>14</td>
<td>24</td>
<td>Cobb</td>
<td>Hartsfield Intl Airport</td>
<td>4/18 –4/21</td>
<td>P4</td>
</tr>
<tr>
<td>15</td>
<td>25</td>
<td>Cobb</td>
<td>SEB</td>
<td>4/21 –4/24</td>
<td>P1</td>
</tr>
<tr>
<td>16</td>
<td>27</td>
<td>Fulton</td>
<td>SEB</td>
<td>4/21 –4/24</td>
<td>P4</td>
</tr>
<tr>
<td>17</td>
<td>28</td>
<td>Fulton</td>
<td>Forsyth St</td>
<td>4/25 – 5/1</td>
<td>P1</td>
</tr>
<tr>
<td>18</td>
<td>29</td>
<td>Fulton</td>
<td>Mason</td>
<td>4/25 – 4/28</td>
<td>P3</td>
</tr>
<tr>
<td>19</td>
<td>30</td>
<td>Dekalb</td>
<td>SEB</td>
<td>4/25 – 4/28</td>
<td>P4</td>
</tr>
</tbody>
</table>
Participant Feedback

As already mentioned, participants were encouraged to report any problems experienced with the equipment or to provide any recommendations they might have with respect to future studies conducted with the equipment packages. Below are a few of the comments and problems reported, along with possible solutions, where appropriate. Additional suggestions provided by the research team member responsible for equipment installation have also been included. The comments fall under four general categories: GPS power supply, GPS antenna and cabling, Palm data cable, and Palm usage and power supply.

GPS Power Supply

• Some American cars have cigarette lighters that are powered on all the time. This means that the electrical power from the car’s battery could be completely drained, even if the vehicle is parked. During this study, participants who had this type of power were asked to disconnect the cigarette lighter adapter cable whenever they turned off the car and to connect it again when they started the engine. However, those participants who did not disconnect the cable reported no vehicle power problems.

GPS Antenna and Cabling

• Before installation of the equipment, installer should explain details and issues of installation and should ask the participant which window is preferred for GPS cable wiring.
• If a vehicle has a sunroof, it is very convenient to use it for cable wiring. However, some cars are designed in such a way that the sunroof closes tightly once engine is shut off, and this could damage the GPS cables.

• In light-duty trucks, use the rear sliding windows (if available) for GPS cable wiring.

• Cabling run through open windows can be damaged if participant closes window tightly on cable, especially if the vehicle has electric windows and the participant wants to minimize the chance of rain entry into the vehicle interior.

• A few participants expressed strong concerns about placing the GPS antenna on the roof of the car and about the window not closing completely. They insisted on putting the antenna inside of their vehicle. This could cause accuracy problems in the GPS data.

• One participant reported that the small patch magnetic-mount antenna scratched the paint surface of the roof of the vehicle. He recommended that a thin piece of soft fabric be used as a layer between the magnetic mount and the roof to eliminate this problem.

**Palm Data Cable**

• The data cable for the Palm IIIx is too short. However, longer cables also become burdensome.

• The connector between the data cable and the Palm is not rugged enough. Velcro strips were added, which partially solved the problem. However, it was still possible for the cable to disconnect.
Palm Usage and Power Supply

- Installer should ask the participant to use the Palm and to verify that he/she understands how to operate the handheld computer and to use the application software.
- Installer should delete collected GPS data after practice exercise is over.
- Some participants forgot to turn on the GPS data logger (the Palm) while traveling. As a result, some trips will not be included in the data file. This problem would be solved if the device were powered remotely and left on all the time.
- Some participants forgot to turn off the Palm, and, as a result, the internal batteries of the Palm went dead and all data were lost -- this is the case for 3 of the 4 equipment package failures. The best solution would be to provide power continuously throughout the survey period.
- It was observed that the Palm could be difficult to turn on and off. This could cause power and data loss problems if unresolved.
- The current Palm software allows the user to exit the GPS data logging application. This option should be removed prior to real-world deployment.
- The LCD display on the Palm screen was difficult to read after the vehicle was parked for a long time in the sun.
CHAPTER VI

DATA PROCESSING

Once the GPS data are collected and transferred successfully to a PC, the GPS comma-delimited text files are ready for processing. These GPS trip files contain the second-by-second date, time, position, speed, heading, HDOP, satellite, and DGPS information for every trip made by each survey participant – assuming that the participant started the GPS data logger for each trip and that the power and data cables were connected securely. The primary goal of the GPS data processing steps is to generate the standard trip details typically recorded on paper by travel survey respondents.

These GPS text files were first translated into DBF files (i.e., a generic database format) for visual inspection within TransCAD, a commercially available GIS developed specifically for transportation applications. Although a more thorough process would include a data cleaning step to identify potential problem data by visual inspection or by macro detection of possible outliers in speed, location, or time, as well as identification of points with poor GPS accuracy measures, this research did not attempt to clean the GPS data files prior to running the trip detection macro. Instead, the trip detection macro was run directly on the raw GPS data file recorded by the GPS data logger. The raw nature of the data was deemed to be a better test (in the sense of a worst case scenario) of the data processing steps.
The four primary steps used to transform the raw GPS second-by-second data into trip-level details are:

1) Trip Detection
2) Land Use and Address Assignment
3) Trip Purpose Derivation
4) Travel Route Determination and Distance Calculation

The purpose of Trip Detection is to identify potential trip ends within the GPS data stream; this is accomplished by locating segments in the data stream containing no vehicle movement for a pre-defined, minimum time period. A TransCAD macro was written to detect these gaps in vehicle movement, with the minimum time threshold initially set at 120 seconds, and then lowered to 90 seconds and 60 seconds. The purpose of these different values was to evaluate the appropriate vehicle non-movement threshold to define a stop without missing too many short stops while simultaneously not mis-identifying periods of extended traffic congestion or signal-related delays as stops. The underlying philosophy of this approach is that it is better to over-predict trips and then correct down than to under-predict trips. Although this approach is more labor intensive, there is less chance of missing trips and it is easier to make corrections. Of course, it also does not make sense to detect too many false trips.

Once this macro was tested at the various thresholds and calibrated using the first six GPS data sets collected by the pre-test participants (those who were not given trip diaries), the macro was then run on each of the 19 survey participant data files at each
threshold. The results at each threshold for each participant were evaluated and one threshold was selected as the best level for detecting true trip ends. The final output of this step was a trip ends file containing the trip number and the starting and ending coordinate information, including date, time, and position.

The trip ends file was next used to meet two separate needs: 1) to identify the nearest approximate land-use codes and addresses for the trip ends within a GIS (i.e., the Land Use and Address Assignment step); and 2) to identify the individual routes taken by the participants along with the distance traveled (i.e., the Travel Route Determination and Distance Calculation step). Consequently, the data processing steps diverged with these two analysis efforts. The results of the Land Use and Address Assignment step are then used to derive trip purpose. Upon successful completion of all steps, the data obtained from each are then combined with the information in the trip ends file to completely describe the derived trip. Results of these steps are summarized at the trip level by respondent and then compared to the recorded paper diary trip data. Table 6-1 contains one trip from one participant with both the reported trip information located on the first line and the GPS-derived trip information on the second line. The results of all comparisons are presented in Chapter 7.

Table 6-1: Preview of Reported versus Derived Trip Results

<table>
<thead>
<tr>
<th>Trip #</th>
<th>Week Day</th>
<th>Trip Start Time</th>
<th>Trip Stop Time</th>
<th>Destination / Place Name</th>
<th>Address / Street &amp; Cross Street</th>
<th>Activity at Destination</th>
<th>Travel Dist.</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Fri</td>
<td>22:15</td>
<td>22:30</td>
<td>home</td>
<td>3433 CasaWoods Ln</td>
<td>home</td>
<td>5.0</td>
</tr>
<tr>
<td>3</td>
<td>Fri</td>
<td>22:13</td>
<td>22:27</td>
<td>Single Family Residence</td>
<td>3433 Casa Woods Ln</td>
<td>Return Home</td>
<td>5.1</td>
</tr>
</tbody>
</table>
Trip Detection

The first step in processing the GPS data is the identification of individual trips within the second-by-second data stream. Software was developed to parse through these individual vehicle travel files and to detect breaks in time and/or periods of vehicle non-movement. This step is clearly non-trivial since it addresses one of the key problems with travel diary studies – that of trip underreporting.

Trip Detection Process

Researchers used TranCAD, a transportation-oriented GIS, to visually inspect the initial GPS data and to write software to pass through the data looking for individual trips. This macro was written in TransCAD’s GISDK (GIS Developer’s Kit), which uses Caliper Script™ (a proprietary macro programming language). The code for the Trip Detection Macro is provided in Appendix D. Periods of vehicle non-movement were used as the primary indicator of a stop.

As mentioned in the review of the Austin GPS travel study, there are two power options currently implemented for logging GPS data during in-vehicle trips – the continuous power option, in which power is provided to the GPS receiver 24-hours per day, and the switched power option, in which the power is only provided when the vehicle is powered on. In either case, if the GPS data logging software is written so that any detection of vehicle non-movement beyond a certain time threshold (e.g., 60 seconds) causes the temporary suspension of data logging until movement is detected again, then the post processing software needs to detect a corresponding gap in the data.
logged. In the case of the continuous power option, each break should be preceded by the non-movement countdown; if the threshold is set at 60 seconds, then there will be 60 records of GPS data with zero or near-zero speeds immediately prior to the break in data points recorded.

Consequently, the Trip Detection Macro, which identifies individual trips, reads each record in the logged file sequentially, looking for gaps in the timestamp between successive records. The macro also looks for zero or near-zero speeds for a continuous period that would indicate a “stop-logging” countdown and that could also indicate breaks in travel. The GPS data captured during the last second of vehicle movement prior to a detected time gap is written to the trip file as a trip end. The macro then writes the first GPS record after the detected gap that contains a speed above the non-movement threshold to the trip file as the next trip start.

The output of the Trip Detection Macro is a file with two records per trip – one with each trip start’s GPS point information followed by the corresponding trip’s endpoint information. Table 6-2 contains an example trip file generated with the trip gap threshold set at 120 seconds. Note that these records also contain the record numbers of the start and end trip records in the second-by-second point data in the original GPS point file; these pointers are used to parse out the individual second-by-second trip files for use the travel route and distance process (as presented in the last section of this chapter).

With the exception of date change conditions, which are easily distinguished in the data stream, there are several reasons for time gaps between successive points in the GPS log file:
Table 6-2: Example of Trip Ends File

<table>
<thead>
<tr>
<th>TRIP #</th>
<th>TYPE</th>
<th>DATE</th>
<th>TIME (Sec)</th>
<th>LATITUDE</th>
<th>LONGITUDE</th>
<th>HEADING</th>
<th>SPEED</th>
<th>DGPS FLAG</th>
<th># SAT</th>
<th>HDOP</th>
<th>AGE</th>
<th>HOUR</th>
<th>MIN</th>
<th>SEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>START</td>
<td>20000407</td>
<td>67321</td>
<td>33.778038</td>
<td>-84.399976</td>
<td>0.0</td>
<td>1.0</td>
<td>2</td>
<td>9</td>
<td>1.3</td>
<td>1</td>
<td>18</td>
<td>42</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>END</td>
<td>20000407</td>
<td>68969</td>
<td>33.850105</td>
<td>-84.256755</td>
<td>134.6</td>
<td>1.9</td>
<td>2</td>
<td>8</td>
<td>1.0</td>
<td>0</td>
<td>19</td>
<td>9</td>
<td>29</td>
</tr>
<tr>
<td>2</td>
<td>START</td>
<td>20000407</td>
<td>75567</td>
<td>33.850160</td>
<td>-84.256759</td>
<td>78.8</td>
<td>3.1</td>
<td>2</td>
<td>6</td>
<td>1.2</td>
<td>1</td>
<td>20</td>
<td>59</td>
<td>27</td>
</tr>
<tr>
<td>2</td>
<td>END</td>
<td>20000407</td>
<td>76081</td>
<td>33.872316</td>
<td>-84.255381</td>
<td>71.9</td>
<td>1.0</td>
<td>2</td>
<td>6</td>
<td>1.6</td>
<td>1</td>
<td>21</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>START</td>
<td>20000408</td>
<td>30917</td>
<td>33.872231</td>
<td>-84.255418</td>
<td>314.5</td>
<td>1.0</td>
<td>2</td>
<td>7</td>
<td>1.7</td>
<td>1</td>
<td>8</td>
<td>35</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>END</td>
<td>20000408</td>
<td>32036</td>
<td>33.775769</td>
<td>-84.398896</td>
<td>1.3</td>
<td>1.0</td>
<td>2</td>
<td>9</td>
<td>1.0</td>
<td>0</td>
<td>8</td>
<td>53</td>
<td>56</td>
</tr>
<tr>
<td>4</td>
<td>START</td>
<td>20000408</td>
<td>63659</td>
<td>33.775808</td>
<td>-84.398930</td>
<td>208.1</td>
<td>1.0</td>
<td>2</td>
<td>7</td>
<td>1.4</td>
<td>1</td>
<td>17</td>
<td>40</td>
<td>59</td>
</tr>
<tr>
<td>4</td>
<td>END</td>
<td>20000408</td>
<td>65465</td>
<td>33.848563</td>
<td>-84.361411</td>
<td>124.1</td>
<td>1.0</td>
<td>2</td>
<td>8</td>
<td>1.0</td>
<td>1</td>
<td>18</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>START</td>
<td>20000408</td>
<td>71460</td>
<td>33.848643</td>
<td>-84.361343</td>
<td>67.8</td>
<td>4.4</td>
<td>2</td>
<td>6</td>
<td>1.2</td>
<td>0</td>
<td>19</td>
<td>51</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>END</td>
<td>20000408</td>
<td>72620</td>
<td>33.872293</td>
<td>-84.255344</td>
<td>114.7</td>
<td>1.0</td>
<td>2</td>
<td>7</td>
<td>1.7</td>
<td>1</td>
<td>20</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>START</td>
<td>20000409</td>
<td>34440</td>
<td>33.872175</td>
<td>-84.255426</td>
<td>71.2</td>
<td>1.0</td>
<td>2</td>
<td>9</td>
<td>1.2</td>
<td>1</td>
<td>9</td>
<td>34</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>END</td>
<td>20000409</td>
<td>35419</td>
<td>33.775658</td>
<td>-84.399008</td>
<td>2.2</td>
<td>1.0</td>
<td>2</td>
<td>9</td>
<td>1.1</td>
<td>0</td>
<td>9</td>
<td>50</td>
<td>19</td>
</tr>
<tr>
<td>7</td>
<td>START</td>
<td>20000409</td>
<td>54185</td>
<td>33.775781</td>
<td>-84.398860</td>
<td>158.6</td>
<td>1.0</td>
<td>2</td>
<td>5</td>
<td>1.9</td>
<td>1</td>
<td>15</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>END</td>
<td>20000409</td>
<td>55282</td>
<td>33.828058</td>
<td>-84.325258</td>
<td>1.3</td>
<td>1.0</td>
<td>2</td>
<td>5</td>
<td>1.7</td>
<td>2</td>
<td>15</td>
<td>21</td>
<td>22</td>
</tr>
<tr>
<td>8</td>
<td>START</td>
<td>20000409</td>
<td>58734</td>
<td>33.827803</td>
<td>-84.325381</td>
<td>303.9</td>
<td>1.0</td>
<td>2</td>
<td>7</td>
<td>1.1</td>
<td>0</td>
<td>16</td>
<td>18</td>
<td>54</td>
</tr>
<tr>
<td>8</td>
<td>END</td>
<td>20000409</td>
<td>59471</td>
<td>33.872178</td>
<td>-84.255401</td>
<td>216.0</td>
<td>1.0</td>
<td>2</td>
<td>7</td>
<td>1.5</td>
<td>0</td>
<td>16</td>
<td>31</td>
<td>11</td>
</tr>
<tr>
<td>9</td>
<td>START</td>
<td>20000409</td>
<td>66758</td>
<td>33.872160</td>
<td>-84.255501</td>
<td>209.9</td>
<td>8.1</td>
<td>2</td>
<td>7</td>
<td>1.8</td>
<td>1</td>
<td>18</td>
<td>32</td>
<td>38</td>
</tr>
<tr>
<td>9</td>
<td>END</td>
<td>20000409</td>
<td>67502</td>
<td>33.903234</td>
<td>-84.271350</td>
<td>191.6</td>
<td>1.4</td>
<td>2</td>
<td>8</td>
<td>1.1</td>
<td>0</td>
<td>18</td>
<td>45</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>START</td>
<td>20000409</td>
<td>68529</td>
<td>33.903194</td>
<td>-84.271354</td>
<td>191.6</td>
<td>1.1</td>
<td>2</td>
<td>8</td>
<td>1.0</td>
<td>1</td>
<td>19</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>END</td>
<td>20000409</td>
<td>69243</td>
<td>33.872233</td>
<td>-84.255463</td>
<td>310.7</td>
<td>1.2</td>
<td>2</td>
<td>7</td>
<td>1.6</td>
<td>1</td>
<td>19</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>START</td>
<td>20000410</td>
<td>29417</td>
<td>33.872193</td>
<td>-84.255458</td>
<td>42.2</td>
<td>1.1</td>
<td>2</td>
<td>5</td>
<td>2.2</td>
<td>0</td>
<td>8</td>
<td>10</td>
<td>17</td>
</tr>
<tr>
<td>11</td>
<td>END</td>
<td>20000410</td>
<td>30639</td>
<td>33.776198</td>
<td>-84.399571</td>
<td>186.4</td>
<td>0.0</td>
<td>2</td>
<td>7</td>
<td>1.4</td>
<td>1</td>
<td>8</td>
<td>30</td>
<td>39</td>
</tr>
</tbody>
</table>
1) Gaps that occur because the vehicle’s power was shut off (which is generally an indication of a trip stop, although some vehicles may stall periodically).

2) Gaps that occur because no GPS signal is available (the data may appear as though the vehicle was powered off, but signal obstructions are not likely to last for one continuous period of 60 seconds)

3) Gaps that occur because the trip logging software detected an extended period of vehicular non-movement beyond the countdown threshold while the vehicle was still powered on; this could have either been caused by congestion or a stop (such as a drop off or quick errand) during which the vehicle was not powered off.

The default non-movement countdown threshold on the Palm device was set to 60 seconds, which was initially defined as a reasonable minimum stop time for a typical mid-trip chain activity, such as a passenger drop-off or pick-up. This allows the post-processing software the flexibility to further test possible stop time-length options equal to or above 60 seconds. For example, in Table 6-3, the number of trips recorded on the memory jogger is listed, along with the number of trips detected by the macro when the stop time gap was defined at 120, 90, and 60 seconds, respectively. Before any comparisons are made between the trips found in each case, one could reasonably expect that the number of self-reported trips would be lower than trips detected due to omitted trips on the paper diary. It is also logical to predict that lower stop duration thresholds such as 60 seconds will result in more possible trips detected, since it is more likely that vehicle delays attributable to heavy congestion, accidents, or extended signals will also be detected due to the lower time threshold for stop classification.
**Table 6-3: GPS Tests – Trips Detected versus Trips Reported**

<table>
<thead>
<tr>
<th>No.</th>
<th>Study ID</th>
<th>Home County</th>
<th>Work Place</th>
<th>Survey Period</th>
<th>Pkg</th>
<th># trips paper</th>
<th># trips macro (120)</th>
<th># trips macro (90)</th>
<th># trips macro (60)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>Dekalb</td>
<td>GT</td>
<td>Fri-Mon</td>
<td>P1</td>
<td>21</td>
<td>19</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>Cobb</td>
<td>GT</td>
<td>Tue-Fri</td>
<td>P4</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>Cobb</td>
<td>GT</td>
<td>Fri-Mon</td>
<td>P3</td>
<td>10</td>
<td>10</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>Cobb</td>
<td>GT</td>
<td>Tue-Fri</td>
<td>P4</td>
<td>7</td>
<td>8</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>Dekalb</td>
<td>GT</td>
<td>Fri-Mon</td>
<td>P1</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>Dekalb</td>
<td>GT</td>
<td>Fri-Mon</td>
<td>P4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>16</td>
<td>Fulton</td>
<td>GT</td>
<td>Tue-Fri</td>
<td>P1</td>
<td>15</td>
<td>15</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>8</td>
<td>17</td>
<td>Fulton</td>
<td>GT</td>
<td>Tue-Fri</td>
<td>P3</td>
<td>20</td>
<td>14</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>9</td>
<td>18</td>
<td>Fulton</td>
<td>GT</td>
<td>Tue-Fri</td>
<td>P4</td>
<td>5</td>
<td>11</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>19</td>
<td>Fulton</td>
<td>GT</td>
<td>Fri-Mon</td>
<td>P1</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>20</td>
<td>Dekalb</td>
<td>GT</td>
<td>Fri-Mon</td>
<td>P3</td>
<td>12</td>
<td>13</td>
<td>17</td>
<td>24</td>
</tr>
<tr>
<td>12</td>
<td>21</td>
<td>Dekalb</td>
<td>GT</td>
<td>Fri-Mon</td>
<td>P4</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>13</td>
<td>23</td>
<td>Dekalb</td>
<td>home</td>
<td>Tue-Fri</td>
<td>P3</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>14</td>
<td>24</td>
<td>Cobb</td>
<td>airport</td>
<td>Tue-Fri</td>
<td>P4</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>15</td>
<td>25</td>
<td>Cobb</td>
<td>GT</td>
<td>Fri-Mon</td>
<td>P1</td>
<td>13</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>16</td>
<td>27</td>
<td>Fulton</td>
<td>GT</td>
<td>Fri-Mon</td>
<td>P4</td>
<td>8</td>
<td>10</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>17</td>
<td>28</td>
<td>Fulton</td>
<td>Downtown</td>
<td>Tue-Mon</td>
<td>P1</td>
<td>21</td>
<td>21</td>
<td>22</td>
<td>34</td>
</tr>
<tr>
<td>18</td>
<td>29</td>
<td>Fulton</td>
<td>Mason</td>
<td>Tue-Fri</td>
<td>P3</td>
<td>12</td>
<td>9</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>19</td>
<td>30</td>
<td>Dekalb</td>
<td>SEB</td>
<td>Tue-Fri</td>
<td>P4</td>
<td>12</td>
<td>15</td>
<td>18</td>
<td>27</td>
</tr>
</tbody>
</table>
The gap threshold analysis started at 120 seconds and was then decreased to 90 seconds and then to 60 seconds. It was apparent in the comparisons of trips reported to trips detected that 120 seconds was the most reasonable threshold when balancing the desire to identify all trips and trip stops against the unnecessary detection of signal and congestion delays. These tradeoffs are evident in the data presented in Table 6-3. No values greater than 120 seconds were tested because of concern that many short errand-type trips, as well as pick ups and drop offs, would go undetected.

The detection of potentially false trips is not nearly as problematic in the proposed process as is the missed detection of stops. Within the implementation framework proposed, trips that have very short stop durations can be analyzed quickly within a GIS to assess the physical location of the stop with respect to the underlying road network, property boundaries (if available), and aerial photography. If it is still not clear whether the detected vehicle stop was due to a traffic signal / congestion delay or to a true trip end, the CATI follow up process can be used to clarify this uncertainty. On the other hand, if the stop gap threshold is set so high that true stops (or trips) are undetected, it is possible this missed detection could go unnoticed throughout the remainder of the data processing steps. However, it is possible that visual inspection of the data could reveal that a stop occurred.

**Trip Detection Results**

During the gap threshold evaluation, the comparison of detected trips to reported trips produced four general classifications of results:
1) **Class 1: Complete Match** – those for which the 120-second threshold resulted in a complete match of detected versus reported trips (7 participants);

2) **Class 2: Detection of Trips Not Reported** – those for which the trip detection software identified additional trips beyond those reported on the paper diary (3 participants);

3) **Class 3: Non-Detection of Stops of Short Duration** – those for which the trip detection software did not identify passenger or item pick ups or drop offs (3 participants); and

4) **Class 4: Potential Problems** – those for which the trip detection results varied greatly from the reported trips, both in over-detecting and under-detecting trips (6 participants).

Table 6-4 summarizes the results of the trip comparisons made between the reported trips and detected trips using the 120-second stop threshold. An analysis of the comparisons made for each participant within each classification follows. Subsequent tables that list each participant will be sorted in order of class first, then study ID to allow better understanding of results by classification.
Table 6-4: Summary of Differences Between Reported and Detected Stops

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Pkg</th>
<th># Trips Paper</th>
<th># Trips Macro</th>
<th>Class</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>P1</td>
<td>21</td>
<td>19</td>
<td>3</td>
<td>Two drop off / pick ups not detected</td>
</tr>
<tr>
<td>9</td>
<td>P4</td>
<td>14</td>
<td>15</td>
<td>2</td>
<td>One extra trip detected – not clear if it is a travel delay or actual stop</td>
</tr>
<tr>
<td>11</td>
<td>P3</td>
<td>10</td>
<td>10</td>
<td>1</td>
<td>Complete match</td>
</tr>
<tr>
<td>12</td>
<td>P4</td>
<td>7</td>
<td>8</td>
<td>2</td>
<td>One extra trip detected – missed trip</td>
</tr>
<tr>
<td>13</td>
<td>P1</td>
<td>11</td>
<td>11</td>
<td>1</td>
<td>Complete match</td>
</tr>
<tr>
<td>15</td>
<td>P4</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>Complete match</td>
</tr>
<tr>
<td>16</td>
<td>P1</td>
<td>15</td>
<td>15</td>
<td>1</td>
<td>Complete match</td>
</tr>
<tr>
<td>17</td>
<td>P3</td>
<td>20</td>
<td>14</td>
<td>3</td>
<td>Six drop off / pick ups not detected</td>
</tr>
<tr>
<td>18</td>
<td>P4</td>
<td>5</td>
<td>11</td>
<td>4</td>
<td>Equipment &amp; cabling problems</td>
</tr>
<tr>
<td>19</td>
<td>P1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>Complete match</td>
</tr>
<tr>
<td>20</td>
<td>P3</td>
<td>12</td>
<td>13</td>
<td>2</td>
<td>One extra trip detected – missed trip</td>
</tr>
<tr>
<td>21</td>
<td>P4</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>Complete match</td>
</tr>
<tr>
<td>23</td>
<td>P3</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>Procedural problems</td>
</tr>
<tr>
<td>24</td>
<td>P4</td>
<td>6</td>
<td>8</td>
<td>4</td>
<td>Equipment &amp; cabling problems</td>
</tr>
<tr>
<td>25</td>
<td>P1</td>
<td>13</td>
<td>13</td>
<td>1</td>
<td>Complete match</td>
</tr>
<tr>
<td>27</td>
<td>P4</td>
<td>8</td>
<td>10</td>
<td>4</td>
<td>Equipment &amp; cabling problems</td>
</tr>
<tr>
<td>28</td>
<td>P1</td>
<td>21</td>
<td>21</td>
<td>3</td>
<td>Two drop off / pick ups not detected Two extra trips detected – both travel delays</td>
</tr>
<tr>
<td>29</td>
<td>P3</td>
<td>12</td>
<td>9</td>
<td>4</td>
<td>Procedural &amp; cabling problems</td>
</tr>
<tr>
<td>30</td>
<td>P4</td>
<td>12</td>
<td>15</td>
<td>4</td>
<td>Equipment &amp; cabling problems</td>
</tr>
</tbody>
</table>
Class 1: Complete Match

For 7 of the 16 participants, the trip detection algorithm resulted in a 100% correct match in the detection of trips reported. There were a total of 61 trips reported by and detected from these participants. Further examination of the GPS points within a GIS indicates that the macro functioned exactly as intended. Figure 6-1 shows the three trips detected for Participant #15. Although this is a rather low number of trips to use as an example, trying to show multiple trips on a single map, with most trips going to the same set of origins and destinations, becomes quite difficult as the number of trips increases.

Figure 6-1: Examples of Trips Detected for Complete Match (Participant #15)
Class 2: Detection of Trips Not Reported

For three of the study participants, the Trip Detection Macro identified one additional trip beyond what was reported. This difference is attributable to either a missed trip by the respondent (which was omitted on the diary) or to an extended delay during travel that was incorrectly identified by the macro as a stop. For the three participants in this group, two of the detected stops were clearly true stops; Participant #12 made a stop lasting 19 minutes and 37 seconds at a shopping center, and Participant #20 made a 4-minute, 36-second stop on the way to work at a strip mall. See Figure 6-2 for the GPS point display of these detected but unreported stops.

Figure 6-2: Detected Missed Trips for Participants # 12 (left) and #20 (right)
The “missed stop” detected from the third participant in this group was questionable. Participant #9 reported a trip made from home to a running trail, but a stop that lasted 2 minutes and 38 seconds was detected at a location near the exit of the participant’s apartment complex; it is possible that this could have been a passenger pick up (i.e., a real stop) or simply a traffic delay. The return trip from the running trail does not contain a stop; however, a passenger drop-off could have occurred in less than two minutes and consequently gone undetected by the macro. This is the kind of detected stop that could be clarified during the CATI follow-up call.

Class 3: Non-Detection of Stops of Short Duration

Recall that stops of short duration, the same type that are often omitted during paper diary surveys, may also not be picked up by the trip detection macro when the non-movement gap detection threshold is set at 120 seconds. Exploration of the differences found between the reported and detected trips in this category uncovered several definite patterns in the data itself that can be used to detect these types of trips since the threshold detection is not triggered for the corresponding stops.

For the three participants who fell into this classification, the first one (#7) had two short duration passenger drop offs that were undetected by the macro. The next participant (#17) in this class had six short passenger drop offs and pick ups that were related to a guest arriving and departing from the airport, and to carpooling during trips to and from work. Figure 6-3 shows two of these trips; the map on the left shows the “home to airport” and “airport to hotel” trips, and the map on the right shows a carpooling pick up after work. In both cases, only one trip was identified due to the short time duration
of the passenger pick ups. Finally, the third participant (#28) made two passenger drop-offs at a MARTA train station that were undetected.

Class 4: Potential Problems

The results of the Trip Detection Macro for the six participant data files contained in this class revealed significant differences between reported and detected trips. Further examination of the data streams within Microsoft Excel and within TransCAD revealed serious data problems that were apparently related to loose cabling which results in incomplete or truncated trip data. All three equipment packages used the same type of
data cable connecting the Palm IIIx to the GPS receiver. This cable has a simple “slip-in” connector with no clasps or screws available to secure the connection. During the pretest weeks, Velcro was purchased and attached to this connector and to the Palm devices to address this weakness in the overall package design. However, even the Velcro did not always prevent data cable disconnects from occurring.

It is worthwhile to note here that these equipment packages were not originally designed for this particular application. They were used in this pilot study because they were readily available and the data logging software was already developed for the original SMARTRAQ passive in-vehicle GPS data logging application -- this application has since been replaced with a package under development for another major research initiative sponsored by the National Highway Transportation Safety Administration, or NHTSA [Wolf 2000b]. It was not anticipated, however, that the Palm data cable connection would be a significant obstacle in the collection of complete GPS data files. Cabling concerns were also discovered in the Netherlands handheld electronic travel diary study [Draijer 2000].

Here is a brief summary of the comparison results for these six “problem” data sets:

1) Participant #18: This GPS data set had numerous stops detected that were not reported. These appeared initially to represent heavy congestion, but are more likely attributable to combination of unreported stops, congestion delays, and loose cabling.
2) Participant #23: There were no GPS data for diary day 1, although two trips were reported. There is a GPS data gap for first half of the first trip on day 2, and a missed stop in the diary for diary day 2.

3) Participant #24: This diary reports six trips – three going to work each morning of the survey period and three returning home at the end of each work day. The data indicate frequent cable disconnects during most of these trips. Only partial trip data were available for beginning and ending segments of the trips, which was most likely related to an initially secure cable upon software start up and then a reconnection made during software shut down at the end of the trip. This resulted in multiple trips (actually trip segments) for each reported trip. Only two of the six total trips reported appear to have logged complete GPS data.

4) Participant #27: This participant reported problems with the GPS receiver on last two days of survey. The GPS data indicate that the receiver and/or cabling was not working properly during these trips; the GPS data trip ends for the last two trips detected were located on a freeway and on major arterial, respectively.

5) Participant #29: There are numerous types of problems suspected as a result of this comparison – it appears that the GPS equipment package was not used during some trips and, at other times, the cable may have been loose, causing the detection of many short trips.

6) Participant #30: This participant also reported experiencing some problems with the GPS receiver and the Palm during numerous trips. The breaks in the GPS data
stream indicate that the equipment and/or the cabling were not functioning properly.

Given the extent of the problems found within the data sets collected by these six participants, a decision was made to eliminate these detected trips from further analysis in this particular study. However, exploratory analyses of these data already conducted have identified the symptoms of various data collection problems that can be used for identification of error sources in other studies, as well as provide insight to equipment, software application, and procedural improvements for future studies. These recommendations are included in Chapter 7.

Loose cabling problems were one such category of problems that were easily identifiable when examining the GPS data in a GIS. For example, the maps in Figure 6-4 show two trips in which a segment of each trip was logged for the beginning of the trip, followed by a large gap in the data stream, and finished with a few GPS points logged at the end of the trip. These patterns indicate that the data cabling became loose during the trip – perhaps while the driver was moving the Palm while performing another task inside the vehicle. The last points recorded at the destination were most likely captured when the participant picked up the Palm to close the data logging application – it is probable the participant secured the connection at that time while looking directly at the Palm and the cable. It is also not likely that these large gaps in the data were due to blocked sky views; both routes were on open roadways with multiple lanes and no continuous overhead obstructions. With the exception of urban canyons, satellite line-of-sight problems tend to be both sporadic and short-term as opposed to cabling problems.
Figure 6-4: Examples of Incomplete Trip Data Due to Loose Cabling
Land Use and Address Assignment

Once the trip ends have been identified within the GPS data stream, it is possible to use a digital land use inventory to derive most, if not all, of the individual trip purposes using the raw GPS trip end data points. Conceptually, all that is needed is a geographically-referenced land use database. The land use of the trip ends can be assigned using a simple point-in-polygon analysis that transfers land use descriptions from the polygon-based land use inventory to the point-based GPS trip ends. The trip purpose can then be estimated by evaluating the origin and destination land use pairs.

While many urban regions maintain a good land use inventory, Atlanta is not one of them. The land use inventory for Atlanta is currently under development and contained several hurdles to overcome for this experiment. Primarily, the land use database is actually a series of tax assessor property databases that were designed for estimating property taxes, not for land use/activity determination. Secondly, the database consisted mostly of center points, not polygons. (Since polygons represent property boundaries by definition, these are much more accurate than a database of property center points alone, which require a manual GPS point to property point matching process.) To overcome this hurdle, supplemental information was used to clarify the land use at the trip ends. The supplemental data included 1993 digital orthophoto quarter quadrangles (DOQQs), Fulton County property boundaries, road centerline data, and a Haines name/address database.
Atlanta Land Use Inventory Development

To prepare for the GIS needs of the SMARTRAQ project, researchers at Georgia Tech’s Center for Geographic Information Systems have been developing a land use inventory for the Atlanta metropolitan region. The heart of this inventory is the 1999 county tax assessors’ databases that were acquired from Property Data Systems Inc. (PDS). These databases (by county) are flat files of properties with over one hundred property variables (physical dimensions, structure descriptions, land values, zoning, owner information, land use characteristics, etc.). Geographic property coordinates are not part of the database, but it does contain street addresses and tax map page coordinates (relative to unique tax map pages). Three different techniques were used to generate geographic coordinates for each property. First, records with the best accuracy levels (within 10 meters, for Fulton and Gwinnett Counties only) were linked to polygon-based geographically referenced databases acquired from the counties themselves. Second, tax map pages were geo-referenced using Land Lot maps for Dekalb and Cobb Counties only (within 50 meters). Thirdly, and least accurately (within 100 meters), tax map pages were geo-referenced using the address of the properties. The final product is a database of property polygons, when available, and property center point coordinates when no polygon data was available.

Because the PDS data are actually a property database and not a land use database, the land use classifications are not ideal for this project. Ideally, the land use inventory should be structured like the Land-Based Classification Standards (LBCS) developed by the American Planning Association. While the Atlanta Land Use Inventory will
eventually be structured in this manner, analysis is currently restricted to the codes developed by each county. These codes vary by county and are difficult to unify for a region. In fact, the land use codes are actually zoning codes. However, there is a business code that provides more insight into the actual land use, and, when combined with the land use codes, allows for the accurate assignment of land use in many cases. Mixed land uses are assigned generic business codes such as “strip mall”, “shopping center”, or “office building” and will need to be processed separately within this study.

Finally, in addition to this parcel inventory, the 1999 DOQQs (i.e., photo overlays) and the Georgia DOT roadway centerline database were also acquired. All of these databases were combined into an ArcView (GIS software by ESRI, Inc.) project, allowing them to be displayed simultaneously with any other geographic data (i.e., GPS coordinates). This setup allows an analyst to display, query, zoom, pan, and spatially manipulate any or all of the data at one time.

**Assignment of Land Use and Address**

As a result of the variation in land use data sources and accuracy, as well as the current development stage of the comprehensive parcel database, a mostly manual investigatory process was used to identify the land use and street address from the land use database. The process was as follows:

1. All of the trips end records were combined into a single file that enabled the simultaneous display of all end points in the GIS. A “status” field was added to the composite file, with each trip end given a value of “1” indicating that the record was not yet processed.
2. All of the trip ends that fell within the boundaries of Fulton or Gwinnett County were batch processed using a point-in-polygon analysis technique in ArcInfo. All trip ends that fell within a property’s boundaries were assigned the land use and street address for that property. For these records, the status field was changed from “1” to “2” indicating that the assignment process was completed.

3. All of the trip ends were then displayed in the GIS along with county boundaries and roads. The color for each trip end displayed on the screen represented the value of the status field.

4. A trip end or trip end cluster (i.e., three or more trip ends in the same general location) with a status of “1” would be selected and examined. Note that clusters are most logically associated with common destinations such as work or home.

5. All of the layers in the GIS were then turned on to display their relative position. The investigator would first use the aerial photo (DOQQ) to determine the trip end location and probable land use. The road database could be used to identify the closest intersection.

6. Based on the relative pattern of parcel points, the appropriate parcel was selected, with its land use and address displayed. If the investigator felt that the right parcel had been identified, the land use and address was transferred to the corresponding trip end (or ends) record within the database. The status would then be changed to “2”.

7. If a trip end was processed but the investigation was inconclusive, a code of “3” was assigned to the corresponding record in the database.
8. Steps 4 through 7 were repeated until all trip ends in the database were evaluated.

For most trip ends, the procedure relied primarily on the aerial photography since it is the most accurate spatially and best conveyed whether the trip end was in a residential area, shopping area, etc. The parcel data were visually compared to the aerial photography to identify the best centroid point. Once the parcel was identified, its land use and address were transferred to the trip end database. For cases in which there appeared to be an error in the land use or address information resulting from either equipment failure or missed trips, the GIS was used to allow further investigation using such elements as the actual travel path, the underlying street network, parcel boundaries and land use details, and aerial photography.

Figures 6-5 and 6-6 show the process used. In Figure 6-5, image “a” shows a cluster of trip destination and origin points for Participant #7 that was selected for analysis. The underlying street network is also displayed, along with the corresponding street name. Next, in image “b”, the property center points for that area have been added to the GIS display. Note the offset between the “page” of property points and the underlying street network – it is evident that the points should be shifted upward and slightly to the left. In image “c”, a visual assignment of the property points offset has identified the best likely candidate for the parcel associated with the GPS trip end points. Finally, in image “d”, the aerial photograph is added to provide the complete picture of the streets, houses, and wooded areas. Given six trips originated at this location, and that the land use was identified as “residential, single family,” this location is most likely the participant’s home.
Figure 6-5: Example of Land Use Determination and Address Assignment (Home)
In the images contained in Figure 6-6, both the trip end points and the GPS travel path points are displayed. Image “a” shows the travel route leading to the end of one trip for Participant #28 and the travel route leading away from that particular stop. In image “b”, the underlying street network is displayed. The individual parcel boundaries are included in image “c”. Finally, in image “d”, the property center points and the identification of the trip destination are provided. Since the land use description for this destination is “retail – multi-occupancy”, it is probable that this was a shopping-related trip.

This process was applied to the entire 414 trip ends obtained for the 212 trips detected by the total sample of 19 participants. However, once it was determined that six of the participants had experienced significant equipment problems or had frequently forgotten to turn on the GPS data logger, which resulted in the creation of many false trips and abbreviated trips as well, as the loss of all GPS data for other trips, the 56 trips detected by those participants were removed from further land use, address, and trip purpose analysis. Consequently, the 302 remaining trip ends for 151 trips detected were used for all further analysis (i.e., classes 1, 2, and 3 as defined in Trip Detection section of this chapter). Figure 6-7 shows the locations of the trip ends for these 151 trips within the 13-county Atlanta metropolitan region.
Figure 6-6: Example of Land Use Determination and Address Assignment (Retail)
Figure 6-7: Trip Ends for 151 Trips Remaining in the Analysis
Analysis of Consecutive Land Use and Address Assignment Results

The next step in the analysis process was to examine the trip end data for other problems associated with GPS start up signal acquisition delays, GPS signal blockages on the trip starts and ends, and cabling disconnects. All three data acquisition problems result in the same GPS data stream characteristics – truncated GPS travel routes caused by the lack of data logged at the beginning or end of the trip. These problems are detected by comparing the land use and street address assigned for each trip destination with the land use and street address assigned for the next trip origin. The results of this analysis are found in Table 6-5. A “non-match” is defined as any destination-next origin pair that does not have identical land uses or addresses. The most likely reasons for match failures based upon follow up investigation were:

- Most occurrences of non-matches were attributable to GPS signal acquisition delays during trip starts, but also could possibly have been caused by loose cabling that was reconnected in transit. These trip end records had land use codes of “UNK-on road”, which indicates that the end points do not fall within a parcel boundary but rather somewhere directly on the roadway network itself.

- There were also a few occurrences of non-matches that were most likely attributable to loose cabling in transit. This was somewhat easy to identify when the next trip clearly started at a “known location”, such as at home or at work.

- There were a few trips that were missing both the beginning and ending trip segments. This was most likely the result of cabling problems as well.
Table 6-5: Trip Destination to Next Trip Origin Analysis Results

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Total Detected Trips</th>
<th>Land Use Non-Match</th>
<th>Notes</th>
<th>Address Non-Match</th>
<th>Notes</th>
<th>Survey Start / End Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>11</td>
<td>0</td>
<td>perfect</td>
<td>0</td>
<td>perfect</td>
<td>GT match</td>
</tr>
<tr>
<td>13</td>
<td>11</td>
<td>0</td>
<td>perfect</td>
<td>0</td>
<td>perfect</td>
<td>GT match</td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>0</td>
<td>perfect</td>
<td>0</td>
<td>perfect</td>
<td>GT match</td>
</tr>
<tr>
<td>16</td>
<td>15</td>
<td>1</td>
<td>start acquisition delay trip 8</td>
<td>0</td>
<td>start acquisition delay trip 8</td>
<td>GT match</td>
</tr>
<tr>
<td>19</td>
<td>3</td>
<td>0</td>
<td>note: first trip is a round trip; trip start and end address are identical</td>
<td>0</td>
<td></td>
<td>home match</td>
</tr>
<tr>
<td>21</td>
<td>6</td>
<td>0</td>
<td>perfect</td>
<td>0</td>
<td>perfect</td>
<td>start acquisition delay trip 1</td>
</tr>
<tr>
<td>25</td>
<td>13</td>
<td>2</td>
<td>start acquisition delay trips 2, 10</td>
<td>4</td>
<td>start acquisition delays trips 2, 3, 5, 10 not all detected by land use code</td>
<td>GT match</td>
</tr>
<tr>
<td>9</td>
<td>15</td>
<td>1</td>
<td>start acquisition delay trip 11</td>
<td>0</td>
<td>street numbers not available for most streets -- therefore unable to assess movement</td>
<td>start at GT should end at GT, but early trip end -- possible loose cabling or signal loss</td>
</tr>
<tr>
<td>12</td>
<td>8</td>
<td>0</td>
<td>perfect</td>
<td>8</td>
<td>perfect</td>
<td>GT match</td>
</tr>
<tr>
<td>20</td>
<td>13</td>
<td>0</td>
<td>perfect</td>
<td>0</td>
<td>trip end 8, start 9, end 9, start 10 have same generic address</td>
<td>GT match</td>
</tr>
<tr>
<td>7</td>
<td>19</td>
<td>1</td>
<td>start acquisition delay trip 15</td>
<td>1</td>
<td>start acquisition delay trip 15</td>
<td>GT match</td>
</tr>
<tr>
<td>17</td>
<td>14</td>
<td>3</td>
<td>start acquisition delay trip 14</td>
<td>3</td>
<td>start acquisition delays trips 4, 7, 14, not all detected by land use code</td>
<td>start at home should end at GT, but early trip end -- possible loose cabling or signal loss</td>
</tr>
<tr>
<td>28</td>
<td>21</td>
<td>3</td>
<td>signal acquisition delay trip 9, street/garage for trip 10 dest., possible cabling or start delay on trip 19</td>
<td>5</td>
<td>signal acq delays and equipment problems (cabling or operational) additional signal delays based on street number for same land use</td>
<td>GT match</td>
</tr>
</tbody>
</table>
There was one trip that had the same exact origin and destination address. GIS investigation revealed that this was actually a round trip – composed of two individual trips. This was not obvious in the original, raw trip ends file.

An example of a truncated trip detected by this process can be seen in Figure 6-8, which shows a complete trip (on the left) made by Participant #17 between work and home and a truncated trip (on the right) recorded between work and home for the same participant. In this case loose cabling is most likely the culprit for both abbreviated trip ends since the travel route had a clear sky view from origin to destination (i.e., there were no urban canyons or dense tree canopies).

Figure 6-8: Example of Complete Trip Data (left) vs. Missing Trip End Data (right)
Once all trip end land uses and addresses were obtained and added to the trip ends database, the land use types were combined in a separate file and standardized to common descriptors for the same land use. For example, different counties could describe a single dwelling unit as “residential, one fam”, “single family res”, or “residence, single family”. For this particular case, all codes were translated to “residence, single family”.

**Trip Purpose Derivation**

Once the land uses were standardized and updated in the trip ends file, the next step in the process was to use the derived land use and address information to determine the trip purpose. To support this task, a primary (or default) trip purpose was identified for each land use description, along with secondary and tertiary trip purposes, if applicable, using the trip purposes from the 1990 Atlanta regional household survey (as presented earlier in Table 4-5). Table 6-6 lists the standardized land use codes and corresponding trip purposes as defined in this process. Note that only those land uses that were actually obtained in this experiment are included (i.e., there are hundreds of other land use codes contained with the overall Atlanta land use inventory). The alternate trip purposes are necessary to accommodate both time of day and duration of stay issues. For example, if a participant went to the airport for 30 minutes and then traveled somewhere else in the same vehicle, the trip purpose would be “drop off / pick up.” However, if they arrived at 8:00 a.m. and departed at 5 p.m., then this would be considered “go to work”. Of course, this could have been a day trip to another location via air travel, which should be coded
Table 6-6: List of Standardized Land Uses and Corresponding Trip Purposes

<table>
<thead>
<tr>
<th>Land Use Description</th>
<th>Purpose 1</th>
<th>Purpose 2</th>
<th>Purpose 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airport</td>
<td>drop off / pick up personal business</td>
<td>change mode</td>
<td>go to work</td>
</tr>
<tr>
<td>Bank</td>
<td>go to work</td>
<td>go to school</td>
<td>drop off / pick up</td>
</tr>
<tr>
<td>College or University</td>
<td>conv store / gas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commercial</td>
<td>drop off / pick up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Convenience Store / Gas</td>
<td>personal business</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School</td>
<td>return home</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improved Gov't Exempt</td>
<td>personal business</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residence, Multi-family</td>
<td>drop off / pick up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Office Building</td>
<td>personal business</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shopping Mall (deck)</td>
<td>return home</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post Office</td>
<td>personal business</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Religious, Church</td>
<td>social / recreation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residence, One family</td>
<td>return home</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restaurant</td>
<td>eat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restaurant - Fast Food</td>
<td>eat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retail, Multiple Occupancy</td>
<td>shop</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retail, Single Occupancy</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shopping Center</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strip Mall</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supermarket</td>
<td>shop</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNK - on road</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vacant Exempt Land</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vacant Land / RailRoad</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vacant Lot, Commercial Warehouse</td>
<td>99</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
as “change mode.” However, the exact physical location of the parked vehicle during the day can provide the necessary supporting detail to differentiate between a remote, employee parking lot and a daily traveler parking lot.

One purpose of this experiment was to determine exactly what types of trips can be properly identified. Examination of the land use codes revealed that some codes cannot be given an automatic trip purpose. This is especially true for all mixed use land parcels, such as shopping centers and strip malls, in which there are a variety of businesses serving a range of purposes, including shopping, eating, personal business, and refueling. In addition, there were some ambiguous land use codes, such as vacant lot, improved government exempt, and warehouse, which could indicate a property that has either been miscoded or on which new development has recently occurred. Finally, there were some trips which had a destination land use code “UNK – on road” (for unknown), which indicates that the final GPS data point was located on the road network and not within property boundaries. For these trips, it is likely that either the GPS data stream terminated prematurely due to cabling problems or that the vehicle parked on the street. A code of “99” was assigned to all land uses that were mixed use in nature, and the trips associated with these codes will be clarified during the CATI household data retrieval call.

Finally, two addition trip purposes were added to provide better detail with respect to specific land uses. These trip purposes are “eat” and “convenience store / gas”, which would typically be coded as “other”. In fact, the “other” trip purpose code was not used in this experiment due to the need to identify follow up actions. It is quite possible that
for some land uses not seen in this study, the “other” purpose would be sufficient and correct in an automated assignment process.

Trip Purpose Derivation Process

After the relationships between land use and trip purpose were defined, the trips for each participant were then processed to assign the appropriate trip purpose based on each trip’s destination land use. The previous trip’s purpose was also examined as a logic check to this assignment process. In addition, the arrival time of day at the destination and the activity duration at the destination were also factored into this assignment. The next step in furthering this research is to develop algorithms to automate the trip purpose assignment.

As proposed earlier in this dissertation, advanced knowledge of home, work, school, and other commonly visited destination addresses will greatly assist this entire process. However, this study intentionally omitted the use of this information so that worst-case analyses could be performed. The only exception to this omission was the assignment of work location since the recruitment for the study occurred at Georgia Tech among Civil Engineering faculty, staff, and associates. Although it was possible that some participants could have been students and not employed at Georgia Tech, it could not be easily determined from the GPS data if they were students or employees. Accordingly, all trips destined to Georgia Tech that were more than 30 minutes in duration were identified as work trips. If the person stopped at Georgia Tech for less than 30 minutes, then the trip was considered a drop off / pick up. Although home address information was not accessed during the land use, address, and trip purpose
assignment process, it was fairly easy to identify home locations by the cluster of GPS
trip end points at one specific residential land use location. Consequently, this process
was used to determine for each participant’s home address. This process could also be
applied to determine work and school locations.

**Trip Purpose Derivation Results**

Table 6-7 contains a summary of the trip purpose derivation analysis for each
participant, along with notes on code “99” and misidentified trip purposes. This table lists
the total trips detected, along with the total “go to work” trips, “return home” trips, code
“99” trips, and incorrectly assigned trip purposes. “Go to work” and “go home” trips
accounted for approximately 54% of the total trips made. Of the 39 trips (26%) coded at
“99”, 26 were trips made to mixed use shopping centers, strip malls, or commercial
zoning; 6 were trips made to ambiguous land uses such as vacant lot, government
exempt, or warehouse; and 6 were trips that terminated prematurely due to equipment
failure, which were coded with an “unknown – on road” land use. The removal of the
equipment failure error source brings the total percentage of follow up calls required
down to 21.9. The remaining trips (i.e., those with mixed use or ambiguous land use
codes) are the type of trips that will require clarification of trip purpose during the CATI
follow up call. As long as the percentage of total trips detected that require this follow up
is minimal, the approach presented in this dissertation is beneficial in terms of improved
travel behavior data quality while maintaining or possibly reducing survey administration
costs.
Table 6-7: Summary of Trip Purpose Derivation Analysis

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Total Trips Detected</th>
<th>Purpose: Go to Work</th>
<th>Purpose: Go Home</th>
<th>Purpose: 99</th>
<th>Wrong Purpose</th>
<th>Notes on Trip Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>10</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>both errors are for the same location; Walmart coded as an office building</td>
</tr>
<tr>
<td>13</td>
<td>11</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>15</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>gas station coded as shopping ctr</td>
</tr>
<tr>
<td>19</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>missed trip -- round trip to store</td>
</tr>
<tr>
<td>21</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>13</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>post office coded as retail, single occupancy</td>
</tr>
<tr>
<td>9</td>
<td>15</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>post office coded as shopping center Publix coded as conv store last trip terminated on roadway</td>
</tr>
<tr>
<td>12</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>both 99's shopping centers</td>
</tr>
<tr>
<td>20</td>
<td>13</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>0</td>
<td>4 shopping centers day care center coded as vacant land</td>
</tr>
<tr>
<td>7</td>
<td>19</td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>residential code instead of store GPS equipment / position error</td>
</tr>
<tr>
<td>17</td>
<td>14</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>two 99's coded as UNK - on road indicating cabling problem parking garage coded as restaurant Publix coded as restaurant Three 99's coded as UNK - on road; 2 are MARTA drop offs, the other most likely a traffic delay</td>
</tr>
<tr>
<td>28</td>
<td>21</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>151</td>
<td>31</td>
<td>50</td>
<td>39</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

20.53% 33.11% 25.83% 6.62%
One significant issue found in these results is the ten incorrect trip purposes, which were misidentified as a result of inaccurate land use assignment. These land use assignment errors can be attributed to GPS position errors (e.g., uncorrected GPS data or premature termination of data stream), inaccurate parcel boundaries in GIS database, inaccurate assignment of parcel to the GPS trip end, or inaccurate coding of land use in the parcel database. If these error sources are not eliminated, inaccurate land use assignments will result in undetectable trip purpose errors in the final trip data derivation and thus threatens the accuracy of the results. However, there are solutions that can be implemented to address each of these error sources; these will be presented in the recommendations contained in Chapter 9.

**Trip Distance**

Although trip distance was not a trip characteristic collected during the 1990 Atlanta survey, the element was included on the memory jogger given to the participants since it was known that algorithms already exist which automate the map matching process of “converting” GPS data points into road network links within a GIS. However, due to incompatibility problems between some of these readily available map matching applications and the TransCAD GIS package, new map matching algorithms were deemed necessary.

Two general approaches for map matching algorithm development were defined and investigated. The first approach is to use buffer analyses around the GPS points to identify links in the road network that fall with this buffer. For example, if the buffer
distance is set to 25 meters, then the algorithm would identify any roadway links that completely fall within 25 meters to either side of the travel path as a link on the actual travel route. The time stamps within the GPS data can be used to indicate travel direction across and through the links. Once the sequence of links has been identified for a given trip, then another algorithm can be run to sum the distances of each link to derive the total trip distance. The other approach is to identify all turns within the GPS data based on a combination of low speed values and vehicle heading changes. Once the sequence of turn coordinates (GPS point data) is obtained, then another algorithm can be applied which searches through the roadway database to find links with ends within a predefined distance from the GPS turn points. The actual travel path can also be derived.

Even though both approaches are quite feasible and it is very likely that map matching and travel route and distance algorithms will be developed in ArcView at Georgia Tech within the next six to nine months, the effort to develop these algorithms in TransCAD proved to be more difficult than anticipated. Given that several of the other GPS data collection projects presented previously in this dissertation have already proven that GPS data can be converted into trip routes and distances [Wagner 1997, Doherty 1999], it was not necessary to fully develop these processes and algorithms as part of this experiment. Therefore, a distance estimation procedure was developed using the instantaneous speeds in the GPS data stream to provide approximate distances for each trip detected.

This procedure used the speed data contained within the GPS data stream to derive travel distances. The GPS receiver used in the P1 and P3 equipment package provides
GPS records for every second of the trip. These records include vehicle speed, which is easily converted into units of miles per second (i.e., distance traveled per second). The sum of all second-by-second distances is obtained to approximate each trip’s travel distance. (The code for this macro can be found in Appendix E.) The Garmin II Plus GPS receiver used in equipment package P4 typically provides position data every two seconds; however, in this study that interval sometimes varied up to 8 seconds. To estimate distances for these data sets, the speeds of consecutive GPS points were averaged and applied as the individual second speeds for all seconds between the two points. Then the second-by-second distances were derived and totaled, as defined above, to get the total trip distances.

To verify that this approach provides a fair approximation of the actual travel distance, random trips were selected for evaluation using TransCAD. A few trip distances were validated by summing the lengths of all links along the GPS point path that were contained in the Georgia DOT street network database. Other trip distances were validated by using the Ruler function with TransCAD to click along the travel path to obtain the sum of a series of straight-line distances. The results of all three methods were within one mile of each other. Consequently, the estimated trip distances calculated as the sum of the second-by-second distances were used as the GPS-derived distance to complete the travel diary comparisons. These distances can be found in the individual trip diary comparisons presented in Chapter 7.
CHAPTER VII

FINDINGS

The findings of this pilot study are provided in two sections. First, the diary comparisons between reported and detected trips for each individual participant are presented. Then, the findings for each participant are aggregated and presented as part of the overall study findings. Next, key issues to be addressed in future related studies are reviewed. Finally, areas of opportunities based on the study findings are provided.

Comparisons by Participant

The trip diary logs containing the comparisons of reported and derived travel data for all 24 survey participants can be found in Appendix F. Figure 7-1 shows the comparison log for participant #12. The information contained in this log includes:

1) The heading (top four rows) contains the participant’s survey descriptions (study ID, start date, starting location, and home address (which was added after the trip derivation process), along with the GPS derived starting location and GPS data filename.

The body of the log contains the trips reported and derived. The details of each trip reported by the participant are contained on the lines shaded in gray (and tagged with a “P” code for “paper.” If the participant did not record anything in a
particular field, then that field contains a question mark “?”. Following each trip reported is the trip information derived from the GPS data as part of this pilot study (tagged with a “G” code for “GPS”). The trip numbers represent the numbers as reported by the participant.

3) Notes are provided at the bottom with some descriptive information about any relevant information with respect to the paper diaries, the GPS data processing, and the trip comparison analysis.

<table>
<thead>
<tr>
<th>Trip #</th>
<th>Day of Week</th>
<th>Trip Start Time</th>
<th>Trip Stop Time</th>
<th>Destination / Place Name</th>
<th>Address / Street &amp; Cross Street</th>
<th>Activity at Destination</th>
<th>Travel Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P</td>
<td>Tues 16:30</td>
<td>17:10</td>
<td>Best Buy</td>
<td>2460 Cobb Pkwy, Smyrna</td>
<td>Shopping</td>
<td>30.0</td>
</tr>
<tr>
<td>1</td>
<td>G</td>
<td>Tues 16:28</td>
<td>16:55</td>
<td>Shopping Center</td>
<td>2400 Cobb Parkway</td>
<td>99</td>
<td>13.6</td>
</tr>
<tr>
<td>2</td>
<td>P</td>
<td>Tues 17:30</td>
<td>18:00</td>
<td>Home (Walton Pk)</td>
<td>1218 Walton Ln, Smyrna</td>
<td>Home Activities</td>
<td>20.0</td>
</tr>
<tr>
<td>2</td>
<td>G</td>
<td>Tues 17:05</td>
<td>17:22</td>
<td>Residence, Multi-Family</td>
<td>Mill Pond Rd.</td>
<td>Return Home</td>
<td>5.3</td>
</tr>
<tr>
<td>3</td>
<td>P</td>
<td>Wed 7:15</td>
<td>7:40</td>
<td>GA Tech</td>
<td>A11 parking zone</td>
<td>Work</td>
<td>20.0</td>
</tr>
<tr>
<td>3</td>
<td>G</td>
<td>Wed 7:08</td>
<td>6:34</td>
<td>College or University</td>
<td>Georgia Tech Campus</td>
<td>Go to Work</td>
<td>11.5</td>
</tr>
<tr>
<td>4</td>
<td>P</td>
<td>Wed 16:45</td>
<td>17:30</td>
<td>Home</td>
<td>1218 Walton Ln, Smyrna</td>
<td>Home Stuff</td>
<td>20.0</td>
</tr>
<tr>
<td>4</td>
<td>G</td>
<td>Wed 16:23</td>
<td>16:26</td>
<td>Shopping Center</td>
<td>10th st / Hemphill</td>
<td>99</td>
<td>0.7</td>
</tr>
<tr>
<td>Missed</td>
<td>G</td>
<td>Wed 16:45</td>
<td>17:09</td>
<td>Residence, Multi-Family</td>
<td>Mill Pond Rd.</td>
<td>Return Home</td>
<td>13.2</td>
</tr>
<tr>
<td>5</td>
<td>P</td>
<td>Thu 7:15</td>
<td>7:42</td>
<td>GA Tech</td>
<td>A11 parking zone</td>
<td>Work</td>
<td>20.0</td>
</tr>
<tr>
<td>5</td>
<td>G</td>
<td>Thu 7:11</td>
<td>7:34</td>
<td>College or University</td>
<td>Georgia Tech Campus</td>
<td>Go to Work</td>
<td>11.8</td>
</tr>
<tr>
<td>6</td>
<td>P</td>
<td>Thu 16:45</td>
<td>17:10</td>
<td>Home</td>
<td>1218 Walton Ln, Smyrna</td>
<td>Home</td>
<td>20.0</td>
</tr>
<tr>
<td>6</td>
<td>G</td>
<td>Thu 18:40</td>
<td>19:02</td>
<td>Residence, Multi-Family</td>
<td>Mill Pond Rd.</td>
<td>Return Home</td>
<td>12.4</td>
</tr>
<tr>
<td>7</td>
<td>P</td>
<td>Fri 7:00</td>
<td>7:20</td>
<td>GA Tech</td>
<td>A11 parking zone</td>
<td>Work</td>
<td>20.0</td>
</tr>
<tr>
<td>7</td>
<td>G</td>
<td>Fri 7:05</td>
<td>7:26</td>
<td>College or University</td>
<td>Georgia Tech Campus</td>
<td>Go to Work</td>
<td>11.6</td>
</tr>
</tbody>
</table>

P: Paper Notes: Participant did not report one stop made during trip 4, apparently to shop or eat.
G: GPS-derived GPS times have been adjusted +1 hour for daylight savings change on Apr 2 Participant provided some times and all distances with @ sign (estimates?)

Figure 7-1: Comparison Log for Participant #12
As shown in Figure 7-2, the derived data elements come from Step 1: Trip Detection Macro (Trip Day of Week, Start Time, and Finish Time); Step 2: Land Use and Address Assignment (Destination Place Name and Address / Street & Cross Street); Step 3: Trip Purpose Derivation (Activity at Destination); and Step 4: Trip Distance (Travel Distance). The derived travel dates and times have been converted to the same units used by the participants to allow for ease in analysis.

<table>
<thead>
<tr>
<th>Trip #</th>
<th>S</th>
<th>Day of Week</th>
<th>Trip Start Time</th>
<th>Trip Stop Time</th>
<th>Destination / Place Name</th>
<th>Address / Street &amp; Cross Street</th>
<th>Activity at Destination</th>
<th>Travel Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>P</td>
<td>Fri</td>
<td>22:15</td>
<td>22:30</td>
<td>home</td>
<td>3433 Casa Woods Ln</td>
<td>home</td>
<td>5.0</td>
</tr>
</tbody>
</table>

Figure 7-2: Derived Trip Elements

If the GPS process detected an additional trip not reported, it appears as two lines after the reported trip, with the addition trip labeled as “Missed” in the Trip # column (as seen in Figure 7-1 for Trip 4). If the participant reported a trip not detected by the trip detection macro, it will appear as seen in the example contained in Figure 7-3. In this figure, the trip reported by the participant was a drop off item at Blockbuster, but the single detected GPS trip for reported trips 20 and 21 starts at the end of trip 19 (at home) with a 10:01 start time and arrives at trip 21’s destination at 10:20.
Overall Study Findings

The proposed process to replace trip diaries in household travel surveys with GPS data loggers is feasible. This pilot study revealed great promise regarding the ability to accurately identify trip ends within a GPS data stream, to perform land use and address look ups with a spatially accurate and comprehensive GIS, and to assign individual trip purposes from the previously derived data. In addition, this study also served to identify the symptoms of data problems related to equipment malfunction, receiver acquisition delays, and manual trip reporting errors. Table 7-1 provides a data summary for the study. The key findings within each step of the trip details derivation process examined in this research effort follow.

Trip Detection Module

- 120 seconds is a reasonable definition for stop detection within the GPS data stream. Lowering the threshold to 90 and 60 seconds picked up additional non-stop delays resulting from traffic signals and congestion. Increasing the threshold would only serve to miss more stops of short duration.

### Table 7-1

<table>
<thead>
<tr>
<th>Trip #</th>
<th>Day of Week</th>
<th>Trip Start Time</th>
<th>Trip Stop Time</th>
<th>Destination / Place Name</th>
<th>Address / Street &amp; Cross Street</th>
<th>Activity at Destination</th>
<th>Travel Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>F</td>
<td>Sun 11:32</td>
<td>11:39</td>
<td>Home</td>
<td>1138 North Highland</td>
<td>home</td>
<td>1.7</td>
</tr>
<tr>
<td>19</td>
<td>G</td>
<td>Sun 11:34</td>
<td>11:40</td>
<td>Residence, Single Family</td>
<td>1138 N. Highland Ave</td>
<td>return home</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>F</td>
<td>Mon 9:59</td>
<td>10:01</td>
<td>Blockbuster Monroe / Piedmont</td>
<td>1138 North Highland</td>
<td>drop off video - errand</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>G</td>
<td>Mon 10:19</td>
<td>10:20</td>
<td>College or University</td>
<td>750 Atlantic Drive</td>
<td>go to work</td>
<td></td>
</tr>
</tbody>
</table>

Figure 7-3: Example of Trip Reported but Not Detected
There were significant problems with the data cable connection between the Palm data logger and the GPS receiver, resulting in the omission of entire trips and trip segments within the GPS data stream.

Overall, this module performed well for the 13 respondents who had complete data sets: for the 156 total trips reported by these participants, 146 were detected successfully (94%), along with the correct identification of an additional four unreported trips. Two other stops were detected that are most likely attributable to extended traffic signal delays.

This module did not detect 10 passenger pick ups or drop offs – all of which had a duration of less than two minutes. However, much was learned about the GPS data...
characteristics of passenger pick up and drop off trips. Visual inspection of the trip
data may be needed to detect these trips.

**Land Use and Address Assignment**

- An automated process to assign land use and address for each trip end was used for Fulton County trip ends. Manual procedures were developed to make the assignments for the remaining counties; these procedures included visual inspection of the GPS trip ends with a GIS, with the underlying road network, parcel, and integrated aerial photograph databases used to make the final assignments.

- Overall, this process worked well. Some trip ends were truncated due to GPS signal acquisition delays or cabling failures, which resulted in incomplete GPS trips, with either the beginning or ending of the actual trip missing from the file. Most of these types of errors were identified within the GIS as an unknown land use with the description “UNK – on road.”

- GPS startup signal acquisition delays can often be overcome by using the information from the last trip’s destination. However, some caution should be exercised with respect to automatically making this assignment; a visual inspection of sequential routes in recommended.

- The current parcel-level database for Atlanta is still under development. Also, due to a variety of data sources and formats, there are different levels of accuracy for different counties in the region. Metropolitan planning organizations in other locations will need accurate parcel-level databases in order to use this same procedure.
Trip Purpose Derivation

- A new trip purpose was introduced to accommodate land uses with the potential for multiple uses. A code of “99” was given to all land uses falling into this category, including shopping centers, strip malls, commercial, and retail – multiple occupants. Trips that are assigned this code will require follow up clarification during the CATI household data retrieval call.

- Of the 151 trips detected by the 13 participants, 50 had a trip purpose of “return home” (33.1%), 31 had a purpose of “go to work” (20.5%), and 39 had a code “99” trip purpose (25.8%) as the result of either a mixed-use destination land use code (26 of 39), an ambiguous land use code such as vacant lot or government exempt (6 of 39), or the premature termination of the GPS data stream due to equipment problems (6 of 39). Hence, only 26% of the trips need to be queried during CATI retrieval, thus reducing CATI burden.

- After removing the 6 equipment-related errors results, only 33 out of 151 trips required CATI follow up questions for trip purpose clarification – which is only 22% of the total trips made. As long as the percentage of total trips detected that require this follow up remains low, it is quite feasible that the approach presented in this dissertation is beneficial in terms of improved travel behavior data quality while maintaining or possibly reducing survey administration costs.

- Ten of the 151 trips were assigned incorrect trip purposes based upon the assignment of an incorrect land use (6.6%). This problem could result from GPS position inaccuracies or missing trip end GPS data which places the trip end in the wrong
parcel, from parcel boundaries that are inaccurate, or from incorrect coding of land use within the parcel database. This type of error is of great concern since it results in undetectable errors that will impact travel demand inputs if left uncorrected. Further research is needed to explore the distribution of these errors across different land uses / trip purposes.

The manual processes used for land use and trip purpose derivations do not detract from the success or failure of this experiment in that, once the land use database is fully developed, much of the process can and will be automated. Of course, there will always be a need for manual (and visual) data examination and exception handling procedures. Regardless of the means, the fact remains that land uses, addresses, and the resulting trip purposes were identified from a GIS database without knowledge of the trip diaries. Even with the issues surrounding the Atlanta land use data, 93.4% of the trips detected were assigned the correct land use.

The benefits gained by using a GIS in combination with the GPS data were numerous. The GIS enabled the collection of trip-related data elements, including trip end land use, trip end street address or nearest cross road, property ownership, the actual travel route, and trip length. Using a GIS allowed the collection of these elements to occur through a mix of automated and manual methods, including visual inspection. Finally, the GIS environment allowed for a visual validation of all results.

There were a few concessions made throughout the trip derivation analysis process, including: 1) the scope of the study was limited to in-vehicle trips only; 2) travel mode details such as driver identification and the number of passengers are not obtainable from
the GPS data; and 3) some trip purposes will have to be collected by CATI due to mixed-use land use codes, ambiguous land use codes, and common parking areas. In fact, there may always be a need for certain follow up questions regarding the derived travel data to be made during the CATI household data retrieval call. Certainly, if driver identification and/or the number of passengers in the vehicle for each trip are required data elements, then they will need to be collected during the CATI call. As an alternative, it is possible that these elements could be reasonably estimated by asking relevant questions during the household recruitment call regarding carpooling-type behaviors. It is also possible that passenger pick up and drop off trip purposes could be used to increment or decrement vehicle occupancy. In addition, due to mixed or ambiguous land uses present in a region’s land use inventory, there will be a need for clarifying questions during the CATI household data retrieval call to obtain the exact land use or trip purpose. Finally, destinations for which the driver parked in common parking areas such as public parking lots or on-street parking will also require further destination information to be obtained from the survey participant. At the very least, however, the CATI process should be shortened significantly once this automated trip derivation process is successfully implemented.

**Key Issues**

There were two key areas for concern discovered in this analysis. First, the equipment packages assembled for GPS data acquisition proved to have many more problems than anticipated. Of the 24 participants who were given both the GPS data
logger and the paper diary, 5 were eliminated during the data collection period due to Palm-related power problems or application errors and another 6 were dropped later during the trip detection step when numerous data errors were discovered indicating problems with equipment or operation. The second problem area was data accuracy; GPS and/or GIS data errors caused the assignment of incorrect land use codes, which resulted in the incorrect trip purpose derivation for the corresponding trips.

Data Logger, the GPS data logging software developed for this project, had a few faults that contributed to the loss of five participants during the data collection phase. The application was designed to allow users to power off the Palm device from within the Data Logger, to power off automatically if the user indicated that the trip ended (by tapping the Stop Logging button on the GPS data logging screen), and to prevent users from exiting the Data Logger application throughout the survey period. However, there were problems experienced with these design objectives. In the case of four participants, the Palm device did not power off, indicating either a conflict between Data Logger and the Palm Operating System, which prevented the user from powering down the unit, or the omission of a trip end entry by the participant. The fifth participant who was dropped from the study during the data collection phase inadvertently exited Data Logger by tapping on one of the Palm icons at the bottom of the screen. Since this should not have happened and since the participants were not trained on Palm operating system usage, the participant did not know how to re-enter the data logging application.

Data Logger contained several user-defined parameters intended to offer flexibility with respect to vehicle non-movement speed thresholds, non-movement
countdown periods, and universal date and time offsets. This last parameter proved to be more troublesome than expected; during the first two weeks of the study, the offset between Eastern Standard Time (EST) and UTC was –5. However, the U.S. eastern time zone switched from EST to EDT (Eastern Daylight Time) at midnight on April 1 and the units were out in the field at the time. When the equipment was collected on April 3 and redeployed on April 4, the parameter offset was not corrected. This caused a one-hour error in the time stamp for all GPS data logged during the first week of April. These GPS times were adjusted on the diary comparisons for the affected participants (as noted on the bottom of the diaries).

The GPS data logging equipment packages were not durable, causing the elimination of 6 additional participants from the study. As mentioned in Chapter 5, the packages assembled for this analysis were not designed for field deployment – the components used were GPS receivers and Palm devices that were available from other research initiatives. Because the packages were small and seemingly simple, problems such as Palm power drainage and chronic cabling disconnects were not expected. However, it was known prior to the data collection effort for this study that a more durable and user-friendly in-vehicle GPS data logger would be developed before the commencement of a larger scale study. Participant concerns about GPS antenna cable placement through open windows were discovered. GPS start up signal acquisition delays were experienced and proved that prior assumptions are valid regarding destination GPS points as more reliable stop location data.
Finally, data completeness and quality problems experienced with both the GPS and GIS data posed significant challenges to trip purpose identification. The trip detection step produced a few false trip ends that were most likely traffic signal delays and missed numerous trip ends associated with short duration drop-offs or pick-ups. Loss of GPS trip destination data due to loose cabling or signal loss resulted in false trip destination location data. Position errors associated with Selective Availability also introduced destination location offsets of 30 to 100 meters from the true location for equipment packages P3 and P4. Finally, there were errors in the parcel data, including missing or incorrect land use codes.

**Key Opportunity: Missed Trip Identification**

Missed trips are a serious problem with traditional travel surveys. Underreporting of trips at the household level can cause significant under-predictions in regional travel demand at the aggregate level. Consequently, much attention has been given lately to improving travel data capture methods. Research indicates that the most common types of trips omitted from travel diaries are those trips imbedded in a trip chain, including passenger or item pick ups or drop offs that are incidental to the traveler’s ultimate trip purpose, and short round trips that are considered by the respondent to be unimportant or non-trip events. GPS data capture and travel behavior derivation from the GPS data can offer assistance in detecting typically omitted trips.

The trip detection macro, as presented earlier in this dissertation, is based on the detection of vehicle non-movement for a predefined period. This study experimented
with various levels of this threshold and found that 120 seconds seemed to be a reasonable threshold for trip detection; thresholds set above this level would result in fewer true trips being detected, whereas thresholds set below 120 seconds start picking up false stops associated with congestion and traffic signal delays. Figure 6-2 (in Chapter 6) shows the GPS data associated with survey participants #20 and #12, who both omitted one trip from the travel diary that was detected by the GPS trip detection macro. These maps show how simple it is to detect stops that are greater than the threshold set within the software. Of course, the detection of false stops is also an issue, and is a primary reason why all questionable stops identified should be confirmed via the CATI process proposed in this study.

**Trip Chains**

Trip chains are a sequence of trips “chained” together, for which the traveler often thinks in terms of the final destination trip purpose or an overall trip purpose descriptor. An example of a final destination trip purpose is “return home” for all trips occurring during the work to home journey, whereas an overall trip purpose descriptor example is “running errands” while conducting a sequence of trips that originate and terminate at the traveler’s home. Because of this thought / perception process, the travel survey participant often forgets these incidental trips or decides that the trips are not really trips at all. A good example of this can be seen in Figure 7-4, which contains a segment of the diary and GPS data collected by Participant #18. Note that the data collected and derived for this participant are not included in the study results; this participant was eliminated from further analysis after the trip detection process revealed numerous data quality
problems. Nonetheless, the data did contain a perfect example of a trip chain containing four intermediate stops not reported and, therefore, has been included here.

This participant was located at Georgia Tech at the start of trip #2 and reported that the next destination was home (even though an address near Georgia Tech was recorded incorrectly). However, the GPS data tells a different story and is confirmed within the GIS (see Figure 7-5). The first stop made by the participant occurred on the southern perimeter of Georgia Tech’s campus at a bank. Next, the participant traveled to a convenience store, stopped for five minutes, continued traveling to a shopping center, stopped for 10 minutes, and then continued to a fast food restaurant. After making these four intermediate stops, the participant then returned home. The sum of the five trip links for this work to home journey was 12.3 miles, which was very close to the 12 mile distance reported by the participant, who reported distances in integer values only.

![Table of Trip Details](image)

**Figure 7-4: Example of Four Missed Trips (Origin at Georgia Tech)**
Figure 7-5: Maps of Four Missed Trip Destinations
**Passenger Pick Up or Drop Off**

Passenger (or item) pick up or drop off presents a special problem given that the associated stops are usually of very short duration. Although it is difficult to identify these short duration drop offs and pick ups using the Trip Detection Macro, there are other automated and visual methods for identifying these trips within the GPS data. For example, a trip that started at a participant’s home, went to the airport to pick someone up at baggage claim (i.e., a very short stop), and then continued to a hotel, will appear in the trip ends file as one trip made from home to a hotel. However, visual inspection of this “trip”, as seen in a GIS (see Figure 7-6), quickly reveals that the trip was not just one trip. Another method for identifying this type of missed GPS trip (i.e., those with short stop times) could be the comparison of trip distance as calculated from the GPS data with a shortest path / cost distance calculated in a GIS between the start and end points. Of course, the distance difference (or ratio) threshold to trigger this identification must be set high enough to account for inefficient route choices made by many travelers every day. Graphical methods which identify some deviation from the shortest path or which find trip “legs” can also be used to locate potentially missed trips.
Figure 7-6: Examples of Trip Chains with Imbedded Passenger Pick Ups

Round Trips

For round trips that do not stop at the destination long enough to trigger trip detection, there are several methods for identification. First, it is very likely that the start and end address for these “out-and-back” trips will be identical. If this is the case, an automated method that looks for this condition could flag these trips for further investigation. This was the case for a trip captured for Participant #19. After detecting that the trip start and end addresses were the same, the GPS data for the trip were examined within a GIS. As seen in Figure 7-7, it is evident that the participant drove to a
strip mall, circled the parking lot a few times, and then returned home. Follow up questions with the participant revealed that, while one person exited the vehicle and went into the grocery store to make a quick purchase, the driver continued to circle the parking lot to keep the baby in the back seat from awakening. The round trip start and end location are marked with the star.

Figure 7-7: Example of Round Trip (Two Trips) Without Vehicle Stop

Short Distance Trips

It is easy to identify short distance trips using GPS data processing techniques as long as there is a meaningful time gap at the stop. This was evident in the data collected for Participant #28; after a trip made to a fast-food drive-through window, the participant
then drove to a parking space located on the same parcel (see Figure 7-8 below). This short trip was not recorded on the paper diary but was picked up within the GPS data. Note that the drift in location between the destination of trip #13 and the start of trip #14 is due to the pseudorandom error introduced by Selective Availability. This type of drift will no longer be an issue since the removal of SA on May 1, 2000.

Figure 7-8: Example of Short Trip Detected Not Reported
CHAPTER VII

CONTRIBUTIONS

Research to date conducted on the use of GPS technology in travel survey data collection have examined GPS data as a supplement to traditional or electronic travel diary methods; this research presents a process which enables the replacement of the paper or electronic diary. Consequently, this research is the first study to examine the total elimination of the travel diary component of a household travel survey by using GPS data collected during travel to derive the essential travel diary data elements. The findings of this study demonstrate that it is feasible to consider this new methodology as an alternative to traditional travel data collect methods. Although other GPS-related travel survey studies have also examined the ability to identify trip ends, travel routes, and travel distances from the GPS data stream, the research presented in this dissertation is the first to date that has examined the feasibility of deriving trip purpose from the GPS travel data. The derivation of trip purpose for each trip recorded with GPS technology is the single most significant contribution made by this research effort.

When implemented, the use of GPS data collection in place of travel diaries will introduce a completely new method for travel data collection that offers important benefits as opposed to traditional methods, including 1) the reduction of respondent burden by eliminating the travel diary instrument and by shortening the telephone interview length; 2) a related reduction in telephone interview costs; 3) the extension of
survey periods to multi-day, multi-week, or multi-year periods; 4) improved accuracy and completeness of existing inputs to travel demand models (specifically for trip generation); and 5) the collection of new travel data elements (routes and speeds) that enable trip assignment model validation, calibration, and/or update.

Beyond travel demand modeling, new travel data elements such as travel speeds, travel routes, and parking durations that are now available as a result of GPS travel data collection present great opportunities for improving air quality models. Specifically, GPS data provides detailed second-by-second speed and acceleration profiles across the roadway network, data that are greatly needed for current air quality models. In addition, the parked vehicle information that can be gleaned from the GPS data, including parked vehicle durations with spatial distributions, provides the details necessary for better modeling evaporative emissions, cold and hot start emissions, and multi-day diurnal emissions.
CHAPTER VIII

CONCLUSIONS AND RECOMMENDATIONS

The purpose of this dissertation research was to investigate the feasibility of replacing the traditional travel diary used in the household travel survey with a GPS data logger. The goal was to transform the second-by-second GPS data collected for all in-vehicle trips into the essential trip-level details typically recorded on a paper or electronic diary and then reported during the CATI data retrieval call. Four data processing steps were defined and implemented to transform the GPS data into trip details. Travel diary and GPS data collected by 19 study participants were processed and analyzed.

This experiment was conducted with several “worst case” elements, including the degradation of GPS position accuracy caused by Selective Availability, the lack of data cleaning after data collection, and the over-simplification of trip detection logic (based on vehicle non-movement time gaps only). Even with these inaccuracies, the analysis performed for this dissertation demonstrated that it is possible to replace travel diaries with GPS data loggers. In fact, the process will work quite well once certain GPS data logging equipment and GIS database problems are addressed; recommendations to reduce these problem areas are included later in this chapter. Also, the removal of Selective Availability on May 1, 2000, an event that occurred after the data for this study were collected, will have a significant, positive impact on the performance of this process. GPS point accuracy improved from a range of 30-100 meters to 3-20 meters overnight.
Conclusions

The results were very promising; after removing 6 participants’ data sets due to faulty equipment or operating problems, the remaining 13 participants reported 156 trips made during the 2.5-day survey period. The Trip Detection Macro identified 151 trips, four of which were unreported. Ten short duration passenger pick-up or drop-off trips were undetected, but much was learned about the data characteristics of these trips. Procedures can be implemented to assist with the detection of these types of missed trips.

Next, the Land Use and Address Assignment step, performed with a spatially accurate GIS that included a land use inventory, was also very successful. Land uses and addresses were found for 145 trip destinations; the other six trips terminated on the road network and could not be assigned a land use. These six unknown destinations were either the result of a premature termination of the GPS data stream due to loose cabling connections or to on-street parking.

The Trip Purpose Derivation step revealed great promise with respect to determining trip purpose by GPS-derived data, specifically from the trip destination coordinate information (including land use and address), the arrival time at the destination, and the activity duration time (i.e., the time gap between trips). Of the 151 trips detected, 54% were trips made either to home or work. Another 22% of the trips were identified as having a mixed-use or ambiguous land use, which indicates that a CATI clarification question is needed to obtain the actual trip purpose. Finally, approximately 7% of the trips were given an incorrect trip purpose – which resulted from
incorrect land uses associated with data quality errors. Much can be done to reduce these types of errors and the next phase of research will concentrate in part on this issue.

Although this research focused on in-vehicle trips only, the results could be extended to other modes of travel as well with the use of personal GPS data loggers. This research concedes that there is no GPS or GIS-related method to determine the driver of the vehicle and the number of passengers. If these details are necessary, they will need to be collected during the CATI household data retrieval call. However, if generalized household travel patterns could be collected during the recruitment call, it would be possible to estimate these variables.

**Recommendations**

The following recommendations for future development and implementation of this new approach for replacing travel diaries with GPS data loggers in traditional household travel surveys are based on the findings of this study:

1) Develop rugged GPS data logging equipment
   - Provide secure connections.
   - Eliminate user interface.
   - Provide continuous external power supply.
   - Ensure ability to withstand temperature extremes using rugged equipment.
2) Improve data logging software

- Do not perform local time zone offset in software; perform it as post processing step instead.
- Log “no signal available” information so that a distinction can be made between this event, loose cabling (partial data streams), and equipment package non-use (no data stream).

3) Continue improving the GPS data processing steps as presented in this dissertation

- Automate manual steps wherever practical, realizing that there will be a need for manual inspection of exceptions along with CATI follow up calls for a small percentage of trips.
- Assume the collection of home, work, and school / day care addresses occurs during the recruitment call so that these locations area available during data processing; modify all three steps accordingly.
- Consider capture of other common destinations in advance as well (grocery store, other shopping locations, gas stations, restaurants, etc) to allow further automation and higher accuracy levels.

A) Improve Trip Detection Module

- Identify and eliminate erroneous trips (i.e., non-movement or false trips).
- Flag questionable trips automatically for exception handling.

B) Improve Land Use and Address Assignment Module

- Improve land use description accuracies within parcel-level database.
• Improve parcel data property boundary information to convert point data to polygon so that GPS points can be automatically assigned to properties using GIS point-in-polygon techniques.

• Automate detection of location inconsistencies between trip ends and subsequent origins.

• Automate identification of round trips recorded without a stop detection.

C) Improve Trip Purpose Derivation Module

• Develop plan to prevent or limit incorrect trip purpose assignments by address error sources (e.g., select power and data cables with secure connections, audit land use inventory for accuracy and consistency, obtain boundary information for property point data).

• Develop plan to minimize the number of code “99” follow up calls.

D) Improve trip route / travel distance module

• Develop trip map matching algorithms for route identification and trip length

4) Develop CATI implementation plans

• Develop end-to-end process plan.

• Develop CATI interface.

• Identify and implement a pilot project using real-world survey data and GPS data loggers to further investigate the automatic derivation of trip details from the GPS data and to evaluate the use and effectiveness of the CATI interface in collecting any missing data elements.
Finally, given the rapid increase in Internet accessibility throughout the world, it is quite feasible that the derived travel data created by the process proposed in this dissertation could be presented back to survey participants directly via a web site. This web site would present the derived trip data, along with highlighted fields prompting the participant to confirm questionable data and to complete missing data. GIS displays of travel routes, including departure and arrival times, could also be made available on the site to assist the participant in remembering the individual trips. Access to each household’s travel data would be password protected. If additional web pages (or screens) are developed to capture household, vehicle, and personal information, then the entire CATI data retrieval process could be eliminated. This would result in substantial cost savings while increasing the amount of data and the quality of data collected.
CHAPTER IX

A LOOK FORWARD

Given the fast-changing pace of technology innovations, it is critical that research continue on GPS and other positioning technologies. On May 1, 2000, President Clinton issued a statement announcing the official termination of Selective Availability effective that night at midnight. He stated that this improvement in accuracy for the civilian population was necessary as more and more application areas had integrated GPS technology into their systems and procedures. The impact of Selective Availability and the removal of this government-induced error are significant with respect to both equipment requirements and performance. Without SA, the remaining error levels in uncorrected GPS data should fall within the 10 to 20 meter range, making this a completely viable option for travel studies. In other words, DGPS is no longer a requirement for any travel studies with some flexibility in the project position and speed accuracy requirements. This accuracy improvement will also allow GPS technology to become widespread as a variety of applications switch to this relatively cheap positioning technology.

If higher accuracy levels are needed, real-time differential corrections (DGPS) available via the NDGPS project, which provides accuracy levels of 3 to 5 meters, is still a viable option. In addition, inverse differential correction technologies, in which
uncorrected GPS field data is transferred back to a base station in real time, are improving rapidly and offer significant advantages with respect to rover data storage for applications that do not require high levels of accuracy in the field.

There are also several other technologies that should be examined for potential use in positional data collection. Dead reckoning devices are now being used in combination with GPS data collection efforts to supplement and/or enhance the positional data for areas in which GPS and/or differential correction signal reception is poor or nonexistent. There has also been much discussion and analysis recently of the use of cellular technologies for determining vehicle position in response to the U.S. government's E911 requirement. This FCC order mandates that cellular phones must convey the caller's position to within 125 meters in 67 percent of all measurements by October 31, 2001. Since the current configuration and/or density of U.S. cellular transmission towers cannot support this level of accuracy, several cellular phone manufacturers are investigating options to integrate GPS and cellular technology. As a result, it may soon be feasible that cellular phones could collect and transmit GPS coordinate information.

As a result of continuing technological improvements, along with ongoing research and applications using these positioning technologies, it is likely that the following events will occur within the next year in the area of travel data collection:

- Production versions of electronic travel diaries and passive in-vehicle GPS systems will become available.
- A variety of spatial and temporal data processing algorithms will be available.
• MPO’s will begin to understand what GPS technology offers in the area of travel studies.

• Extended travel surveys using GPS data loggers will occur.

• Pilot studies will be conducted to test the feasibility of using cellular phones in the collection of travel data.

• Internet-based travel surveys will be implemented.

• Multiple travel survey data collection methods will be used simultaneously across a survey population to increase response rates.

• Research focusing on the collection of GPS travel data to replace the paper travel diary portion of traditional household travel surveys will continue.

Because passive GPS data logging enables extended travel studies, it is highly probable that the development of a “Nielson” family of vehicles will occur soon. In this type of study, a sample of households could be recruited to participate for a multi-year period. Each vehicle of the household would be instrumented with a GPS data logger and a communications link to transfer the GPS data back to a central location on a regular basis. In a method similar to that used for television’s Nielson ratings, the travel and activity patterns of this sample could be analyzed and monitored over time to assess the impact of congestion or transportation control measures on travel behavior.
APPENDICES

APPENDIX A: FOUR EXAMPLE TRAVEL DIARIES

Example 1: Trip-Based Diary
Example 2: Portland Activity-Based Travel Survey
Example 3: Transportation Futures Place-Based Travel Survey
Example 4: Memory Jogger

APPENDIX B: NMEA 0183 GPS SENTENCE FORMATS

APPENDIX C: INSTRUCTION SHEET FOR EQUIPMENT PACKAGE P1

APPENDIX D: MACRO FOR DETECTING INDIVIDUAL TRIPS

APPENDIX E: MACRO FOR CALCULATING TRIP DISTANCES

APPENDIX F: TRIP DIARY COMPARISONS
APPENDIX A

FOUR EXAMPLE TRAVEL DIARIES

Example 1: Trip-Based Diary

Journey 16
PLEASE FILL IN A SEPARATE PAGE FOR EACH PART OF YOUR JOURNEY

<table>
<thead>
<tr>
<th>PLACE</th>
<th>FINISH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ROAD &amp; NUMBER</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AREA/TOWN</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>POSTCODE</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TIME (Please tick)</th>
<th>AM</th>
<th>PM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>WHAT DAY WAS IT?</th>
<th>WHAT WAS YOUR METHOD OF TRAVEL? (Tick one)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday</td>
<td>Caravan driver</td>
</tr>
<tr>
<td>Tuesday</td>
<td>Motorcycle or mopeds</td>
</tr>
<tr>
<td>Wednesday</td>
<td>Caravan passenger</td>
</tr>
<tr>
<td>Thursday</td>
<td>Goods vehicle</td>
</tr>
<tr>
<td></td>
<td>Pedal Cycle</td>
</tr>
<tr>
<td></td>
<td>Walk</td>
</tr>
<tr>
<td></td>
<td>Bus</td>
</tr>
<tr>
<td></td>
<td>Taxi</td>
</tr>
<tr>
<td></td>
<td>Train</td>
</tr>
<tr>
<td></td>
<td>Other</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>WHAT WAS THE MAIN REASON FOR THIS PART OF YOUR JOURNEY?</th>
<th>IF DRIVER/PASSENGER OF CAR/VAN/MOTORCYCLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Tick main one only)</td>
<td>Which vehicle were you in? (See page 10 of diary for code)</td>
</tr>
<tr>
<td>To go to work/education/for work</td>
<td>Was you delayed? Yes No</td>
</tr>
<tr>
<td>To go home</td>
<td>by traffic congestion?</td>
</tr>
<tr>
<td>To take a passenger somewhere (incl. kids to school)</td>
<td>Did you park On-street Off-street (Please tick)</td>
</tr>
<tr>
<td>Social/entertainment</td>
<td>Was the parking free? Paid</td>
</tr>
<tr>
<td>Shopping or personal business (e.g. doctor, bank)</td>
<td>(Please tick)</td>
</tr>
<tr>
<td>Other</td>
<td>How long did you spend searching/queuing for parking?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IF SHOPPING OR PERSONAL BUSINESS</th>
<th>HOW MANY ADULTS ( traveld/ travelled? Include self)</th>
<th>HOW MANY CHILDREN WERE WITH YOU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food/grocery shopping</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other shopping (incl. petrol)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal Business</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Example is from K.W. Axhausen's Travel Diaries: An Annotated Catalogue, 1994.)
Example 2: Portland Activity-Based Travel Survey

Day _______ Activity ________

1. What was your activity?

2. Where did your activity take place?
   Name of place: ___________________
   Address: _________________________   City: ____________________________

3. What time did your activity start?
   ___________ am / pm

4. What time did your activity end?
   ___________ am / pm

5. Did you have to travel to get to this activity?
   Yes [ ] No [ ]

6. What time did your travel start?
   ___________ am / pm

7. What time did your travel end?
   ___________ am / pm

8. How did you travel to the activity? (Circle one and follow instructions)
   Private Vehicle [ ] Public [ ] Train [ ] Walk [ ] Bicycle [ ] School [ ] Other [ ]

9. Which vehicle did you use?
   Household [ ] Other [ ]

10. Were you the
    Driver [ ] Passenger [ ]

11. How many people were in the vehicle (including driver)?

12. Where did you park?
    Street [ ] Drive-through [ ] Driveway [ ] Parking lot/garage [ ] Other [ ]

13. How did you pay for the parking?
    Did not pay [ ] Hourly [ ] Weekly [ ] Semesterly [ ] Daily [ ] Monthly [ ] Other [ ]

14. How much did you pay for parking?
    $ __________

15. Who subsidized your parking?
    No-One [ ] Employer [ ] Business/store [ ] Other [ ]

16. What did you have a vehicle available?
    Yes [ ] No [ ]

17. How would you have paid for parking if you went by car?
    Would not pay [ ] Hourly [ ] Weekly [ ] Semesterly [ ] Daily [ ] Monthly [ ] Other [ ]

18. How much would you have had to pay for parking if you went by car?
    $ __________

19. How much did you pay for parking?
    $ __________

20. Who subsidized your transit fare?
    No one [ ] Employer [ ] Business/store [ ] Other [ ]

21. Did you transfer to another bus or train?
    Yes [ ] No [ ]

22. To what line did you transfer?

23. How many people were in your party?

24. How did you pay for your trip?
    Cash [ ] Fareless square [ ] Ticket [ ] Transfer [ ] Pass [ ] Other [ ]

25. How did you get to the stop?
    Drove & parked [ ] Dropped off [ ] Carpoled [ ] Walked [ ] Other [ ]

26. How did you get from the stop to your destination?
    Drove & parked [ ] Dropped off [ ] Carpoled [ ] Walked [ ] Other [ ]

27. What was the first transit route taken?

28. Where did you board?
    Address/place ______________________
    Cross streets: ______________________
    City: _____________________________

29. Did you have a vehicle available?
    Yes [ ] No [ ]

30. How would you have paid for parking if you went by car?
    Would not pay [ ] Hourly [ ] Weekly [ ] Semesterly [ ] Daily [ ] Monthly [ ] Other [ ]

31. How much would you have had to pay for parking if you went by car?
    $ __________

32. How much did you pay for parking?
    $ __________

33. Who subsidized your parking?
    No-One [ ] Employer [ ] Business/store [ ] Other [ ]

34. What was the full unsubsidized price to park?
    $ __________

35. How much did you have to pay for parking if you went by car?
    $ __________

36. How many people were in your party?

37. To what did you change?
    Walk [ ] Public bus [ ] Train [ ] Private vehicle [ ] Bicycle [ ] School bus [ ] Other [ ]

38. Where did you change travel modes?
    Address/place ______________________
    Cross streets: ______________________
    City: _____________________________

Answer Q. 9 – 16 Answer Q. 17 – 20 Answer Questions 29 – 32
If you traveled by public bus/ train answer Q. 17 – 20, then 33 – 34

166
Example 3: Transportation Futures Place-Based Travel Survey

Place #2

- My Home
- My Regular Workplace
- My School
- Other Place (address already provided)
- A New Place

Name of Place (if any): ________________________________

Street Address: ______________________________________
	City: __________ State: __________ Zip: __________

Nearest Cross Streets: ________________________________

At WHAT TIME did you ARRIVE at Place #2?

: ___________________________ am/pm

HOW did you get from Place #1 to Place #2?

Show ALL the methods of travel you used to make this trip.

1st 2nd 3rd 4th 5th

If you used TRANSIT Bus, Rail, Subway, Ferry, Other

First board: __________ at: __________ Line # Service __________________

1st transfer: __________ at: __________ Station Name (if Rail or Subway)

2nd transfer: __________ at: __________

Last Station at: __________

WHAT did you do at Place #2? (Check all that apply)

- Drop off/pick-up someone
- Visit friends/relatives
- Eat meals
- Social/recreational/entertainment
- Shop
- Doctor/dentist/other professional
- Other family or personal business
- Religious or civic
- Other activities not-at-home (Specify): __________________________

From Place #2 did you go to another place during your 24-hour day?

NO- This was your LAST place for the 24-hour day.

Check here: ☐ DONE

YES- At what time did you leave Place #2 to go to Place #3?

: ___________________________ am/pm

NEXT PLACE #3
Example 4: Memory Jogger

Please fill out during the DIARY DAY

<table>
<thead>
<tr>
<th>Trip #</th>
<th>Activity at Destination</th>
<th>Destination / Place Name</th>
<th>Address / Street &amp; Cross Street</th>
<th>Trip Stop Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Activity: Write down each thing you did in a different location

Location: Write down the name and address (or cross streets) of each place you went

Start Time: Write down when you started doing each activity

Stop Time: Write down when you finished doing each activity

Do not include travel as an activity
APPENDIX B

NMEA 0183 GPS SENTENCE FORMATS

**Recommended Minimum Specific GPS / Transit Data (RMC)**

$$GPRMC,<1>,<2>,<3>,<4>,<5>,<6>,<7>,<8>,<9>,<10>,<11>*hh<CR><LF>$$

- **<1>** UTC time of position fix, hhmmss format
- **<2>** Status, A= Valid Position, V= NAV receiver warning
- **<3>** Latitude, ddmm.mmmm format (leading zeros will be transmitted)
- **<4>** Latitude hemisphere, N or S
- **<5>** longitude, ddmm.mmmm format (leading zeros will be transmitted)
- **<6>** Longitude hemisphere, E or W
- **<7>** Speed over ground, 0.0 to 999.9 knots
- **<8>** Course over ground, 0.0 to 359.9 degrees, true (leading zeros will be transmitted)
- **<9>** UTC date of position fix, ddmmyy format
- **<10>** magnetic variation, 000.0 to 180.0 degrees (leading zeros will be transmitted)
- **<11>** Magnetic variation direction, E or W (westerly variation adds to true course)

**Global Positioning System Fix Data (GGA)**

$$GPGGA,<1>,<2>,<3>,<4>,<5>,<6>,<7>,<8>,<9>,M,<10>,M,<11>,<12>*hh<CR><LF>$$

- **<1>** UTC time of position fix, hhmmss format
- **<2>** Latitude, ddmm.mmmm format (leading zeros will be transmitted)
- **<3>** Latitude hemisphere, N or S
- **<4>** Longitude, ddmm.mmmm format (leading zeros will be transmitted)
- **<5>** Longitude hemisphere, E or W
- **<6>** GPS quality indication, 0 = fix not available, 1 = Non-differential GPS fix available, 2 = Differential GPS (DGPS) fix available
- **<7>** # of satellite in use, 00 to 12 (leading zeros will be transmitted)
- **<8>** Horizontal dilution of precision, 1.0 to 99.9
- **<9>** Antenna height above / below mean sea level, - 9999.9 to 99999.9 meters
- **<10>** Geodial height, -999.9 to 9999.9 meters
- **<11>** Differential GPS (RTCM – SC104) data age, # of seconds since last valid RTCM transmission (null if non – DGPS)
- **<12>** Differential Reference Station ID, 0000 to 1023 (leading zeros will be transmitted, null if non– DGPS ).
GPS DOP and Active Satellites (GSA)
$GPGSA,<1>,<2>,<3>,<3>,<3>,<3>,<3>,<3>,<3>,<3>,<3>,<3>,<3>,<3>,<3>,<3>,<3>,<4>,
<5>,<6>,*hh<CR><LF>

<1> Mode, M = manual, A = automatic
<2> Fix type, 1 = not available, 2 = 2D, 3 = 3D
<3> PRN Number, 01 to 32, of satellite used in solution, up to 12 transmitted (leading
zeros will be transmitted)
<4> Position dilution of precision, 1.0 to 99.9
<5> Horizontal dilution factor of precision, 1.0 to 99.9
<6> Vertical dilution of precision, 1.0 to 99.9

GPS Satellites in View (GSV)
$GPGSV,<1>,<2>,<3>,<4>,<5>,<6>,<7>*hh<CR><LF>

<1> Total # of GSV sentences to be transmitted
<2> # of current GSV sentence
<3> Total # of satellites in view, 00 to 12 (leading zeros will be transmitted)
   <4> Satellite PRN number, 01 to 32 (leading zeros will be transmitted)
<5> Satellite elevation, 00 to 90 degrees (leading zeros will be transmitted)
<6> Satellite azimuth, 000 to 359 degrees, true (leading zeros will be transmitted)
<7> Signal to noise ratio (C/No) 00 to 99 db, null when not tracking (leading zeros will
   be transmitted)

Note: Items <4>, <5>, <6> and <7> repeat for each satellite in view to a maximum of
four (4) satellites per sentence. Additional satellites in view information must be sent in
subsequent sentences. These fields will be null if unused.

[Sources: Garmin 1996, NMEA 1998]
APPENDIX C

INSTRUCTION SHEET FOR EQUIPMENT PACKAGE P1

Instructions for Travel Study Participants

Thank you for participating in our study. We are collecting data on travel patterns so that we can create data processing algorithms for the Atlanta metropolitan travel survey, which will start in the fall of 2000.

The equipment package that will be installed in your car is very simple – it has a Palm IIIx data logger, a GPS receiver and antenna, a FM antenna and receiver, and a power connector to the cigarette lighter. Here are a few guidelines for installation and usage:

Attach both GPS antennas to roof of vehicle.
Select the door on your vehicle that is seldom used. Wind down the window for that door, place both antennas on the roof of your vehicle directly above that window with 1-2 feet of space between them, and then place the rest of the equipment through the open window into the interior of your vehicle.
Do not slide the antennas across your roof – you should place them directly where you want it to be.
Close the window as much as possible without crimping the wires to the antennas – leave a small gap for these cables. If you have automatic windows – do not let them close all the way! Also, do not allow the door to close on the cables.
If you need to open the door at any time during the survey, remember to remove the antennas first, then open the door, then close the door, and finally place the antennas back on the vehicle.
If it rains while you are driving and water is entering your vehicle, you can remove the antennas from the roof and place the small one securely on your dashboard or rear shelf, whichever is more convenient. The longer antenna could also be placed on the rear shelf.

Plug the cigarette lighter connector into your cigarette lighter.
If you smoke or power your cell phone by the cigarette lighter, we will provide you with a power splitter for your lighter.

Each time you turn on the vehicle, you need to turn on the Palm IIIx.
To start the Palm, open the Palm cover and press the green button on the lower left side – this should turn on the Palm. The top of the screen should have the words “Connect GPS and Click Ok” – you should remove the stylus from the top of the Palm, tap firmly on the OK button, verify that data is displayed on the Palm screen, return the stylus to its storage slot, and then close the Palm cover. If you do not see data refreshing on the Palm
screen after tapping OK, verify that the power connector is inserted completely into the cigarette lighter and that the data cable is securely connected to the bottom of the Palm.

If, during your data collection period, a message appears on the Palm indicating that your battery level is low, please stop logging data, shut off the unit, and quickly replace the batteries in the Palm with the two extra batteries provided in the equipment bag. Then restart the data logging process.

Each time you turn off the vehicle, you should turn off the Palm.
To shut off the Palm, open the cover, remove the stylus, tap firmly on the Stop button, and press the green button to turn off the power.

Returning the equipment
Upon arrival to GA Tech on the morning that you are scheduled to return the equipment, simply shut off the Palm as explained above, remove the power connector from the cigarette lighter, and remove the antennas from the roof. Note: do not slide the antennas; tilt them until you are able to remove them easily. Please take the equipment to your office and someone will stop by to pick it up in the morning. Or, if you prefer, one of the researchers can assist you with removing the equipment.
APPENDIX D

MACRO FOR DETECTING INDIVIDUAL TRIPS

Macro "Read_TRIPS"

// Allow user to specify input and output file names

inp_file_name=choosefile(["logfiles","*.txt"],"Choose a Garmin35 log file.
out_file_name=choosefilename(["output","*.dbf"],"Choose the Output dbf file.

//Initialize variables
field_info=["ID","Integer",8,null,"Yes"],
   ["TRIP","Integer",8,null,"Yes"],
   ["TYPE","String",12,null,"No"],
   ["DATE","String",12,null,"No"],
   ["TIME","Real",20,9,"No"],
   ["LAT","Real",12,6,"No"],
   ["LONG","Real",12,6,"No"],
   ["HEAD","Real",6,1,"No"],
   ["SPEED","Real",6,1,"No"],
   ["DIFF_FLAG","Integer",2,null,"Yes"],
   ["SATELLITES","Integer",12,"No"],
   ["HDOP","Real",5,1,"No"],
   ["DIFF_AGE","Integer",2,null,"Yes"],
   ["HOUR","Integer",8,null,"No"],
   ["MIN","Integer",8,null,"no"],
   ["SEC","Integer",10,null,"no"]

gap = 120
timediff=0
speed=0
trip=0
Start="START"
End="END"
id=0

table_name=createtable(out_file_name,out_file_name,"DBASE",field_info)
inp_file=openfile(inp_file_name,"r")
txt_line=readline(inp_file)

// read first record
items=parsestring(txt_line,",")
date=items[1]
time_str=items[2]
hour=stringtoint(substring(time_str,1,2))
minute=stringtoint(substring(time_str,3,2))
second=stringtoint(substring(time_str,5,2))
time=(hour*3600)+(minute*60)+second
lat=stringtoreal(items[3])
long=stringtoreal(items[4])
head=stringtoreal(items[5])
speed=stringtoreal(items[6])
diff=stringtoint(items[7])
sats=stringtoint(items[8])
hdop=stringtoreal(items[9])
dage=stringtoint(items[10])

// increment record pointer
id=id+1

// Detect and print trip start as first record with speed above the non-movement threshold
begin_loop:
while((speed<3 and diff=1) or (speed<1 and diff=2)) do
  if (FileAtEOF(inp_file)) then do
    goto quit_loop
  end
  txt_line=readline(inp_file)
  items=parsestring(txt_line,",")

  date=items[1]
  time_str=items[2]
  hour=stringtoint(substring(time_str,1,2))
  minute=stringtoint(substring(time_str,3,2))
  second=stringtoint(substring(time_str,5,2))
  time=(hour*3600)+(minute*60)+second
  lat=stringtoreal(items[3])
  long=stringtoreal(items[4])
  head=stringtoreal(items[5])
  speed=stringtoreal(items[6])
  diff=stringtoint(items[7])
  sats=stringtoint(items[8])
  hdop=stringtoreal(items[9])
  dage=stringtoint(items[10])
  id=id+1

end

// Trip start record found. Increment trip counter and write start record.
trip=trip+1
addrecord(table_name,{{"ID",id},}
//Detect and print trip end.  If end of file detected, go to write last record as trip end

if (FileAtEOF(inp_file)) then do
    goto quit_loop
end

// Increment record pointer and read next record.  Store values in *1 variable names for forward comparisons
id=id+1
txt_line=readline(inp_file)
items=parsestring(txt_line","

date=items[1]
date1=date
day=stringtoint(substring(date,7,2))
day1=day
time_str=items[2]
hour=stringtoint(substring(time_str,1,2))
minute=stringtoint(substring(time_str,3,2))
second=stringtoint(substring(time_str,5,2))
time=(hour*3600)+(minute*60)+second
time1=time
hour1=hour
minute1=minute
second1=second
lat=stringtoreal(items[3])
lat1=lat
long=stringtoreal(items[4])
long1=long
head=stringtoreal(items[5])
head1=head
speed=stringtoreal(items[6])
speed1=speed
diff=stringtoint(items[7])
diff1=diff
sats=stringtoint(items[8])
sats1=sats
hdop=stringtoreal(items[9])
hdop1=hdop
dage=stringtoint(items[10])
dage1=dage
tempt=time1
id1=id

// read next record and check for time gap or date gap. Check first for end of file condition.

stop_loop:

if (FileAtEOF(inp_file)) then do
  goto quit_loop
end

txt_line=readline(inp_file)
items=parsestring(txt_line","")

date=items[1]
day=stringtoint(substring(date,7,2))
day2=day
time_str=items[2]
    hour=stringtoint(substring(time_str,1,2))
    minute=stringtoint(substring(time_str,3,2))
    second=stringtoint(substring(time_str,5,2))
    time=(hour*3600)+(minute*60)+second
lat=stringtoreal(items[3])
long=stringtoreal(items[4])
head=stringtoreal(items[5])
speed=stringtoreal(items[6])
diff=stringtoint(items[7])
sats=stringtoint(items[8])
hdop=stringtoreal(items[9])
dage=stringtoint(items[10])
id=id+1

// look for date change, if yes, check for time gap between records

if datechange<>0 then do
  timenew = time + (3600*24)
timechg = timenew - tempt
  if timechg > gap then do
goto end_loop
end
end

// look for time gap
timediff=time-tempt
if timediff>gap then do
goto end_loop
end

// if no time gap, determine if vehicle has stopped but is still logging GPS data
if (diff=1 and speed<3 ) then do
while(speed<3 and diff=1) do
if (FileAtEOF(inp_file)) then do
goto quit_loop
end

txt_line=readline(inp_file)
items=parsestring(txt_line,"",""

date=items[1]
day=stringtoint(substring(date,7,2))
day2=day
time_str=items[2]
hour=stringtoint(substring(time_str,1,2))
minute=stringtoint(substring(time_str,3,2))
second=stringtoint(substring(time_str,5,2))
time=(hour*3600)+(minute*60)+second
lat=stringtoreal(items[3])
long=stringtoreal(items[4])
head=stringtoreal(items[5])
speed=stringtoreal(items[6])
diff=stringtoint(items[7])
sats=stringtoint(items[8])
hdop=stringtoreal(items[9])
dage=stringtoint(items[10])
id=id+1

// check for date change between records. If yes, check for time gap.
datechange=day2-day1
if datechange<>0 then do
timenew = time + (3600*24)
timechg = timenew - tempt
if timechg > gap then do
goto end_loop
end
end
end
// Vehicle is moving. Check for time gap since previous movement. If yes, write previous
// movement as trip end and write current movement as the next trip start.
timediff=time-tempt
if timediff>gap then do
    addrecord(table_name,{{"ID",id1},
          {"TRIP",trip},
          {"TYPE",End},
          {"DATE",date1},
          {"TIME",time1},
          {"LAT",lat1},
          {"LONG",long1},
          {"HEAD",head1},
          {"SPEED",speed1},
          {"DIFF_FLAG",diff1},
          {"SATELLITES",sats1},
          {"HDOP",hdop1},
          {"DIFF_AGE",dage1},
          {"HOUR",hour1},
          {"MIN",minute1},
          {"SEC",second1} })
    trip=trip+1

    addrecord(table_name,{{"ID",id},
          {"TRIP",trip},
          {"TYPE",Start},
          {"DATE",date},
          {"TIME",time},
          {"LAT",lat},
          {"LONG",long},
          {"HEAD",head},
          {"SPEED",speed},
          {"DIFF_FLAG",diff},
          {"SATELLITES",sats},
          {"HDOP",hdop},
          {"DIFF_AGE",dage},
          {"HOUR",hour},
          {"MIN",minute},
          {"SEC",second} })
end
// Store current record as base record for next set of forward comparisons
date1=date
day1=day
time1=time
hour1=hour
minute1=minute
second1=second
lat1=lat
long1=long
head1=head
speed1=speed
diff1=diff
sats1=sats
hdop1=hdop
dage1=dage
tempt=time1
id1=id

// Look for next trip stop
goto stop_loop

// End of trip detected based on time gap between records. Write trip end record
// and go back to begin loop to detect next trip start.
end_loop:
addrecord(table_name,{{"ID",id1},
                 {"TRIP",trip},
                 {"TYPE",End},
                 {"DATE",date1},
                 {"TIME",time1},
                 {"LAT",lat1},
                 {"LONG",long1},
                 {"HEAD",head1},
                 {"SPEED",speed1},
                 {"DIFF_FLAG",diff1},
                 {"SATELLITES",sats1},
                 {"HDOP",hdop1},
                 {"DIFF_AGE",dage1},
                 {"HOUR",hour1},
                 {"MIN",minute1},
                 {"SEC",second1} })
goto begin_loop

// End of File detected, write last record as trip end
quit_loop:
addrecord(table_name,{{"ID",id},
                     {"TRIP",trip},
                     {"TYPE",End},
                     {"DATE",date},
                     {"TIME",time},
                     {"LAT",lat},
                     {"LONG",long},
                     {"HEAD",head},
                     {"SPEED",speed},
                     {"DIFF_FLAG",diff},
                     {"SATELLITES",sats},
                     {"HDOP",hdop},
                     {"DIFF_AGE",dage},
                     {"HOUR",hour},
                     {"MIN",minute},
                     {"SEC",second} })
"TYPE",End},
{"DATE",date},
{"TIME",time},
{"LAT",lat},
{"LONG",long},
{"HEAD",head},
{"SPEED",speed},
{"DIFF_FLAG",diff},
{"SATELLITES",sats},
{"HDOP",hdop},
{"DIFF_AGE",dage},
{"HOUR",hour},
{"MIN",minute},
{"SEC",second} )

closefile(inp_file)
closeview(table_name)

dendmacro
APPENDIX E

MACRO FOR CALCULATING TRIP DISTANCES

//This macro will calculate trip distances from GPS data logged at 1 second intervals

Macro"calc_dist"

inp_file_name=choosefile({"logfiles","*.csv"},"Choose a trip csv file.")
out_file_name=choosefilename({"output","*.dbf"},"Choose the Output dbf file.")

//Initialize variables
field_info={"ID","Integer",8,null,"Yes"},
{"TRIP","Integer",8,null,"Yes"},
{"TYPE","String",12,null,"No"},
{"DATE","String",12,null,"No"},
{"TIME","Real",20,9,"No"},
{"LAT","Real",12,6,"No"},
{"LONG","Real",12,6,"No"},
{"HEAD","Real",6,1,"No"},
{"SPEED","Real",6,1,"No"},
{"DIFF_FLAG","Integer",2,null,"Yes"},
{"SATELLITES","Integer",12,"No"},
{"HDOP","Real",5,1,"No"},
{"DIFF_AGE","Integer",2,null,"Yes"},
{"HOUR","Integer",8,null,"no"},
{"MIN","Integer",8,null,"no"},
{"SEC","Integer",10,null,"no"}}

speed=0
dist = 0
totaldist = 0
id=0

table_name=createtable(out_file_name,out_file_name,"DBASE",field_info)
inp_file=openfile(inp_file_name,"r")

//Calculate second-by-second travel distances and keep total
read_loop:
while(!(FileAtEOF(inp_file)))do

txt_line=readline(inp_file)
items=parsestring(txt_line,"",""

id=items[1]
date=items[2]
time_str=items[3]
lat=stringtoreal(items[4])
long=stringtoreal(items[5])
head=stringtoreal(items[6])
speed=stringtoreal(items[7])
diff=stringtoint(items[8])
sats=stringtoint(items[9])
hdop=stringtoreal(items[10])
dage=stringtoint(items[11])

dist = speed/3600
totaldist = totaldist + dist

der

dodd record(table_name,{{ "LAT" , totaldist} })

done:
closefile(inp_file)
closeview(table_name)
endmacro
APPENDIX F

TRIP DIARY COMPARISONS

The diary comparisons are provided on the following pages in Study ID order. Refer to the first section of Chapter 7 for a complete description of the contents of each diary log.

The class assignment for each log based on the trip detection comparisons can be found in the upper right-hand corner of each log. These classes are:

Class 1: Complete Match

Class 2: Detection of Trips Not Reported

Class 3: Non-Detection of Stops of Short Duration

Class 4: Potential Problems
33 pages of diary comparisons
REFERENCES


Jean Wolf has extensive experience in the development and management of transportation-related technology projects. In her current role as the automation manager for SMARTRAQ, she developed three instrumentation packages to automate the capture of travel and vehicle information. SMARTRAQ will assist with the Year 2000 Household Travel Survey for the Atlanta metropolitan region. Ms. Wolf is also a co-principal investigator, serving as the instrumentation manager, for the NHTSA-sponsored Speed / Crash Study that is being conducted in the Atlanta metropolitan region. In this study, 1100 vehicles will be instrumented with GPS receivers, antennas, and data loggers, as well as with tri-axial accelerometers and a CPU, for controlling the storage and transmission of speed and crash detection data.

In pursuit of further innovations in the field, Ms. Wolf is launching a new company, GeoStats, to concentrate on the development and use of emerging technologies in the collection, processing, display, and presentation of transportation data. These technologies will focus on GPS and GIS, and will extend to handheld data collectors, real-time wireless data transfer, and internet-based media.
Ms. Wolf is an active member of the Transportation Research Board’s Committee on Urban Transportation Data and Information Systems and is a co-author of the committee’s white paper entitled “The Future of Urban Transportation Data (Transportation in the New Millennium – A Look Forward).” In this publication, she provided a state-of-the-art review of the use of electronic travel diaries, GPS receivers, and vehicle instrumentation to enhance the quality and quantity of urban travel data. Ms. Wolf has made numerous presentations of her GPS-related research at several annual Transportation Research Board conferences, as well as at other travel survey and travel behavior conferences. In 1999, Ms. Wolf’s NCHRP Synthesis Report request, entitled “GPS to GIS – Turning Data into Information” was accepted as one of only 12 projects awarded in that year.

Prior to pursuing her Ph.D., Ms. Wolf was employed by United Parcel Service as a technology manager on a variety of UPS operations and customer service automation projects. These projects utilized technologies ranging from handheld data collectors and radio-frequency scanners to stand-alone PCs, LANs, and mainframes. Her responsibilities included development and support of UPS’s international computer systems located around the world, which are used to forecast and clear shipments through customs prior to their arrival in the destination country. Her yearly information systems budget from 1992 through 1994 averaged more than $5 million and her staff included 30 managers and supervisors.