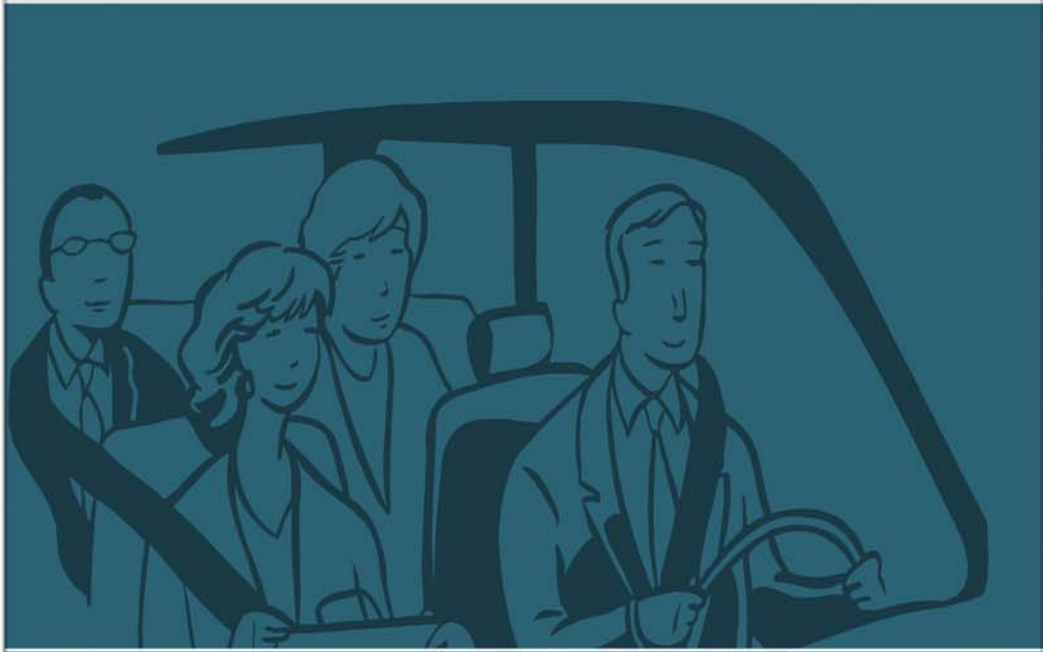


Vehicle Occupancy Data Collection Methods

Final Report
Contract No. BD-015-9

February 2005



Prepared for
Transportation Statistics Office
Florida Department of Transportation



Prepared by
Lehman Center for Transportation Research
Florida International University

Final Report

Vehicle Occupancy Data Collection Methods

Contract No. BD-015-9

Prepared by:

Albert Gan, Ph.D., Associate Professor
Rax Jung, Ph.D., Research Associate
Kaiyu Liu, Senior Software Engineer
Xin Li, Graduate Research Assistant
Diego Sandoval, Graduate Research Assistant

Lehman Center for Transportation Research
Department of Civil and Environmental Engineering
Florida International University
10555 West Flagler Street, EC 3680
Miami, Florida 33174
Phone: (305) 348-3116
Fax: (305) 348-2802
E-mail: gana@fiu.edu

in cooperation with the

State of Florida Department of Transportation
605 Suwannee Street, MS 30
Tallahassee, FL 32399-0450

“The opinions, findings and conclusions expressed in this publication are those of the authors and not necessarily those of the State of Florida Department of Transportation”

February 2005

Technical Report Documentation Page

1. Report No. Final Report for BD-015-9		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle Vehicle Occupancy Data Collection Methods				5. Report Date February 2005	
				6. Performing Organization Code	
7. Author(s) Albert Gan, Rax Jung, Kaiyu Liu, Xin Li, and Diego Sandoval				8. Performing Organization Report No.	
9. Performing Organization Name and Address Lehman Center for Transportation Research Florida International University 10555 W. Flagler Street, EC 3680, Miami, FL 33174				10. Work Unit No. (TRAIS)	
				11. Contract or Grant No. BD-015-9	
12. Sponsoring Agency Name and Address Office of Research and Development State of Florida Department of Transportation 605 Suwannee Street, MS 30, Tallahassee, FL 32399-0450				13. Type of Report and Period Covered Final Report October 2003 – February 2005	
				14. Sponsoring Agency Code 99700-3596-119	
15. Supplementary Notes Mr. Gordon Morgan of the Office of Transportation Statistics at the Florida Department of Transportation served as the project manager for this project.					
16. Abstract As congestion management strategies begin to emphasize more person movements than vehicle movements, vehicle occupancy data are becoming increasingly important. With this increasing need for occupancy data comes the need to examine and reexamine the ways in which these data have been, and will be, collected. This project reviews the existing methods of vehicle occupancy data collection, examines issues related to geographic, temporal, and vehicle coverage design of occupancy data collection, and develops guidelines for performing occupancy data collection as well as analyzing occupancy data. Appropriate sampling plans for site-specific, corridor, and regional studies are presented. Potential new methods for collecting occupancy data are discussed. A user-friendly software system that can estimate occupancy rates from multiple years of crash records on the Florida state roadway system was developed as part of this study. The system can estimate occupancy rates for select roadway segment, corridor, or regional level for specific time periods for different types of vehicles. The system also includes a stand-alone GIS interface to facilitate the selection of geographic features and display of occupancy rate estimates. Also developed is a Pocket PC application that can facilitate field data collection based on the commonly used windshield method. A companion program for this Pocket PC application was also developed to compute the average vehicle occupancy rates and related statistics from the field data.					
17. Key Word Vehicle Occupancy, Data Collection Methods, Sampling, Accident Data Applications				18. Distribution Statement	
19. Security Classif. (of this report) Unclassified		20. Security Classif. (of this page) Unclassified		21. No. of Pages 83	
				22. Price	

ACKNOWLEDGEMENTS

This research was funded by the Research and Development Center of the Florida Department of Transportation (FDOT). The authors are grateful to the project manager, Mr. Gordon Morgan, of the FDOT Transportation Statistics Office for his guidance and assistance. Ms. Yujing Xie, a Graduate Research Assistant with the Lehman Center for Transportation Research (LCTR), served as the tester for the FAVORITE software system. Special thanks are due to Mr. David R. Rae, Vice President of URS Corporation, for providing the vehicle occupancy data used in Chapters 4 and 8.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	ii
TABLE OF CONTENTS.....	iii
EXECUTIVE SUMMARY	vi
CHAPTER 1. INTRODUCTION	1
1.2. Research Needs.....	1
1.3. Objectives	2
1.4. Report Organizations	2
CHAPTER 2. EXISTING METHODS AND STUDIES	3
2.1. Vehicle Occupancy Collection Methods	3
2.1.1. Field Data Collection Methods	3
2.1.1.1. Roadside Windshield Method.....	3
2.1.1.2. Carousel Method.....	3
2.1.1.3. Video Surveillance Method	4
2.1.1.4. Computer Vision Method	4
2.1.2. Existing Data Methods.....	4
2.1.2.1. Crash Records.....	4
2.1.2.2. Household Survey.....	5
2.1.2.3. Vehicle Intercept Survey.....	5
2.1.2.4. Census Transportation Planning Package (CTPP).....	5
2.1.2.5. Nationwide Personal Transportation Survey (NPTS).....	6
2.2. Transit Occupancy	6
2.3. Existing Studies	6
CHAPTER 3. STUDY DESIGN CONSIDERATIONS.....	12
3.1. Geographic Coverage.....	12
3.1.1. Geographic Units	12
3.1.2. Geographic Density	12
3.2. Temporal Coverage.....	12
3.2.1. Time-of-Day	12
3.2.2. Day-of-Week.....	12
3.2.3. Month/Season-of-Year.....	13
3.2.4. Temporal Trends.....	13
3.3. Facility Types.....	13
3.4. Data Collection Cycles	14
3.5. Data Collection Locations.....	14
3.6. Vehicle Types	14
3.7. Vehicle Percentages.....	15

CHAPTER 4. STUDY PROCEDURE	16
4.1. Defining Objectives and Selecting Data Collection Method	16
4.2. Establishing Sampling Plan	17
4.2.1. Field Observational Methods	17
4.2.1.1. Major Factors Affecting AVO	17
4.2.1.2. Simple Collection	20
4.2.1.3. Stratified Collection	21
4.2.2. Mail-Out Questionnaires	23
4.2.3. Crash Records	24
4.3. Comparing Costs	24
4.4. Sampling Randomly	25
4.5. Computing AVO	25
4.5.1. Field Observational Methods	25
4.5.2. Mail-Out Questionnaires	27
4.5.3. Crash Records	27
CHAPTER 5. EXTRACTION OF VEHICLE OCCUPANCY RATES FROM ACCIDENT RECORDS	28
5.1. Overview	28
5.2. Installation	29
5.3. Input Specifications	29
5.3.1. Accident Data Years	30
5.3.2. Location Selection	30
5.3.3. Filters	31
5.4. Output Options	31
5.4.1. Table Display	32
5.4.2. Chart Display	33
5.4.3. GIS Display	33
5.5. Variable Re-categorization	37
5.6. Validations	38
5.6.1. Reasonableness Checks	38
5.6.2. Field Data Comparisons	42
CHAPTER 6. AUTOMATED FIELD DATA COLLECTION SYSTEM	44
6.1. Overview	44
6.2. Installation	44
6.3. Main Screen	45
6.4. General Screen	45
6.4.1 Create a New File	46
6.4.2. Open a Existing File	46
6.4.3. Input Location Information	48
6.5. Data Entry Screen	48
6.6. Editing Data Table	50
6.7. Import and Export Database	51
6.7.1. Import from Desktop to Windows CE	51
6.7.2. Export an ADOCE Database to a Desktop	52

6.8. Post-Processing Program	52
CHAPTER 7. POTENTIAL DATA COLLECTION METHODS	55
7.1. Photographic Detection.....	55
7.2. In-Vehicle Detection.....	56
7.2.1. In-Vehicle Systems	56
7.2.2. Information Transmission.....	56
7.3. Remarks	57
CHAPTER 8. STUDY GUIDELINES	58
8.1. General.....	58
8.2. Determining Time Periods.....	58
8.3. Determining Sampling Plan.....	59
8.3.1. Number of Observation Sessions.....	59
8.3.2. Sampling Procedure.....	60
8.4. Field Operation Design.....	62
8.4.1. Locate Counting Sites.....	62
8.4.2. Work Plan	63
8.4.3. Personnel Training.....	63
8.4.4. Data Collection Tool.....	64
8.5. Data Collection Plan	64
8.6. Data Analysis	65
CHAPTER 9. SUMMARY AND RECOMMENDATIONS	69
REFERENCES	72

EXECUTIVE SUMMARY

Introduction

The Intermodal Surface Transportation Efficiency Act (ISTEA) of 1991 expanded the role of vehicle occupancy data by requiring continued monitoring of congestion management strategies, many of which emphasize more person movements than vehicle movements. With this increasing need for vehicle occupancy data comes the need to examine and reexamine the ways in which these data have been, and will be, collected. This report reviewed the existing methods of vehicle occupancy data collection, examined issues related to geographic, temporal, and vehicle coverage design of occupancy data collection, and presented study guidelines for performing occupancy data collection as well as analyzing occupancy data.

Occupancy Data Collection Methods

Existing methods for vehicle occupancy data can be grouped into those that are collected in the field for the sole purpose of computing vehicle occupancy rates and those that are collected for other purposes, but may be used to estimate vehicle occupancy rates. In general, the field collection methods are more suitable for collecting data at the site-specific or corridor level, while the existing data methods are more suitable for area-wide or regional studies. Field data collection methods are more commonly used because they can be tailored to specific application needs in terms of location, sample size and accuracy, etc. Current field collection methods include roadway windshield, carousel, and video surveillance. Existing databases that can be used to generate vehicle occupancy include crash records, vehicle and household surveys, Census Transportation Planning Package (CTPP), Nationwide Personal Transportation Survey (NPTS), etc. The accuracy and scope of applications of existing data are constrained to what have been collected.

Study Design Considerations

Important design considerations for vehicle occupancy data collection include geographic and temporal coverages, facility types, collection cycles, locations, and vehicle types. The general geographic units include site-specific, corridor, and area-wide. Temporal variability in AVO is a common issue across all data collection methods. Existing studies show significant variations in vehicle occupancy rates by time-of-day, day-of-week, and season-of-year. Different types of roads typically have different occupancy levels. For example, roadways of the higher functional hierarchy would typically be expected to have lower AVO. It is important to thus sample all roadway types in order to generate a representative estimate of regional vehicle occupancy. AVOs can also differ from one location to another. The spatial variations of AVO are related to the distribution of household types and work places. It is thus necessary to select many different locations in order to measure AVO variability adequately.

The vehicle types included in a data collection are determined by the purpose for which the data are to be collected. Different study purposes may utilize different criteria for interpreting AVO.

In most vehicle occupancy studies, only data from passenger vehicles or light vehicles (private passenger automobiles, pickups, vans, recreational vehicles and motorcycles) are usually counted. Buses are typically excluded or counted separately because it is difficult to count all the occupants using the roadside windshield method or the carousel method. Trucks are generally excluded because they are used mainly for goods movement and have little to do with people mobility.

Study Procedure

A typical procedure for vehicle occupancy studies include defining study objectives, selecting a data collection method, establishing a sampling plan, comparing costs, performing random sampling, and computing the average vehicle occupancy (AVO). The first step in conducting a vehicle occupancy study is to define the study objectives, which form the basis for further study planning and design. Once the objectives are defined, inappropriate collection methods can be screened out. The remaining methods can then be considered based on cost comparisons. After the data collection method is selected, the actual survey procedures and the sampling plan can be designed. Appropriate sampling plans for site-specific, corridor, and regional studies were presented and discussed in detail. Statistical sound collection techniques are a major concern and should be properly designed. A sound sampling procedure is needed to ensure that the AVO estimates meet the desired precision with a certain level of confidence. In this study, standard deviations for several factors identified by Ferlis (1981) were derived from vehicle occupancy data collected in a Florida statewide study from 1996 to 1999. In the absence of standard deviations from local data, these standard deviations are recommended for use in determining the appropriate sample sizes for corridor and area-wide studies.

Software System for Estimating Vehicle Occupancy from Crash Records

As part of this study, a user-friendly software system called FAVORITE (**F**lorida **A**ccident **V**ehicle **O**ccupancy **R**ate **I**nforma**T**ion **E**stimator) was developed to estimate the occupancy rates from multiple years of crash records on the Florida state roadway system. The system can estimate occupancy rates for select roadway segments, corridors, or regions for specific time periods for different types of vehicles. FAVORITE comes with the 1990-2002 accident data and includes passengers for up to two vehicles for each accident. In addition, the database also includes a number of variables that can be used for various analyses, including district, county, hour of day, day of week, month of year, vehicle type, facility type, area type, and crash severity. Because the system makes use of a comprehensive statewide database, it can potentially be a highly cost-effective means for monitoring statewide, regional, and site-specific vehicle occupancy trends. While a preliminary assessment of the system show outputs that are consistent with the expected data trends, an enhanced version of the system would require additional research that takes into account over- and under-involvement of certain types of accidents, e.g., young drivers and higher-occupancy vehicles are more prone to traffic accidents. Failing to correct for these factors could produce biased AVO estimates.

Field Data Collection Tool

To facilitate field data collection and processing, an automated field data collection tool designed for use with a handheld Pocket PC was developed as part of this study. The tool eliminates the need for manual data post-processing by allowing the user to make use of the touch-screen interface on a Pocket PC to record the number of passengers for different types of vehicles and different lane numbers. In addition, a companion program that can calculate the average occupancy rates from the data collected from the automated tool was also developed. The research team first investigated the possibility of applying voice recognition technology in lieu of screen input on a Pocket PC. This was found to be impractical due to the lack of an applicable commercial voice recognition system that could work well with the Windows CE operating system used by Pocket PCs. It was also found that traffic noise in the field could interfere with voice recording, resulting in erroneous data.

Potential Methods for Occupancy Data Collection

Current research into new methods for collecting vehicle occupancy has mainly been motivated by the needs for automated enforcement of high occupancy and managed lanes. These methods can be divided into photographic and in-vehicle detection. With advances in image processing and pattern recognition, a number of researchers have explored the use of photographic systems primarily for automated enforcement of HOV lanes. While photographic systems have achieved some success in counting vehicle occupants and have been shown to have some potential for further improvements, an operational, cost-effective system for occupancy data collection does not currently exist. The use of different types of in-vehicle electronics in combination with wireless communications can be a future source of occupancy data. These include systems that can detect the presence and weight of a passenger for safer deployment of air bags; in-vehicle cameras to detect passenger location and position (also for safer deployment of air bags); automated detection of seat belt usage by each passenger (already implemented for the driver); etc., all of which hold potential to provide more accurate vehicle occupancy information than any of the existing methods.

Study Guidelines for Occupancy Studies

One finding of this study was that fully automated methods of vehicle occupancy data collection are still largely a distant reality due to technological, cost, and institutional barriers. Methods with a lesser degree of automation that record and analyze occupancy data electronically remain the current methods of choice. As part of this study, a set of study guidelines was developed for the manual counting methods for corridor and area-wide studies. The guidelines address design issues related to time periods, sampling plan, field operation, work plan, personnel training, use of data collection tool and equipment, data collection plan, and data analysis.

CHAPTER 1 INTRODUCTION

1.1. Vehicle Occupancy Rates

Traditionally, vehicle occupancy rates have been used to convert person trips to vehicle trips in the four-step travel demand forecasting process and to determine the required parking spaces for fixed-seat facilities such as sporting facilities and performing centers. The Intermodal Surface Transportation Efficiency Act (ISTEA) of 1991 expanded this traditional role of vehicle occupancy rates by requiring continued monitoring of congestion management strategies, many of which emphasize person movements rather than vehicle movements. Today, traffic engineers use vehicle occupancy data to compute person delays; transportation planners use vehicle occupancy rates to derive person-miles traveled and to set policies for high-occupancy and managed lanes; transit planners use transit occupancy rates to identify routes that need service expansion; etc. More applications of vehicle occupancy data can be expected. For example, transit advocates are proposing person volumes as the basis for traffic signal warrants, transit signal priority, and transit preferential lanes.

1.2. Research Needs

With the increasing need for vehicle occupancy data comes the need to examine and reexamine the ways in which these data have been, and will be, collected. Unlike counting vehicles, which can be automatically recorded when vehicles run over pneumatic road tubes, counting the number of persons in a vehicle in the field remains largely the task of human observers. In the face of budget reductions, agencies must find better ways and define acceptable practices of collecting vehicle occupancy data that not only meet the needed accuracy, but also the limits of a restricted budget.

A somewhat limited number of studies have been found to examine methods and issues related to collection of vehicle occupancy data. The most comprehensive study to date has been performed by Heidtman *et al.* (1997) for the Federal Highway Administration. The study evaluated various data collection methods with field data and provided a relatively detailed comparison of the methods. In Florida, the Florida Department of Transportation (FDOT) undertook a statewide pilot study in 1996 and 1997 to examine alternative methods of vehicle occupancy data collection, observation locations, field procedures, treatments of commercial vehicles, etc. (Liu and Desai 1998; Reed *et al.* 1998). As part of the Florida study, over 2,000 hours of vehicle occupancy data from 21 sites covering different types of roadways throughout the state were collected.

While these prior studies have been relatively comprehensive, much remain to be done. No formal guidelines have been developed to assist users in selecting the proper methods and the associated geographic, temporal, and vehicle coverage for specific applications. In addition, it is necessary to explore the use of new technology to improve existing methods, to investigate alternative methods, and to develop tools that can ease data collection, processing, and analysis.

1.3. Objectives

This project was initiated and funded by the Florida Department of Transportation (FDOT) to accomplish the following objectives:

1. To research and evaluate existing methods of vehicle occupancy data collection with respect to their geographic, temporal, and vehicle coverage design.
2. To identify new methods as a potential source of vehicle occupancy data.
3. To present recommended practices as Florida's guidelines for the collection of vehicle occupancy data.
4. To develop tools to facilitate the collection, processing, and analysis of vehicle occupancy data.

1.4. Report Organizations

The rest of this report is organized as follows. Chapter 2 summarizes the existing methods for vehicle occupancy data collection. Existing studies on vehicle occupancy collection methods were also summarized. Chapter 3 discusses various factors that need to be considered in occupancy data collection studies. Chapter 4 presents a study procedure for vehicle occupancy collection. Chapter 5 introduces an automated system developed as part of this study to extract average vehicle occupancy rates from traffic crashes on Florida's state roadway system. Chapter 6 describes an automated tool, also developed as part of this study, for collecting vehicle occupancy data in the field. A post-processing program that can compute the average vehicle occupancy rates and the related statistics from data collected from the automated tool is also presented. Chapter 7 reviews potential methods for vehicle occupancy data collection. Based on findings from the previous chapters, Chapter 8 presents study guidelines for vehicle occupancy data collection and analysis. Finally, Chapter 9 summarizes this study and recommends further research.

CHAPTER 2

EXISTING METHODS AND STUDIES

2.1. Vehicle Occupancy Collection Methods

Existing methods for vehicle occupancy data can be grouped into those that are collected in the field for the sole purpose of computing vehicle occupancy rates and those that are collected for other purposes, but may be used to estimate vehicle occupancy rates. In general, the field collection methods are more suitable for collecting data at the site-specific or corridor level, while the existing data methods are more suitable for area-wide or regional studies. This is obviously because existing data were, by nature, collected for larger areas and with relatively small sample sizes.

2.1.1. Field Data Collection Methods

Field data collection methods are more commonly used because it can be tailored to specific application needs in terms of location, sample size and accuracy, etc. However, they are also more costly, thus sampling is a major design issue for this method. Current field collection methods include roadway windshield, carousel, and video surveillance.

2.1.1.1. Roadside Windshield Method

Roadside windshield observation is the most widely used method for collecting vehicle occupancy data. It involves stationing observer(s) along the roadside to perform physical counts of vehicles and occupants on different lanes. Data are manually recorded using paper forms or manual count boards. To reduce potential errors and the time required for data entry and analysis, electronic counter boards or laptop computers have been used. In general, this method is labor intensive and sampling methods are normally used to keep cost at a reasonable level. Because data collected by the windshield observers are often affected by such factors as human fatigue, weather conditions, amount of daylight, vehicle mix, and traffic volume and speed, the sampling procedure needs to accommodate for these factors by taking measures such as scheduling periodic breaks and possibly sampling only a portion of the traffic stream.

2.1.1.2. Carousel Method

The carousel method positions observers in vehicles traveling on multi-lane highways to collect vehicle occupancy data on neighboring vehicles. During data collection, the observer vehicle drives slightly slower than the general traffic, resulting in the continuous flow of traffic by the observation vehicle. The observer may use an electronic counter or a laptop computer to record the vehicle occupancy data. The observer vehicle begins a cycle traveling in one direction along the survey route, then turn around to drive the same route in the opposite direction to the beginning point on the roadway segment before another run is started. The average traffic volumes and speeds along the selected roadway segment are examined in each study to calculate the number of observers and observer vehicles needed to collect an adequate sample size.

With the carousel method, although fewer vehicles are actually surveyed (about one-fourth of those observed by the windshield method), the ability to discern the number of persons including small children in each passing vehicle is greater because the survey vehicle travels along with the vehicles under observation. Obviously, more observed vehicles can be captured by deploying more survey vehicles. However, this may require careful planning and coordination among the survey vehicles to avoid double counting of vehicles.

2.1.1.3. Video Surveillance Method

This method uses video cameras mounted on overpasses or along the side of a roadway to capture passing vehicles and then uses a human observer to view the capture videos to extract occupancy data. Because this method allows the observers the needed time to make a better judgment about occupancy, the counts can be more accurate than the direct field observation methods. In addition, video segments skipped for the purpose of sampling or fast-forwarded to skip large vehicle gaps can provide significant time savings especially for low-volume roadways. The video surveillance method is preferred when it is impractical to station a field observer during the entire exercise because of physical conditions, or there is a need for a large amount of data to be collected continuously. The cost from equipment purchase and setup can be justified if the existing video cameras installed for traffic surveillance and management purposes allow the collection of multiple data types from the monitoring period. Other than the high initial implementation cost, this method has some limitations. The windshield glare from observed vehicles or vehicles with tinted windows could prevent vehicle occupants from being accurately counted. There may also be difficulty recording data in poor weather conditions or during darkness.

2.1.1.4. Computer Vision Method

This method applies computer vision techniques to automatically recognize people in a vehicle from captured images. It offers a potential solution to the time-consuming task of viewing video images to count vehicle occupancy. Although a few researchers have attempted this method, an operational system is not currently available. More information related to current advances in this method is presented in Chapter 7 of this report.

2.1.2. Existing Data Methods

The accuracy and scope of applications of vehicle occupancy data extracted from the existing data are constrained to what have been collected. Obviously, the sampling method and size are not a design issue with existing data. The objective of the users in this case is mainly to extract the maximum amount of information at the highest level of details possible. In addition, the accuracy of estimation that can be achieved with the existing data needs to be determined. The following subsections introduce several sources of data that can be used to extract vehicle occupancy information.

2.1.2.1. Crash Records

This method extracts vehicle occupancy estimates from police crash records for a particular road

segment, corridor, or metropolitan area for specified time periods. Crash reports typically record information on type of vehicle involved, number of occupants in each vehicle, and time and location of the crash. This makes it possible to compute average vehicle occupancy (AVO) that is stratified by these variables. Also, the continuous nature of crash records allows analysis of trends on vehicle occupancy. However, adjustments may be needed to account for over-involvement of certain subpopulations in terms of driver gender or age in crashes. In addition, delays in compiling crash records prevent the method from providing up-to-date AVO. An automated system developed as part of this study to extract average vehicle occupancy rates from traffic crashes on Florida's state roadway system is presented in Chapter 5.

2.1.2.2. Household Survey

Household survey is conducted by randomly sampling on the general resident population of the urban transportation study area. Comprehensive information is obtained regarding socioeconomic data, trip purposes, and modes of travel. Household surveys are generally conducted through telephone interview, personal interview, or questionnaire mail-out/mail-back procedures. A common type of household survey is origin-destination (O-D) survey, which collects not only trip origins and destinations, but also vehicle occupancy.

2.1.2.3. Vehicle Intercept Survey

Vehicle intercept survey collects origin-destination and basic characteristics of the trips based on roadside interview or postal questionnaire. The vehicles intercepted for survey can be randomly or systematically selected. Roadside interview surveys can be unsafe and expensive to conduct in many situations. Where practical, however, interview surveys will yield higher and less biased response rates for a set of limited data items. On facilities with high volume and where traffic cannot be stopped long enough for an interview, postal questionnaire can be used instead. Typically, pre-paid mail-back questionnaires are distributed to drivers in vehicle stopped at traffic signals or toll booths to reduce delay to traffic. However, the typical low response rates (20-35%) of this method may produce biased estimates (Pietrzyk, 1996).

2.1.2.4. Census Transportation Planning Package (CTPP)

Census Transportation Planning Package (CTPP) is a set of comprehensive data extracted from the decennial census. The journey-to-work data collected as part of CTPP reports travel behavior and characteristics for work trips throughout the country. These data are compiled and reported for every Metropolitan Statistical Area (MSA) in the United States. It may serve as a good source of data for examining the relationship between vehicle occupancy and demographic characteristics such as the number of hours worked a week, sex, age, income, and vehicle available. Because CTPP is available only on a decennial basis, the data do not allow annual monitoring of trends. In addition, census data provide vehicle occupancy information only for journey-to-work trips. With the relative decline of home-based work trips, journey-to-work data may not be adequate to measure system performance.

2.1.2.5. Nationwide Personal Transportation Survey (NPTS)

Nationwide Personal Transportation Survey (NPTS) is a national household survey that is based on a random telephone survey designed to collect household information on the daily trips made by household members. NPTS can provide vehicle occupancy information by trip purpose, household characteristics, and other factors. However, it cannot provide information on vehicle occupancy by type of highway or time of day.

2.2. Transit Occupancy

Transit occupancy data is difficult to collect with either the windshield or the carousal method. Also, traffic crash data do not usually contain a sufficient number of bus crashes. An approximation method that has been used is to determine if a bus is empty (only the driver), one-quarter full (or about 10 occupants), one-half full (or about 20 occupants), and full (or about 40 occupants). Variation on the associated passenger estimates can be made depending on typical bus sizes in the study area (Levine and Wachs, 1994). Existing data sources for estimating transit occupancy data include:

1. *Manual Checker or Automatic Passenger Counter (APC) Data*: Transit agencies regularly use human checker to count passengers on select transit vehicles on a sampling basis to estimate passenger trips and passenger miles. More recently, APC technology has been deployed to count passengers automatically as they board transit vehicles. Vehicle occupancy rates (i.e., passenger loads) may be obtained at any particular point on a transit route from these data. Currently, the Jacksonville Transportation Authority (JTA) is the only agency known to have some of its buses equipped with APCs; however, most transit agencies are expected to eventually have APCs on all of their buses.
2. *National Transit Database (NTD)*: NTD is the most comprehensive database collected by and for the transit industry and can be used to estimate only the system-wide transit occupancy rate.

2.3. Existing Studies

Various studies have been conducted to collect vehicle occupancy data for different purposes. For example, Rutherford *et al.* (1990) reviewed the methods used by agencies for short-term and long-term monitoring of high-occupancy vehicle (HOV) lane violations. Levine and Wachs (1994) presented a methodology for conducting regional and corridor-level vehicle occupancy surveys. Heidtman *et al.* (1997) compared various vehicle occupancy data collection methodologies based on field studies and cost consideration. A number of concerns and issues reported in these studies include:

- Observation locations and days were chosen arbitrarily, rather than sampled randomly in order to represent vehicle occupancy accurately for a particular purpose.
- Locations with higher traffic volume are more likely to be chosen to the exclusion of lower volume locations.

- Sites near central business districts (CBDs) or major trip attraction areas where congestion is assumed to be highest were also chosen frequently.
- Many agencies appear to acquire AVO estimates using the simple mean calculation procedure. Vehicle occupancy observations taken under these conditions will create biased estimates and in turn unreliable and inaccurate findings would be drawn.

Table 2-1 provides a chronological summary of studies conducted in terms of study objectives, study methodology, and use of temporal and spatial coverages. It can be concluded from this summary that:

- Roadside windshield is the most commonly used method.
- There is a trend to save data entry time by switching recording from manual count boards to laptop computers or electronic counter boards.
- Measurements were usually taken in AM peak period, or with PM peak period.
- Observers generally counted the short periods rather than continuously counting within one hour.
- Observation were typically taken in the middle of the week (Tuesdays, Wednesdays, and Thursdays) to avoid unusual conditions of recreational travel prevailing on Mondays, Fridays, and weekends.

Table 2-1. Summary of Vehicle Occupancy Collection Practices in Different Study Areas

City/State	Sponsoring Agency	Report Year	Study Purpose	Methodology Used	Spatial Coverage	Temporal Coverage
Detroit, MI	Southeast Michigan Council of Governments	1980	To test Ferlis' sampling procedure	Roadside windshield observation with random sampling	Area-wide: - 69 counting sessions stratified by geographic area, highway class	- 15-minute count with 5-minute break - 2-hour AM peak, 2-hour midday, and 4-hour PM peak - Weekday - The first three weeks of May
Atlanta, GA	Georgia DOT	1980	To test Ferlis' sampling procedure	Roadside windshield observation with random sampling	Area-wide: - 64 counting sessions stratified by geographic area type and highway class Corridor: - 22 locations at freeways and arterials crossing the railroad cordon - 5 locations along the highway adjacent to proposed park-and-ride lots	- 45-minute count with 15-minute break - 12-hour daylight period - April and May
Washington, D.C.	FHWA Office of Research and Development	1981	To test methodology to measure vehicle occupancy by prototype camera system	Photographic Surveillance	Corridor: - Two HOV lanes and a ramp on Interstate Highway	- times of day - 9 days - Different weather condition
Portland	Oregon DOT	1982	To determine the effectiveness of HOV lanes	Roadside windshield observation with traffic counter board	Site-specific: - Location on freeway with HOV lanes	- 10 minutes in the peak direction and 5 minutes in non-peak direction in each hour - 3-hour peak period - Tuesday through Thursday - During 2nd week of the month

Table 2-1. Summary of Vehicle Occupancy Collection Practices in Different Study Areas (Cont.)

City/State	Sponsoring Agency	Report Year	Study Purpose	Methodology Used	Spatial Coverage	Temporal Coverage
George Washington Bridge, NJ	New Jersey DOT	1987	To evaluate the impact of expansion the bus-only lane into a longer bus-carpool lane	Roadside windshield observation with traffic counter	Site-specific: - 3 toll plazas on the bridge with HOV lane	- 20-minute period each half-hour - 3-hour AM peak - Tuesday through Thursday - During 3rd week of the month - Once a month
New York City, NY	New York State DOT	1988	To determine impacts of proposed development scenarios for a route reconstruction	Mail-back questionnaire and associated roadside windshield observation	Corridor: - Cordon line sounding the impact zones	- 2-hour AM peak - Monday, Tuesday, and Thursday - November
Puget Sound Area, WA	Washington State DOT	1988	To test methodology for long-range VOC program planning	Roadside windshield observation	Area-wide: - 10 locations including employment sites, arterials, and freeways.	- 15-minute periods with 15-minute breaks and 30-minute periods with 15-minute breaks - 3-hour AM, 3-hour PM peak, and 3-hour evening - Weekdays and weekends
Puget Sound Area, WA	Washington State DOT for Puget Sound Council of Governments	1989	To update travel demand model for vehicle occupancy program planning	Roadside windshield observation with random sampling	Area-wide: - 54 sampling sites	- AM and PM peak - One year period
New York City and surrounding areas	New York Metropolitan Transportation Authority	1989	To support transit agency planning	Telephone survey of random, stratified sample on travel pattern	Area-wide: - New York City and surrounding areas	- Over 24-hour day - Spring
Phoenix, AZ	Maricopa Association of Governments	1989	To calibrate the shared-ride component of the mode choice model	Roadside windshield observation and associated Mail-back questionnaire with random sampling	Area-wide: - 36 locations stratified by area type, facility type - 33 parking lots and garages	- 2-hour AM, 5-hour midday, 2-hour PM peak, 1-hour evening - Weekdays - February and March

Table 2-1. Summary of Vehicle Occupancy Collection Practices in Different Study Areas (Cont.)

City/State	Sponsoring Agency	Report Year	Study Purpose	Methodology Used	Spatial Coverage	Temporal Coverage
Oklahoma, Canadian, and Cleveland Counties	Association of Oklahoma Governments	1991	To update survey methodology to better meet planning needs including CAAA requirements	Roadside windshield observation with clustering technique	Area-wide: - 20 stations of different facility types in three counties	- 1-hour AM peak, 1-hour midday, and 1-hour PM peak - Monday, Tuesday, and Thursday - Two seasons - Two years in a row
Elmira, NY	Executive Transportation committee for Chemung County	1991	To support general planning and development strategies	O-D survey and pullover roadside interview	Corridor: - 13 point-of-entry sites at high volume roads bisecting cordon line	- AM peak - six weekdays - two-week period
Providence, RI	Rhode island DOT Division of Planning	1991	To study traffic patterns to alleviate congestion on an Interstate Highway	License plate sample and mail-back postcard survey of sample	Corridor: - two high volume facilities bisecting the corridor	- AM peak - Weekday
Selected counties in New Jersey and Pennsylvania	Delaware Valley Regional Planning Commission	1992	To design statistically valid telephone survey method to measure vehicle occupancy	Telephone interview with random sampling	Area-wide: - 4,800 interview across 13 counties	
Minnesota	Minnesota DOT	1993	- To conduct an annual Metropolitan Occupancy Data Collection program - To monitor vehicle occupancy on both the HOV and mix-flow lanes	Direct observation field counts from overpasses and elevated barricade positions	Statewide: - 18 count locations stratified by facility type and area type Corridor: - 6 count locations along an Interstate highway with HOV lane	- 3-hour AM and 3-hour PM peak - Tuesday through Thursday - Twice a year during the spring (April and May) and fall (September and October)
Dallas-Fort Worth, TX	Texas DOT	1993			Area-wide: - 26 locations stratified by area type	- once a year

Table 2-1. Summary of Vehicle Occupancy Collection Practices in Different Study Areas (Cont.)

City/State	Sponsoring Agency	Report Year	Study Purpose	Methodology Used	Spatial Coverage	Temporal Coverage
Southeastern area, VA	Virginia DOT	1993	To regularly monitor vehicle occupancy	Roadside windshield observation with intersection turning movement recorder	2 locations on a HOV system	- 3-hour AM and 3-hour PM peak - four times a year
Washington, D.C.	Washington, D.C. Council of Government	1994		Roadside windshield observation with laptop computers	Corridor: - Freeways and principal arterial highways bisecting cordon lines	- Spring and fall
Nationwide	U.S. DOT National Highway Traffic safety Administration	1994	To propose method to measure level of seatbelt use nationwide	Roadside windshield observation with clustering and stratification techniques	Nationwide: - 50 primary sampling units - 500 tracts - 4,000 sites	To be probability determined
New York	New York State DOT	1996	To investigate the capacity of traffic accident database to predict AVOs	Traffic accident records	Statewide - Accident on interstates	- Can be by hour of the day, week, season, or year - 3-years period
Florida	Florida DOT	1997	To determine efficient procedures for the collection, analysis and use of vehicle occupancy data	Roadside windshield observation	Statewide: - 21 sites stratified by geographic area and facility type	- 12-hour daylight period - Five consecutive weekdays - Two weekdays in each month for twelve consecutive months
Connecticut	Connecticut Public Transportation Commission	1998	To monitor the efficiency of roadway system	Traffic Accident Database	Statewide: - Accidents on interstates	Can be by hour of the day, week, season, or year
Kansas City, MO	Mid-America Regional Council	2002	To monitor impacts of increasing occupancy and compare the results with assumptions used in the long-range transportation planning process	Roadside windshield observation	Regional: - 44 sites stratified by geographic area (inner and outer urbanized areas) and four functional classifications of roadways	- 1.5-hour AM peak and 2-hour PM peak - Tuesday, Wednesday, and Thursday

CHAPTER 3

STUDY DESIGN CONSIDERATIONS

This chapter reviews several design considerations for vehicle occupancy data collection, including geographic and temporal coverages, facility types, collection cycles, locations, and vehicle types and percentages.

3.1. Geographic Coverage

3.1.1. Geographic Units

Geographic units used to evaluate vehicle occupancy may include the entire state, air quality non-attainment area, urban area, subarea, region, corridor, activity center, functional class, and external cordon line (Heidtman *et al.* 1997). Depending on the specific study objectives, one or more of these units may be appropriate. The traditional interest has been to estimate vehicle occupancy at either site-specific (e.g., factory entrance and critical roadway section) or area-wide (e.g., metropolitan and county) level. The more recent focus, however, has been on the intermediate level of geographic coverage, which is referred to as the corridor level, e.g., at cutline and CBD cordon line.

3.1.2. Geographic Density

Geographic density can be minimized to reduce cost if clusters of geographic features can be identified. Spatial analysis using a GIS can help identify clusters and detect changes in vehicle occupancy rates across geographic features. Due to its comprehensive nature, crash records may be used to evaluate spatial changes in vehicle occupancy rates.

3.2. Temporal Coverage

Temporal variability in AVO is a common issue across all data collection methods. Existing studies show significant variations in vehicle occupancy rates by time-of-day, day-of-week, and season-of-year.

3.2.1. Time-of-Day

The AM peak period tends to be heavily dominated by home-based work trips, primarily with single-occupant vehicles. As the day progresses, increasing variety of other trip purposes occur which are likely to shift vehicle occupancy upward. In the PM peak period, vehicle occupancy may be higher than the AM peak period due to a multitude of different trip purposes, many of which involve more than one person.

3.2.2. Day-of-Week

Vehicle occupancy levels vary from day to day, with Saturday typically being higher. However, studies did not find a consistent day-of-week pattern for variations in AVO. In general, it is a common practice to exclude Monday, Friday, and weekend counts because these days are

assumed to contain atypical and non-recurrent trips. Measurements were typically taken on Tuesday, Wednesday, and Thursday only, as these days are considered most representative of average weekday travel behavior and commute conditions.

3.2.3. Month/Season-of-Year

Vehicle occupancy rates have been found to vary with month and season. It is thus important that observations be taken throughout the year (and not restricted to any one season) in order to properly represent average vehicle occupancy (Heidtman *et al.*, 1997).

3.2.4. Temporal Trends

The following temporal trends in AVO have been observed in various studies (Roach and Lester, 1978; URS Corporation, 1997):

- Weekday AM peak AVO is normally lower than weekday midday and PM peak AVOs.
- Weekend AVO is normally higher than weekday AVO.
- Off-peak AVO is normally higher than AM and PM peak AVOs.
- Summer AVO is higher than winter AVO for the northern cities.

Based on data from Florida, Liu and Desai (1998) made the following general recommendations with regards to temporal coverage:

- The data collection period should desirably be within the period of interest (e.g., peak hours, etc.).
- Counts of one to two hours would produce data with sufficient accuracy and precision for most purposes.
- Tuesdays and Wednesdays are normally adequate for data collection and the best days for collecting data are Thursdays.
- Mondays and Fridays should be avoided.
- Time of day chosen for data collection is important.
- As a rule of thumb, mid-morning to mid-afternoon counts are adequate for most purposes. However, if the 5-6 pm period is chosen for performance monitoring, adjustments would be necessary to derive the AVO for the day.

3.3. Facility Types

Different types of roads typically have different occupancy levels. Roads of the higher functional hierarchy would typically be expected to have lower AVO. Freeways and major arterials are heavily used for home-based work trips and commercial trips, particularly in the AM and PM peak periods. These two trip purposes tend to have lower AVO. Conversely, minor arterials, collectors, and local streets are utilized more extensively for trips of home-based shop, home-based school, and social/recreational, resulting in higher AVO. In addition, freeway segments with high occupancy vehicle (HOV) lanes will have higher AVO than mix-flow freeway segments. However, these phenomena are subject to large variations in different study

areas. It is thus important to sample all roadway types in order to generate a representative estimate of regional vehicle occupancy for an area-wide estimate of AVO.

3.4. Data Collection Cycles

A literature search did not reveal specific practices or guidelines for data collection cycles. It is, however, an important design issue that needs to be addressed as it has a major impact on the cost of data collection. Given the fact that vehicle occupancy data are rather stable and that the cost of collecting them are currently high, a collection cycle of one year or less would generally be both unnecessary and unfeasible. Especially for an area-wide level, it is not necessary to collect data every year but to establish it with sufficient precision to produce a reliable indicator of regional road efficiency. An initial area-wide study should hence be conducted to establish a statistical sound baseline. Once the initial study has been conducted, area-wide studies can be conducted every couple of years to monitor trends in vehicle occupancy. However, cycles greater than five years should be avoided if the data are to be included in the FDOT's Roadway Characteristics Inventory (RCI) database, which has a maximum collection cycle of five years.

3.5. Data Collection Locations

AVO can differ from one location to another. The spatial variations of AVO are related to the distribution of household types and work places. It is necessary to select many different locations in order to measure variability adequately. Site selection for traffic data collection is dependent on the purposes and the overall expectation of the study (Liu and Desai, 1998). To ensure that a representative sample of the population will be selected, data collection locations should be sampled randomly from all possible highway segments for area-wide studies. If an initial study has been performed, the entire random selection procedure should be repeated again for subsequent studies. The same sampled locations should not be chosen unless they are randomly selected. In contrast to the randomly selected locations for area-wide studies, locations are judgmentally selected due to the nature of corridor/site-specific studies. Therefore, it would be appropriate to use the same locations for every collection cycle. This ensures that the data can be used for continued monitoring, that any changes in the data trend may be attributed to factors other than location. This is consistent with the state's practice for telemetry and portable traffic monitoring sites, which uses the same locations from year to year.

3.6. Vehicle Types

The vehicle types included in a data collection are determined by the purpose for which the data are to be collected. Different study purposes may utilize different criteria for interpreting AVO. If the purpose is to evaluate the overall efficiency of a road system, then all vehicles should be counted regardless of vehicle types in order to assess the entire load on the road system. On the other hand, if the purpose is to measure the impact of a ridesharing program, for example, may be only passenger vehicles should be included because they are most sensitive to ridesharing incentives.

In most vehicle occupancy studies, only data from passenger vehicles or light vehicles (private passenger automobiles, pickups, vans, recreational vehicles and motorcycles) are usually

counted. Buses are typically excluded or counted separately because it is difficult to count all the occupants using the roadside windshield method and carousel method. Trucks are generally excluded because they are used mainly for goods movement and have little to do with people mobility.

The definition of vehicle type remains an issue in vehicle occupancy data collection. A simple breakdown of passenger vehicles, buses, and trucks are generally sufficient for vehicle occupancy studies. However, there is no consistency in definition across different regions or different transportation models. Therefore, the vehicle type definition should be determined by the data collection objective and purpose, as well as the consistency with existing traditional transportation model requirements. For example, if the vehicle occupancy data were used to validate the modal split in FSTUMS, the vehicle type definition of vehicle occupancy data should be consistent with the vehicle type definition used in FSTUMS.

3.7. Vehicle Percentages

Instead of recording data from all vehicles traversing a link during a selected time period, a subset of vehicles that are assumed to reflect the same characteristics as the target population is generally observed. Systematic short-count procedure, in which observations are made for a fixed interval in each hour of the day, can be used to produce relatively accurate daily estimates while conserving manpower resources. The following three basic types of short-count procedures were suggested by Ferlis (1981):

- Position one or more observers to count all vehicles that pass by during a fixed interval within each hour (e.g., count for 45 minutes and rest for 15 minutes, thus representing a 75 percent systematic sample).
- Position one observer to count vehicles that pass by on each lane during a fixed interval within each hour (e.g., count each of three lanes during successive 15-minute periods and rest for 15 minutes within each hour, thus representing a 25 percent systematic sample).
- Position one or more observers to systematically observe two or more locations concurrently by counting all vehicles passing a particular location during the same time interval within each hour (e.g., count vehicles at one location from 7:00 to 7:15, 8:00 to 8:15, etc., and at another location from 7:30 to 7:45, 8:30 to 8:45, etc., thus representing a 25 percent systematic sample).

CHAPTER 4 STUDY PROCEDURE

This chapter describes the procedure for conducting a vehicle occupancy study. A significant portion of the materials presented was taken from Ferlis (1981) and Heidtman *et al.* (1997). The procedure include defining study objectives, selecting a data collection method, establishing a sampling plan, comparing costs, performing random sampling, and computing the average vehicle occupancy (AVO).

4.1. Defining Objectives and Selecting Data Collection Method

The first step in conducting a vehicle occupancy study is to define the study objectives, which form the basis for further study planning and design. Once the objectives are defined, inappropriate collection methods can be screened out. The remaining methods can then be considered based on cost comparisons. After the data collection method is selected, the actual survey procedures and the sampling plan can be designed. A sample of objectives for vehicle occupancy estimates may include:

- Evaluate the effectiveness of the transportation system management (TSM) program.
- Verify compliance with state regulation.
- Analyze transportation related air quality and energy efficiency measures.
- Validate the adequacy of urban transportation planning models.
- Assess the impact of a new transportation system.
- Monitor the general trends in travel characteristics.
- Formulate transportation strategies.

Based on the desired objective, the target collection population can be defined in terms of geographic scope, temporal coverage, facility types, etc. Stratification of collection population may also be needed to provide estimates at a finer level. For example, if the estimates of AVO are desired for the central business district (CBD) and suburban area as well as for the entire region, a stratified sampling plan will then be required. Depending on the geographic and temporal coverages for data collection, collection methods that are not suited for the intended collection design can be discarded from the list. The general applicability in terms of geographic scope for different occupancy data collection methods are summarized in Table 4-1. The level of precision is viewed as an acceptable range of error about the mean AVO estimated from the data collection. The minimum sample size and collection cost will be directly affected by the need for AVO estimates at predetermined levels of precision.

Table 4-1. Applicability of Occupancy Data Collection Methods by Geographic Scope

Method	Geographic Scope		
	Area-Wide	Corridor	Site-Specific
Roadside Windshield	√	√	√
Carousel Observation	√	√	
Video Surveillance		√	√
Mail-Out Questionnaires	√	√	
Crash Records	√	√	

4.2. Establishing Sampling Plan

After the objectives are determined and the method of data collection selected, a proper sampling procedure can be used to construct a sampling plan, which applies the sampling theory to determine the sample size needed to obtain the desired level of precision in the AVO estimates. The sample size required to attain a certain level of precision depends on the degree of variation in AVO measures, which is usually measured in standard deviation. For example, a common approach is to set the tolerance level for an estimate to be within ± 0.03 of the true AVO with 95 percent confidence, i.e., there is a 95 percent likelihood that the estimated AVO will fall within a range of ± 0.03 from the true population average. Tolerance is defined as the acceptable difference between the estimated AVO and the true AVO, while the level of confidence represents the probability that the sample estimate will fall within this range. Thus, a more precise estimate will require a higher sample size.

Once the sample size is determined, the sample units are randomly selected from a frame where all possible individuals are listed. The random sampling of sample unit will result in an unbiased estimate of AVO. Different sampling procedures are required depending on whether the AVO estimates are site-specific, corridor, or area-wide. Multiple objectives requiring different levels of AVO measures are often of interest to transportation agencies. The choice of the area-wide vehicle occupancy study is recommended to achieve the multiple objectives through a single collection effort. In addition, later uses of occupancy data for different purposes when selecting a data collection method should be considered.

4.2.1. Field Observational Methods

The sampling process that is conventionally used in the field observational methods includes simple collection and stratified collection. Simple collection involves selection of sample units at random from the entire population so that each sample unit has an equal probability of being selected. Stratified collection, on the other hand, begins by dividing the entire sample frame into mutually exclusive strata, then selecting sample units from each stratum using simple collection. The following sections begin with the introduction of key factors related to the computation of variance in AVO. The estimation of minimum sample size on the basis of the derived variance for two types of collection methods can then be established.

4.2.1.1. Major Factors Affecting AVO

A number of factors described in Chapter 3 could affect the variation in the AVO estimates. Due to limited resources available for this study, it is not possible to examine the effect of each of these actors. However, it is possible to examine major factors that are believed to have a major influence on vehicle occupancy. The sampling scheme developed by Ferlis (1981) is to estimate the number of sampling units on the basis of spatial and temporal variation and the precision required. To account for the impact of spatial and temporal factors upon the accuracy of AVO estimates, Ferlis (1981) suggested the following primary sources of variation:

- The variation across link-days within a season (σ_L)
- The variation from day to day within a season (σ_D)

- The variation from season to season (σ_S)
- The variation from time interval to time interval during a day as a result of short-counts (σ_W)

Additional sources of variations have been suggested by other studies. Ulberg and McCormack (1988) examined some potential sources of error and concluded that observer counting error (σ_O) should be included. In addition, the degree of variation in AVO derived in Ferlis' study may not be reflective of the traffic characteristics today. With a relatively small variance derived from his study, the required sample size would be too small to adequately represent the regional total. The variability of AVO obtained from local data should be used instead.

A study was conducted in 1997 by URS Corporation for the Florida Department of Transportation to collect vehicle occupancy data from 21 individual sites throughout the State of Florida. These sites were selected for roadways ranging from two-lane rural routes to four- and six-lane urban arterial routes, including interstate highways. Occupancy data were collected for two directions from 17 conventional sites for twelve hours during daytime on five consecutive weekdays. Accordingly, variations in occupancy data of private vehicles due to time periods during a day resulting from the use of short-counts, day of week, and month of year could be established. Four seasonal sites were observed twice each month over a one-year period to determine the monthly and seasonal variation in vehicle occupancy.

Vehicles were classified as passenger cars, vans, trucks, and buses. Vans and buses were not counted due to difficulty in getting an accurate passenger count of these vehicles, especially at high speed. Since a very high proportion of the vehicles were passenger cars, variation in vehicle occupancy attributed to other vehicle types was relatively insignificant.

Table 4-2 shows the results computed with local data for different sources of variation along with the suggested values from studies by Ferlis (1981) and Ulberg and McCormack (1988). The table shows that the standard deviations observed in the data from the URS study is generally larger than those suggested by Ferlis. It should be noted that the values for the within-day term, σ_W , are based on an assumed systematic sampling rate of 25 percent.

Table 4-2. Sources of Variation in Average Occupancy

Geographic Scope	Standard Variation	Ferlis (1981)	Ulberg and McCormack (1988)	Data from URS Study (1997)
Area-wide	σ_L	0.063	0.006	0.076
	σ_S	0.015		0.068
	σ_W	0.017		0.008
	σ_O			
Site-specific or Corridor	σ_D	0.015	0.006	0.028
	σ_S	0.015		0.068
	σ_W	0.017		0.008
	σ_O			

The formulas and computational procedures for the different standard deviations in Table 4-2 are given below:

$$\sigma_L = \left[\frac{N_i \times \sum_i (P_i - AVO_L \times V_i)^2}{\left(\sum_i V_i \right)^2} \right]^{1/2}$$

$$AVO_L = \frac{\sum_i P_i}{\sum_i V_i}$$

where N_i = number of locations,
 P_i = number of persons counted at location i , and
 V_i = number of passenger vehicles counted at location i .

$$\sigma_S = \left[\frac{N_k \times \sum_k (P_k - AVO_S \times V_k)^2}{\left(\sum_k V_k \right)^2} \right]^{1/2}$$

$$AVO_S = \frac{\sum_k P_k}{\sum_k V_k}$$

or

$$\sigma_S = \frac{MAX(P_k/V_k) - Min(P_k/V_k)}{4}$$

where N_k = number of seasons,
 P_k = number of persons counted during season k , and
 V_k = number of passenger vehicles counted during season k .

$$\sigma_w = \left[\frac{N_m \times \sum_m (P_m - AVO_w \times V_m)^2}{\left(\sum_m V_m \right)^2} \right]^{1/2}$$

$$AVO_w = \frac{\sum_m P_m}{\sum_m V_m}$$

where N_m = number of possible time intervals within each hour,
 P_m = number of persons counted during time interval m , and
 V_m = number of passenger vehicles counted during time interval m .

$$\sigma_D = \left[\frac{N_d \times \sum_d (P_d - AVO_D \times V_d)^2}{\left(\sum_d V_d \right)^2} \right]^{1/2}$$

$$AVO_D = \frac{\sum_d P_d}{\sum_d V_d}$$

where N_d = number of days,
 P_d = number of persons counted on day d , and
 V_d = number of passenger vehicles counted on day d .

4.2.1.2. Simple Collection

Simple collection can meet a wide range of needs for the area-wide and site-specific occupancy studies. The composite standard deviation of AVO should be estimated before the minimum sample size can be computed. Following the discussion in the previous section, Ferlis' framework is modified by defining the composite standard deviation as:

Area-wide Level: $\sigma = \left(\sigma_L^2 + \sigma_S^2 + \sigma_w^2 + \sigma_O^2 \right)^{1/2}$

Site-specific Level: $\sigma = \left(\sigma_D^2 + \sigma_S^2 + \sigma_w^2 + \sigma_O^2 \right)^{1/2}$

where σ = composite standard deviation of AVO,
 σ_L = standard deviation of AVO across link-days within a season,

σ_D = standard deviation of AVO across days within a season,
 σ_S = standard deviation of AVO across seasons,
 σ_W = standard deviation of AVO across time periods within a day, and
 σ_O = standard deviation of AVO due to observer error.

Depending on the sampling procedure, some of these sources of variance may not apply. The link-day variation, σ_L , reflects the fact that different AVO measurements can vary by measured locations as well as by different days within a season. For an area-wide study, σ_L accounts for a much greater variability than the other terms. The daily variation term, σ_D , should be included for corridor and site-specific studies because it can vary substantially from one day to the next. The seasonal variation, σ_S , should be included if the estimated occupancy is intended to represent more than one season and the data collection is therefore extended throughout this period. Similarly, the within-day variation, σ_W , should be included only if a short-count method is used.

For the first year, the results of a previous vehicle occupancy survey can be used to estimate the individual standard deviation terms by using the formulas described above. If agencies have not conducted prior studies or the standard deviations is difficult to estimate before the initial study is conducted, the values derived in this study as shown in Table 4-2 may be used. The default composite standard deviation for area-wide studies is 0.102. For corridor and site-specific studies, the default value is 0.074. For subsequent years, the composite standard deviation estimated from the first-year study can be used directly.

The basic sampling unit is a “link-day”, representing an estimate of the survey measures made for a particular link on a particular day for a specified time period. After the composite standard deviation is estimated, the number of link-days required to reliably estimate AVOs within a desired tolerance and confidence level can be computed as follows:

$$N = \left(\frac{Z \times \sigma}{T} \right)^2$$

where N = number of link-days to be sampled,
 Z = normal variant for the specific level of confidence, and
 T = desired tolerance.

The Z values are 1.645 and 1.960 for 90% and 95% confidence levels, respectively. At the site-specific level, the sample size indicates the number of days of data collection due to one location/link. Every location must be sampled on at least one day.

4.2.1.3. Stratified Collection

Agencies often place emphasis on collecting separate AVO estimates for various subsets of the area-wide transportation network (e.g., separate estimates for individual counties, freeways and arterials, and freeways by HOV and mix-flow lanes). In this case, the collection population is stratified for separate sampling. The set of links within each stratum is considered a separate population. On the other hand, a corridor-level study can also be treated with stratified sampling plans in which each location represents a different stratum. The minimum sample size of link-

days needed to estimate AVO for each stratum can then be computed. A minimum sample size of one day per location is also required. Once the sample size for each stratum is known, the level of precision that the sampling is to achieve can be directed toward the level of each stratum and the total estimate from all strata. The samples of link-days are then independently drawn at random from each stratum population.

In contrast to the simple estimating method, the composite standard deviation is estimated for each stratum as:

Area-wide Level:
$$\sigma_h = \left(\sigma_{Lh}^2 + \sigma_{Sh}^2 + \sigma_{Wh}^2 + \sigma_{Oh}^2 \right)^{1/2}$$

Corridor Level:
$$\sigma_h = \left(\sigma_{Dh}^2 + \sigma_{Sh}^2 + \sigma_{Wh}^2 + \sigma_{Oh}^2 \right)^{1/2}$$

where σ_h = composite standard deviation of AVO in stratum h ,
 σ_{Lh} = standard deviation of AVO across link-days within a season in stratum h ,
 σ_{Dh} = standard deviation of AVO across days within a season in stratum h ,
 σ_{Sh} = standard deviation of AVO across seasons in stratum h ,
 σ_{Wh} = standard deviation of AVO across time periods within a day in stratum h , and
 σ_{Oh} = standard deviation of AVO due to observer error in stratum h .

The definition and logic of these variations are equivalent to those described for the simple occupancy study except they apply to individual strata. The results of a prior study should preferably be used for the initial study. Otherwise, the composite standard deviation can be judgmentally estimated from the default values given in Table 4-2. The composite standard deviation computed from the first year's survey can then be used in a later year.

The sample size of link-days required to estimate stratum AVO within a desired tolerance can be computed as:

$$N_h = \left(\frac{Z \times \sigma_h}{T_h} \right)^2$$

where N_h = number of link-days required for stratum h , and
 T_h = desired tolerance for stratum h .

Alternatively, if the objective of a corridor study is to estimate the AVO across all locations within a tolerance rather than to obtain precise estimates for each location, the total sample size of days for all locations and the sample size for each location can be computed as follows:

$$N = \left[\frac{Z \times \left(\sum_h W_h \times \sigma_h \right)}{T} \right]^2$$

and

$$N_h = N \times \frac{W_h \times \sigma_h}{\sum_h W_h \times \sigma_h}$$

where W_h is the proportion of total traffic volumes across all locations occurring in location h .

4.2.2. Mail-Out Questionnaires

For a mail-out survey, questionnaires are mailed to a particular subpopulation to inquire the residents' typical trips about the number of occupants traveling with them to or from work. Mail-out questionnaires are most often used for area-wide studies. The sampling of the region often involves stratifying the population into strata (e.g., zip code area), and to send out questionnaires in proportion to the driver population in each stratum. The sampling unit is the returned questionnaire. The observation is the number of persons that travel with the respondent during the day on the trips described. With prior estimates of the standard deviation in vehicle occupancy for each stratum, the number of returned questionnaires needed to estimate the regional AVO within a desired tolerance can be computed as follows:

$$N_h = \frac{Z^2 \times D \times \left(\sum_h D_h \times \sigma_h^2 \right)}{D^2 \times T_h^2 + Z^2 \times \left(\sum_h D_h \times \sigma_h^2 \right)}$$

where N_h = number of returned questionnaires required for stratum h ,
 Z = normal variant for the specific level of confidence,
 D = total driver population in the region,
 D_h = driver population in stratum h ,
 σ_h = composite standard deviation for stratum h , and
 T_h = desired tolerance for stratum h .

If no stratification is needed or no variability in vehicle occupancy patterns exists, the following formula can be used:

$$N = \frac{Z^2 \times D \times \sigma^2}{D \times T^2 + Z^2 \times \sigma^2}$$

where N = number of returned questionnaires required,
 σ = composite standard deviation for the region, and
 T = desired tolerance.

Because not all survey respondents return the surveys, it is important to achieve a maximum response rate, r , to reduce the no-response bias. Depending on the population and the length and

complexity of the questionnaire, it is reasonable to assume a 10 to 15% response rate. After the number of returned questionnaires N_h (or N) is determined, the number of questionnaires necessary to distribute N_{dh} (or N_d) can be calculated as follows:

$$N_{hd} = \frac{N_h}{r} \text{ or } N_d = \frac{N}{r}$$

4.2.3. Crash Records

Instead of determining the sample size to be collected, crash records are used to determine if the database contains a sufficient number of records to make the required estimation. The sampling unit is the vehicle(s) involving in an accident. Based on the defined population of interest in terms of geographic scope and temporal coverage, the number of vehicles required to estimate AVO within a specified precision level can be computed for each stratum by using the variability in occupancy rates:

$$N_h = \left(\frac{Z \times \sigma_h}{T_h} \right)^2$$

where N_h = number of vehicles required for stratum h ,
 σ_h = composite standard deviation for stratum h , and
 T_h = desired tolerance for stratum h .

The standard deviation estimates of vehicle occupancy for different strata of crashes can be obtained from existing studies. If the number of vehicles is not enough to provide estimation of AVO within a specified level of precision, another method of collecting vehicle occupancy data should be used.

4.3. Comparing Costs

Costs are calculated and compared for the remaining methods that are suited for the prospective data collection design. Heidtman *et al.* (1997) suggested the historical practice of disaggregating costs by survey stage for cost comparison. Because factors affecting the total cost vary among different collection methods, a procedure for establishing the relative costs was established. The costs associated with each method are broken down into the five stages involved in a collection effort: survey planning, survey design, data collection/entry, data analysis, and reporting. Table 4-3 shows a cost breakdown by data collection methods and study activities. The cost of each activity is ranked in terms of the relative level of costs associated with the given collection method. Typically, the data collection/entry activity is the most costly stage, which includes field personnel wages, data entry personnel wages, and equipment costs. Heidtman *et al.* (1997) suggested that the final cost estimates to collect vehicle occupancy data depend on a number of decisions made during the planning and design processes. Based on the specific study objectives, all methods but video surveillance may be the most cost-effective.

Table 4-3. Comparison of Occupancy Data Collection Costs by Method and Activity (Heidtman *et al.*, 1997)

Activity	Method				
	Roadside Windshield	Carousel Observation	Video Surveillance	Mail-Out Surveys	Crash Records
Planning	M	M	M-H	H	L
Design	L	M	H	H	L
Collection/Entry	study design (e.g., site-specific vs. area-wide) and regional labor costs greatly impact collection costs				
Analysis	L	M	M	L	H
Reporting	L	L	L	L-M	M

H = relatively higher cost; M = relatively medium cost; L = relatively lower cost

4.4. Sampling Randomly

The random sampling process should be applied to select a representative subset of the population such that every individual in the population has the same probability of being drawn. Ferlis (1981) suggested that the target population should include all possible highway segments and every possible date on which the sample could be randomly chosen. For area-wide studies, the locations and dates should be selected based on random sampling. For corridor and site-specific studies, the locations should be selected judgmentally, but the selection of dates should be randomly sampled. A link list and a date list are then generated for the area-wide studies, while corridor/site-specific studies only require a date list. In practice, there is a tendency to arbitrarily select higher volume roads such as freeways and major arterials, which can result in biased area-wide AVO estimates and limited extrapolation of the finding outside the collection sites.

The sample of link-days can then be drawn from the available lists using standard random selection techniques. The links are randomly selected first with a probability proportional to vehicle mile traveled (VMT), which is a function of the traffic volume and the length of link. All links could further be divided into multiple basic units with equal length (e.g., the minimum link length of all links). As a result, there is no difference in the weight assigned to a link from traffic volume or VMT (Levine and Wachs, 1994). Consequently, the higher volume roads have a greater probability of being selected over the lower volume road in this weighted random sampling process. After the locations are decided, the dates within a year for each location are then selected either randomly (data to be collected at several locations on same days might occur) or systematically (data collection effort is spread out evenly throughout the study period). If a stratified sampling plan is used, a separate sample of link-days and days should be selected for each stratum of area-wide studies and corridor/site-specific studies, respectively.

4.5. Computing AVO

4.5.1. Field Observational Methods

If short-count data collection procedures are used, an expansion factor, as defined below, should be computed to expand person and vehicle counts to total “day” estimates for specified time

period:

$$f_{ij} = \frac{N_i}{n_{ij}}$$

where f_{ij} = expansion factor for lane j at location i ,
 N_i = number of possible count periods during a session at location i , and
 n_i = number of actual count periods during a session on lane j at location i .

The simple (or unweighted) estimate of AVO can be computed by dividing the total number of occupants by the total number of vehicles for a specified geographic area and the time period. This calculation assumes that the traffic volumes observed at each location is proportional to the actual traffic flow at the corresponding location. The simple estimate of AVO and its corresponding composite standard deviation can be computed as follows:

$$AVO = \frac{\sum_i P_i}{\sum_i V_i}$$

$$\sigma = \left[\frac{N \times \sum_i (P_i - AVO * V_i)^2}{\left(\sum_i V_i \right)^2} \right]^{1/2}$$

where P_i = factored number of persons counted in session i , and
 V_i = factored number of vehicles counted in session i .

These same formulas can also be applied to data collected for each stratum of a stratified data collection to compute the average vehicle occupancy AVO_h and its composite standard deviation σ_h .

After computing the composite standard deviation of the data collection result, the actual precision of the estimates derived from the survey can be assessed. The actual precision (or tolerance) of the estimates obtained from the data collection can be computed as below:

$$T = \left(\frac{Z^2 \times \sigma^2}{N} \right)^{1/2}$$

The weighted mean of the strata AVOs by VMT (for area-wide studies) or traffic flow (for corridor studies) can be very important, especially if there are significant differences between AVOs on high- and low-volume roads. Weighted estimates are used to provide an overall measure representing multiple strata. If a particular location was only surveyed on one day, the composite standard deviation cannot be estimated from the survey results. The assumed value of

the composite standard deviation should be used to estimate precision. A weighted estimate of the AVO and corresponding tolerance can be computed as:

$$AVO = \sum_h W_h \times AVO_h$$

$$T = Z \times \left(\sum_h W_h^2 \times \sigma_h^2 / N_h \right)^{1/2}$$

where W_h is the estimated proportion of total VMT of all the selected links occurring in stratum h or of total traffic volumes across all selected locations occurring in location h . A level C (e.g., 95%) confidence interval is calculated by:

$$\text{Level } C \text{ Confidence Interval of } AVO = AVO \pm T$$

4.5.2. Mail-Out Questionnaires

Since mail-out questionnaires are used frequently to collect area-wide information with stratification technique, a stratified estimator of AVO is as follows:

$$AVO_h = \sum_{j=1}^{N_h} \frac{P_{hj}}{N_h}$$

where P_{hj} = number of occupants recorded on questionnaire j in stratum h , and
 N_h = number of questionnaires returned in stratum h .

To estimate the regional AVO, the weighted estimate of AVO should be calculated as follows rather than using the simple mean procedure:

$$AVO = \frac{\sum_h D_h \times AVO_h}{D}$$

where D_h = driver population in stratum h , and
 D = total driver population in the region.

4.5.3. Crash Records

In calculating AVO estimates from crash records, an approach analogous to the simple mean procedure for estimating AVOs of field observational method is normally applied. This method is not applicable given the fact that crash data are likely to be a biased sample of the population. To accommodate these biases, some adjustments and filtering of the data as well as a weighted means procedure are required to remove biases inherent in these types of data.

CHAPTER 5

EXTRACTION OF VEHICLE OCCUPANCY RATES FROM ACCIDENT RECORDS

This chapter describes a user-friendly system that applies multiple years of Florida accident records to derive AVO estimates. The system, named **Florida Accident Vehicle Occupancy Rate InformaTion Estimator**, or FAVORITE, is able to generate AVO estimates at the district, county, and corridor levels.

5.1. Overview

FAVORITE is a typical Windows program designed to run on Microsoft Windows operating systems. The program is fully stand-alone and was developed using Visual Basic, Microsoft Access, and ESRI MapObjects Developer Library. The database that comes with FAVORITE includes the 1990-2002 accident data for the Florida state highway system. The data include the number of passengers of each accident for up to two vehicles, which are used to calculate the average vehicle occupancy (AVO). In addition, the database also includes the following variables:

- District
- County
- Section, subsection, beginning and ending mileposts
- State road number
- Hour of day
- Day of week
- Month of year
- Type of vehicle
- Type of road
- Type of area
- Accident severity

The major functionalities of the current version of FAVORITE include:

- Calculate and display AVOs in a cross table of two variables, for example, the AVO for different days of a week for different FDOT districts. The total AVOs are also provided for all rows and columns of the cross table.
- Calculate and display AVOs on line, bar, or pie charts.
- Calculate and display AVOs on a GIS map at the district, county, and corridor level.
- Export tables and charts to Excel.
- Apply filters for time of day, day of week, month of year, type of vehicle, type of road, type of area, and accident severity.
- Allow variables to be re-categorized. For example, you can define the Spring season by combining the months of January, February, and March.

5.2. Installation

The FAVORITE setup was packaged using the InstallShields install software. To install FAVORITE, insert your FAVORITE CD and wait several seconds for the install to automatically start the setup program. You can then follow the instructions on the screen to complete the installation. If an existing version of FAVORITE is detected on your computer, you will be prompted to remove it before you can install a new version.

5.3. Input Specifications

Figure 5-1 shows the main screen of the FAVORITE program. The screen allows you to make the following three major selections:

- Accident data years
- Locations
- Filters

In addition, the main screen also allows you to re-categorize the variables. These functions are further described in the following sections.

The screenshot shows the 'FAVORITE' application window with the title 'Florida Accident Vehicle Occupancy Rate InformaTion Estimator'. The interface is organized into several sections:

- Year:** 'From: 99' and 'To: 02' with a 'Set Default' button.
- District:** A list of districts (1-8) with 'District 1' selected. Buttons for 'Select All', 'Clear All', and 'Map' are present.
- County (optional):** A list of counties (01-91) with '01 Charlotte' selected. Buttons for 'Select All', 'Clear All', 'Map', and 'Sort' are present.
- Segment (optional):** A table with columns 'County/Sec/Subsec', 'From Milepost', and 'To Milepost'. It contains one row with an asterisk (*). Buttons for 'Clear All' and 'Map' are present.
- Filters:**
 - Severity:** Checkboxes for 'Fatal', 'Injury', and 'PDO' are all checked.
 - Day:** Checkboxes for 'Mon', 'Tue', 'Wed', 'Thu', 'Fri', 'Sat', and 'Sun' are all checked.
 - Hour:** A dropdown menu showing '12:00 - 1:00 AM'.
 - Month:** A dropdown menu showing '1 Jan'.
 - Area Type:** A list of area types (1-9) with '1 Outside City, Outside Urban' selected.
- Vehicle Type:** A list of vehicle types (0-5) with '0 Unknown/ Not Coded' selected.
- Road Type:** A list of road types (0-5) with '0 Not Divided' selected.

At the bottom of the window are buttons for 'Table', 'Chart', 'GIS', 'Exit', and 'Help'.

Figure 5-1. FAVORITE Main Screen

5.3.1. Accident Data Years

The current version of FAVORITE includes the complete 1990-2002 accident records for Florida's State Roadway System. You can select any number of years of data to include in the analysis using the **From** and **To** dropdown lists on the main screen.

5.3.2. Location Selection

Location selection can be made at three different levels: district, county, and corridor. More than one item in each level can be selected. For example, you can select Districts 4 and 6 to cover the southeast Florida region. When a district is selected, all the counties in it will be listed under the **County** list box. Selection of counties is optional. If no counties are selected, the selection is considered to have been made at the district level only, that all counties in a selected district will be included. When a county is selected, only data for that county is included.

The selection of districts or counties is made by checking the appropriate checkboxes. Alternatively, you can click the **Map** button beside the **District** or **County** list box to select by map. Selection by map is convenient when you want to select a contiguous area such as selecting two adjacent counties. Figure 5-2 shows a map screen in which the Miami-Dade and Broward counties are selected.

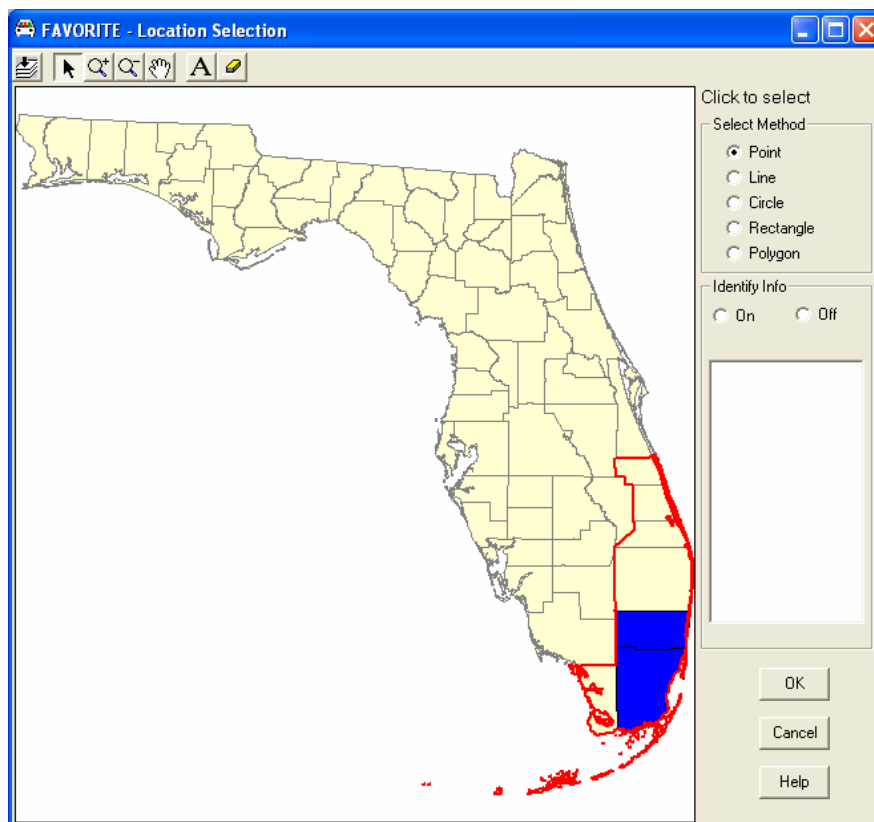


Figure 5-2. Select Locations by Map

In this screen, you can:

- Select a county by clicking the county polygon on the map. Selected counties are shown in blue. Note that only the counties in the selected districts can be selected. The selectable counties are enclosed by red borders.
- Click a selected county to unselect it.
- Click  to clear all selections.
- Click  to display county or district names.
- Click , , and  to zoom in, zoom out, or pan the map.
- Click  to display the complete map.
- Choose a different selection method such as “Rectangle” to select multiple features at once by drawing a rectangle with the mouse cursor.
- Click the “**On**” radio button to display information of a feature pointed by mouse cursor.
- Click **OK** to exit the map view and return to the main screen with your selections.

To select a specific corridor, you can either enter the county/section/subsection standard roadway ID and the beginning and ending mileposts, or you can also click the adjacent **Map** button to select a particular roadway by pointing and clicking on the map.

5.3.3. Filters

In addition to accident data years and analysis locations, FAVORITE includes filters for the following variables:

- Accident severity
- Hour of day
- Month of year
- Day of week
- Vehicle Type
- Area type
- Road type

This allows you to include only a subset of these variables by checking or unchecking the appropriate checkboxes for the variable options listed for each variable. By default, all options of a variable are included. At least one variable option must be selected.

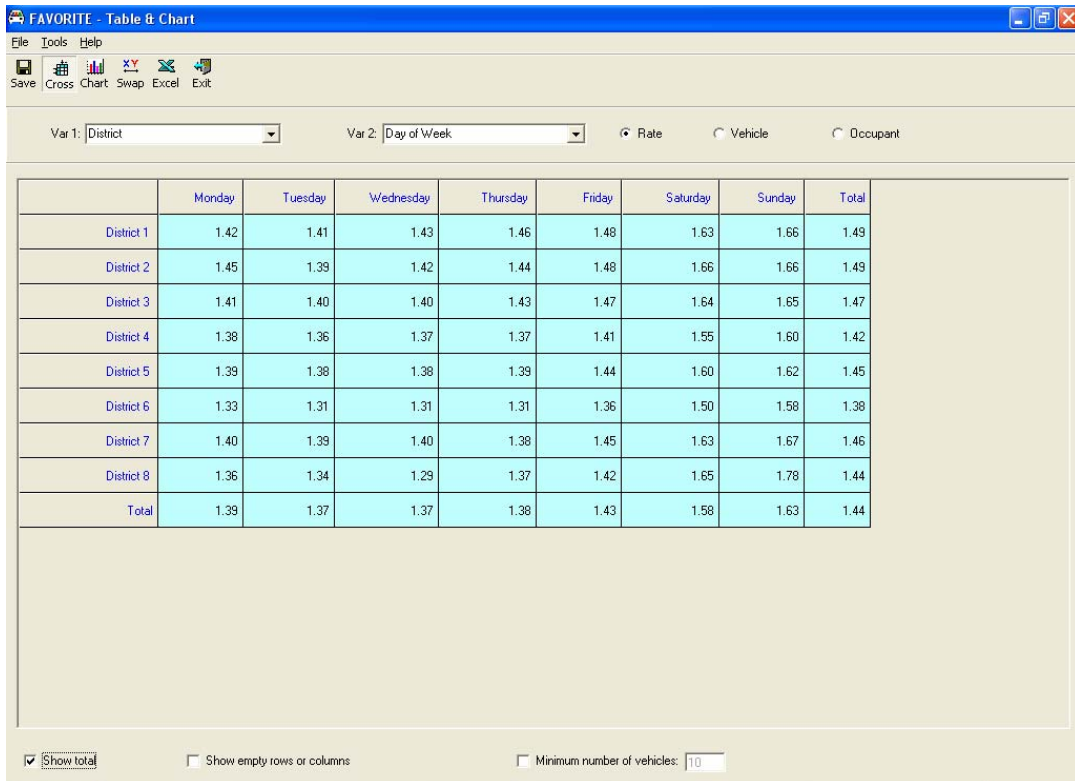
5.4. Output Options

Once the input specifications described in the previous section are completed, you can click the **Table**, **Chart**, or **GIS** button at the bottom of the main screen to start computing and displaying AVOs on a cross table, a chart, or a map, respectively.

5.4.1. Table Display

This option allows you to display the AVOs on a cross table. Figure 5-3 shows a table that is cross-classified by District and Day of Week. The following functions are available:

- The **Var 1** and **Var 2** dropdown lists allow you to select up to two variables to cross classify the AVO estimates. By default, AVO estimates are displayed when the table is first displayed.
- Click the **Vehicle** and **Occupant** buttons to display the corresponding cross tables for number of vehicles and number of occupants, respectively, by clicking the appropriate radio buttons.
- Click  to export the current view to Excel.
- Click  to swap the rows and columns of the cross table.
- Click  to exit to the main screen.
- Uncheck the **Show total** option to exclude the summary column and row in the table.
- Check the **Minimum number of vehicles** option and then enter a threshold number to exclude cells that do not meet the minimum vehicle sample size.
- Check the **Show empty rows or columns** option to show rows and columns that are empty (i.e., with zero sample size for all cells across a row or a column).
- Select from the **View|X-Long or Short** or **View|Y-Long or Short** dropdown menu items to toggle between whether to show the variable option name after each option code.



The screenshot shows the 'FAVORITE - Table & Chart' application window. The interface includes a menu bar (File, Tools, Help), a toolbar with icons for Save, Cross, Chart, Swap, Excel, and Exit, and a control panel with dropdowns for 'Var 1: District' and 'Var 2: Day of Week', and radio buttons for 'Rate' (selected), 'Vehicle', and 'Occupant'. The main area displays a table with 9 rows and 9 columns. The first column lists districts (District 1 to District 8) and a 'Total' row. The first row lists days of the week (Monday to Sunday) and a 'Total' column. The table cells contain numerical AVO estimates. At the bottom, there are checkboxes for 'Show total' (checked), 'Show empty rows or columns' (unchecked), and a text input for 'Minimum number of vehicles' set to 10.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Total
District 1	1.42	1.41	1.43	1.46	1.48	1.63	1.66	1.49
District 2	1.45	1.39	1.42	1.44	1.48	1.66	1.66	1.49
District 3	1.41	1.40	1.40	1.43	1.47	1.64	1.65	1.47
District 4	1.38	1.36	1.37	1.37	1.41	1.55	1.60	1.42
District 5	1.39	1.38	1.38	1.39	1.44	1.60	1.62	1.45
District 6	1.33	1.31	1.31	1.31	1.36	1.50	1.58	1.38
District 7	1.40	1.39	1.40	1.38	1.45	1.63	1.67	1.46
District 8	1.36	1.34	1.29	1.37	1.42	1.65	1.78	1.44
Total	1.39	1.37	1.37	1.38	1.43	1.58	1.63	1.44

Figure 5-3. AVOs Cross-Classified by FDOT District and Day of Week

5.4.2. Chart Display

The **Chart** option allows you to display AVOs on a chart. The same interface is shared by both the **Cross Table** and **Chart** displays. Once you have selected one of the two display options, you can simply click  and  to switch between the two displays. Figure 5-4 shows a chart for automobile AVOs for different FDOT districts for different days of a week. The overall AVOs are also shown for each district and all district combined.

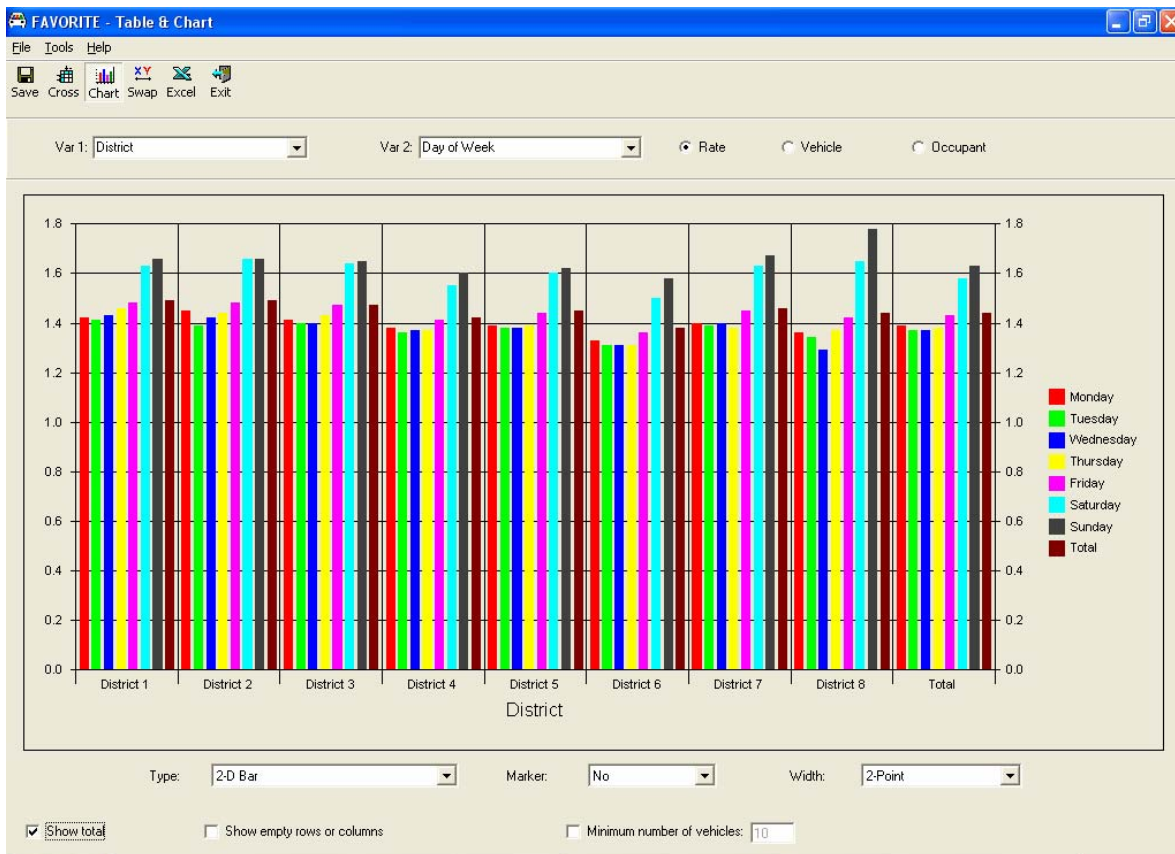


Figure 5-4. AVOs Cross-Classified by FDOT District and Day of Week

5.4.3. GIS Display

While tables and charts are able to display AVOs that are cross-classified by variables, they cannot show the AVOs spatially. FAVORITE provides a GIS interface that can display AVOs by district, county, and segment. Figure 5-5 shows the main interface of the GIS display. The left side of the screen allows you to make various selections while the right side of the screen displays a map view of the state. The top-left corner lists the GIS layers and the corresponding colors used for display. Below this layer box is a **Variable** dropdown box that allows you to specify the variable options to include in the calculation of AVOs. Once a variable is selected, all options will be displayed at the box below it. You can select the options to include by checking the appropriate checkboxes.

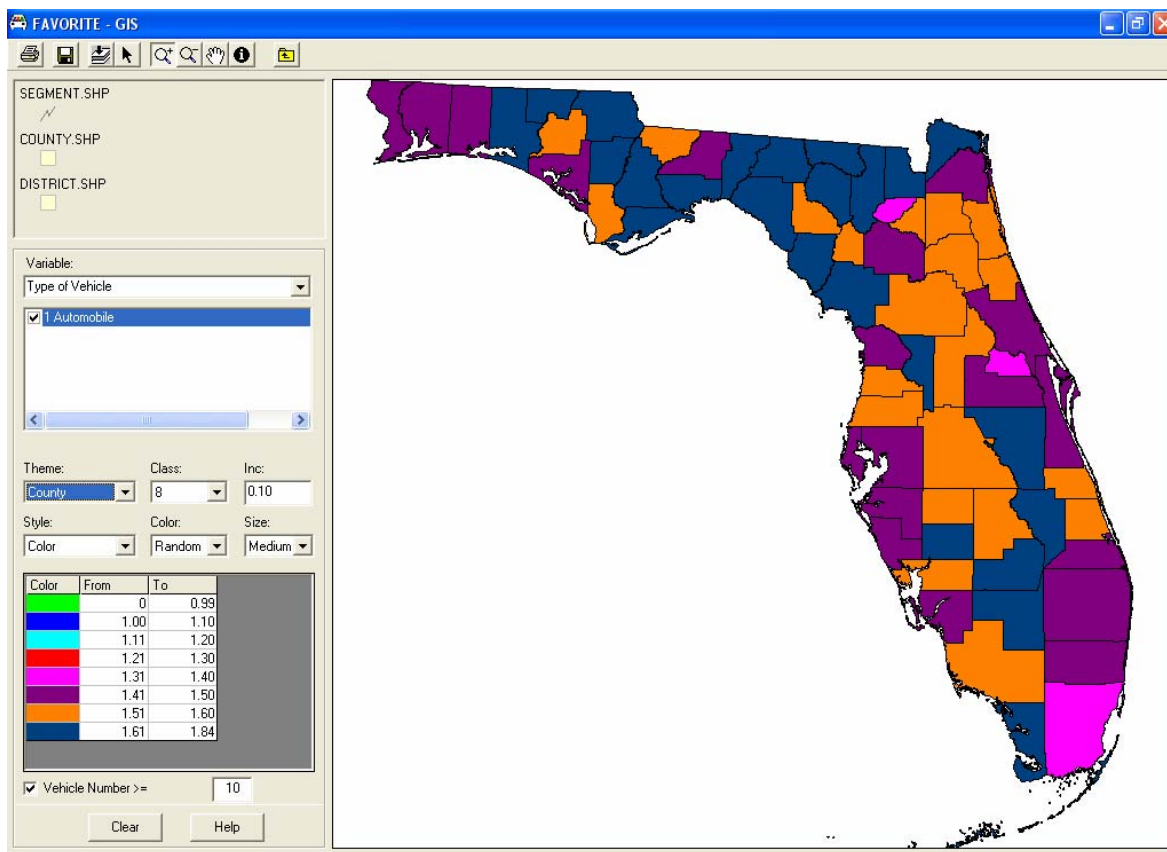


Figure 5-5. Automobile AVOs Displayed by County

Once a variable option is selected, the six specification boxes below the option list box will become active. The first of these boxes allows you to select a theme to display by county, district, or segment. By default, the **Segment** option is selected.

The **Class** dropdown list box allows you to select the number of classes for the theme. The default number of classes is seven. The first class is always assigned to 0.00 to 0.99, which normally include features that have no accidents. The last class includes any number above a certain threshold. All classes between these two boundary classes are divided based on the increment specified to the right. By default, the increment is 0.1.

The **Style** option allows you to specify either to show themes by color, line width or a combination of both. The default option is to display by different colors. Obviously, the line width option applies only to the **Segment** theme. The **Color** option allows you to select a color scheme. By default, random colors are used. You may choose to use the **Blue** and **Red** color schemes, which display features from gradual light to dark colors to indicate low to high AVOs. The GIS display is automatically refreshed as soon as any one of these specifications is changed. Figure 5-6 shows an example of AVOs displayed by segments.

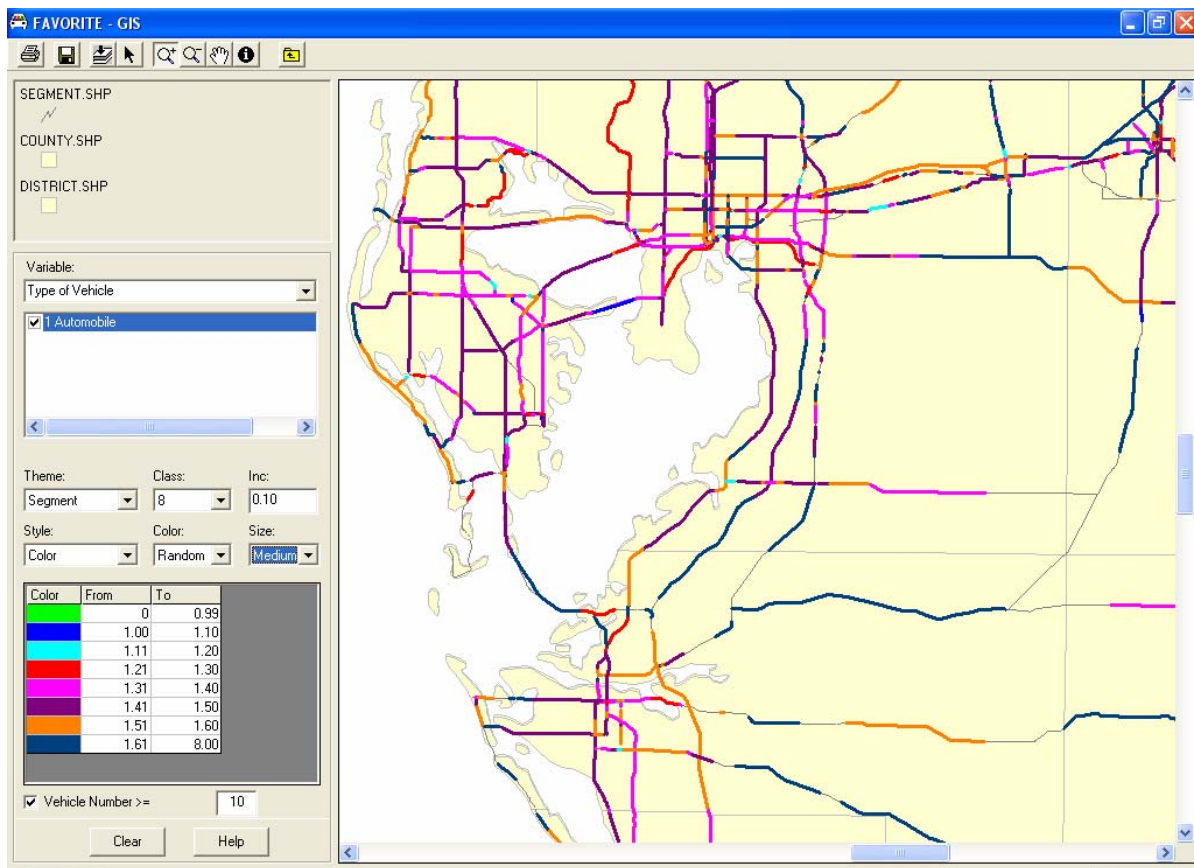


Figure 5-6. Automobile AVOs Displayed by Segment

A number of tool buttons are available on the GIS screen. Clicking the  button will bring you to the **Print Layout** screen shown in Figure 5-7. In this screen, you can:

- Single click any object on the screen and then drag it to re-position the object.
- Single click on an object to highlight the object and then drag the mouse cursor to enlarge or shrink the object.
- Double click “title” or “footnote” to enter a title or a footnote, respectively.
- Click  to print the map to the default printer.
- Click  to toggle between the “portrait” and “landscape” page orientation.
- Click  to copy the print layout to clipboard.
- Click  to add a text box.
- Click  to remove a text box.
- Click  to add or remove the page border.

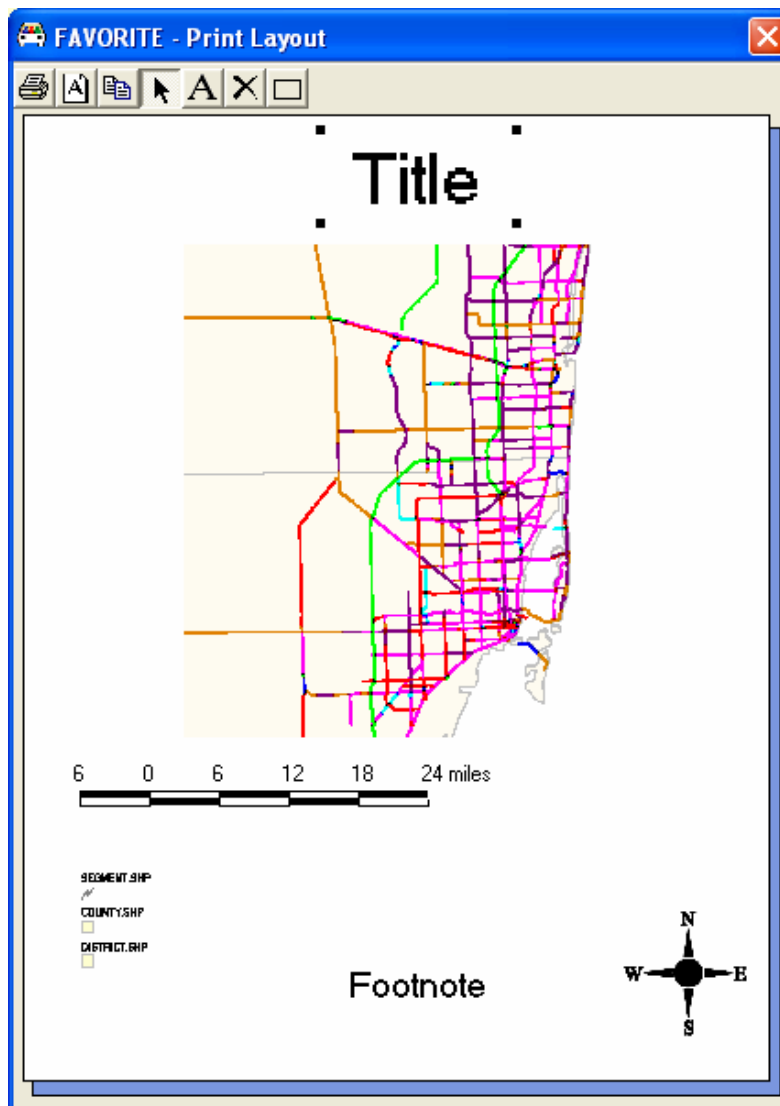


Figure 5-7. Print Layout

In addition to the **Print** button, you can:

- Click  to save the map as a new shape file.
- Click  to return to the mouse pointer.
- Click  to display county or district names.
- Click , , and  to zoom in, zoom out, or pan the map.
- Click  to display the complete map.
- Click  to identify features by mouse cursor.
- Click  to exit to the FAVORITE main screen.

5.5. Variable Re-categorization

FAVORITE allows you to re-categorize the options of each variable. To add a new category, select the **Setup|Add Category** dropdown menu item from the main screen to invoke the screen shown in Figure 5-8.

Group Name	Equivalent Codes
Spring	1,2,3
Summer	4,5,6
Fall	7-9
Winter	10, 11, 12

Figure 5-8. Screen for Adding a Category

To create a new category, follow these steps:

- Enter a name for the new category.
- Select an original variable to be re-categorized.
- Enter the names of groups under the **Group Name** column.
- Enter the code numbers to be included in each group. Code numbers are separated by commas. Enter **1, 2, 3** if to assign January, February, and March from the **Month of Year** variable. Alternatively, you can also enter **1-3**.
- If you are not familiar with the code numbers, click the cell for which codes are to be entered and then press the **F2** function key. This will invoke the screen shown in Figure 5-9, which lists all the codes and option names for the selected variable. Check the boxes to be included and click **OK** to return the selected options to the **Add Category** screen.
- Click the **Save** button to save the new category and exit to the FAVORITE main screen.

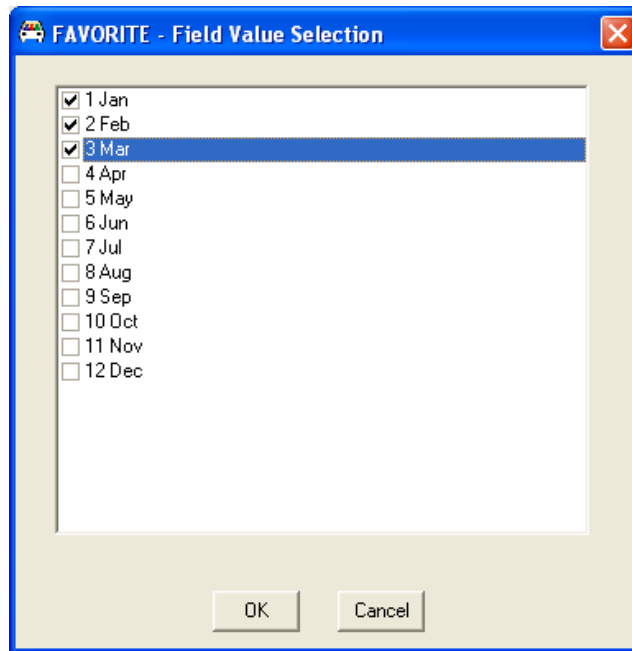


Figure 5-9. Screen for Specifying Options for Group

To edit an existing category, select the **Setup|Edit Category** dropdown menu item. This will bring up a screen similar to the one in Figure 5-8. The screen allows you to specify an existing category to edit.

5.6. Validations

This section attempts to validate the AVO estimates from FAVORITE via both reasonableness checks and field data comparisons.

5.6.1. Reasonableness Checks

Reasonableness checks involve examining the AVO estimates to determine if they match expectations and known trends. A number of cases are presented below:

Case 1: AVO Trends by Year and Month

Figure 5-10 shows that, as expected, weekend AVOs for automobiles are higher than weekday AVOs, with Sunday having the highest AVOs, that Mondays and Fridays tend to have slightly higher AVOs than Tuesdays, Wednesdays, and Thursdays.

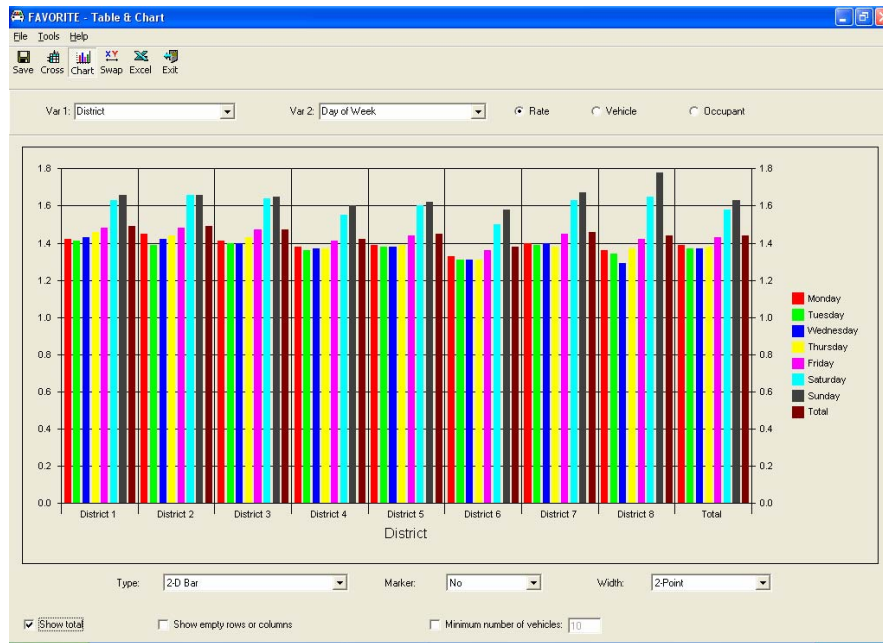


Figure 5-10. AVO Trend by Year and Day of Week

Case 2: AVO Trends by Month

Figure 5-11 shows that, as expected, the summer months of July and August have the highest AVOs while the months of September, October, and November have the lowest. Overall, AVOs do not change significantly over the year.

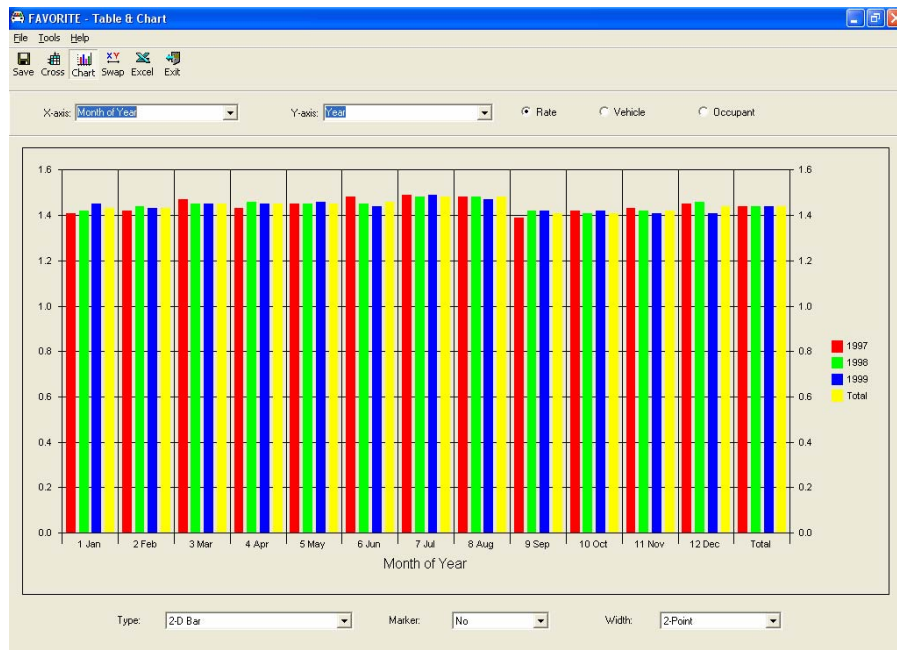


Figure 5-11. AVO Trend by Month

Case 3: AVO Trends by Hours

Figure 5-12 shows that, as expected, AVOs during the morning peak hours are lower than that of the afternoon peak hours, that the daytime AVOs are lower than the nighttime AVOs. An interesting observation is that the AVOs continue to drop over the period of 9:00 pm to 1:00 am, with the hour just past the mid-night having a significantly lower AVO than any other hours. However, after 1:00 am, the AVOs increased significantly.

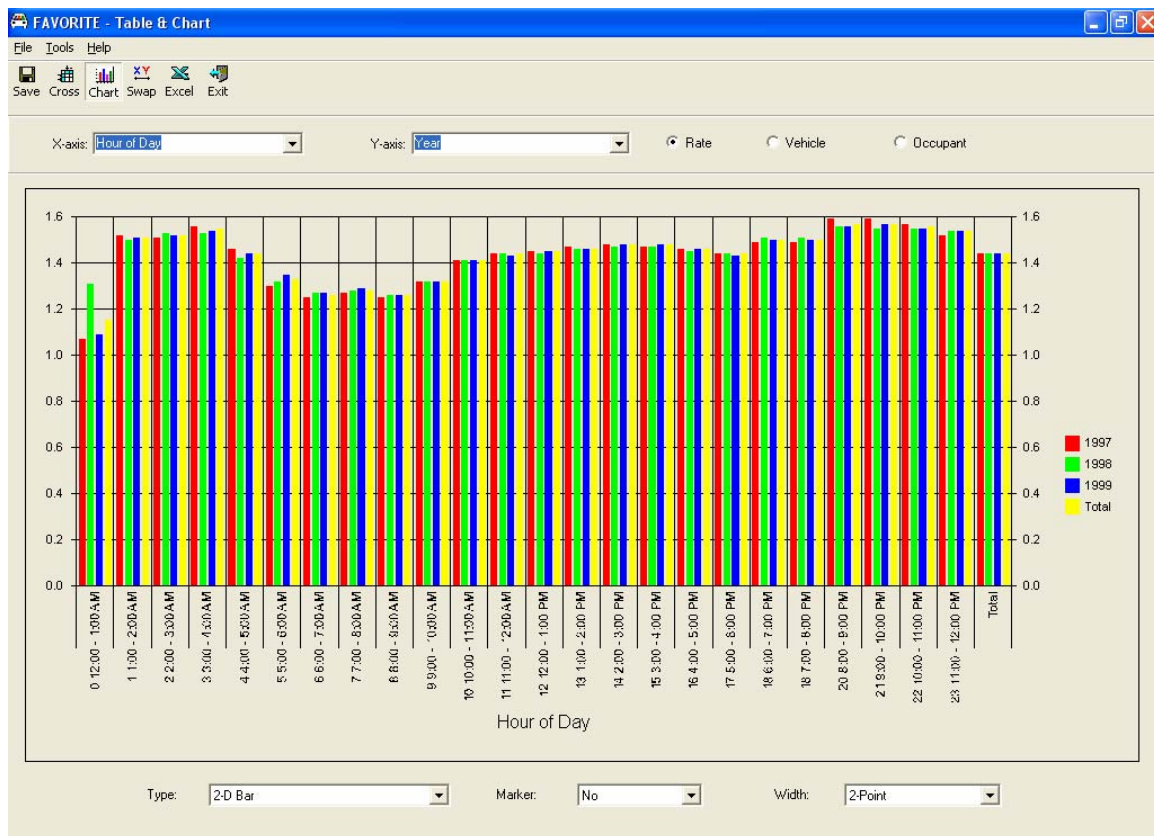


Figure 5-12. AVO Trend by Time of Day

Case 4: AVO Trends by Area Type

Figure 5-13 shows a GIS thematic map of AVO distribution by county. The darker the color, the higher the AVO. The map shows that, as expected, the more rural counties have higher AVOs than those of the more urbanized counties.

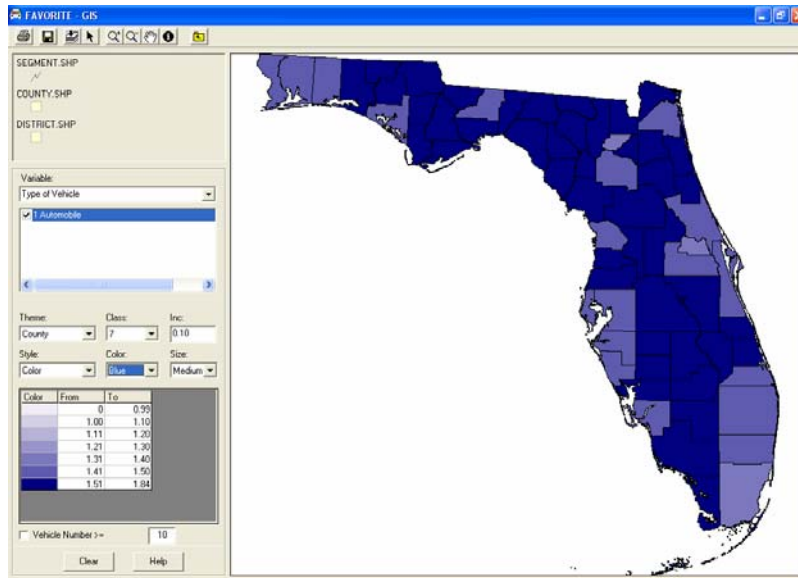


Figure 5-13. AVO Trend by County

Case 5: AVO Trends by Vehicle Type

Figure 5-14 shows the AVOs for different types of vehicles. The AVOs obtained appear to be logical compared to the expected AVOs for different vehicle types. Some observations include:

- AVOs for passenger vans are higher than those for passenger cars.
- Buses have a significantly higher AVO.
- Bicycles have the lowest AVOs.

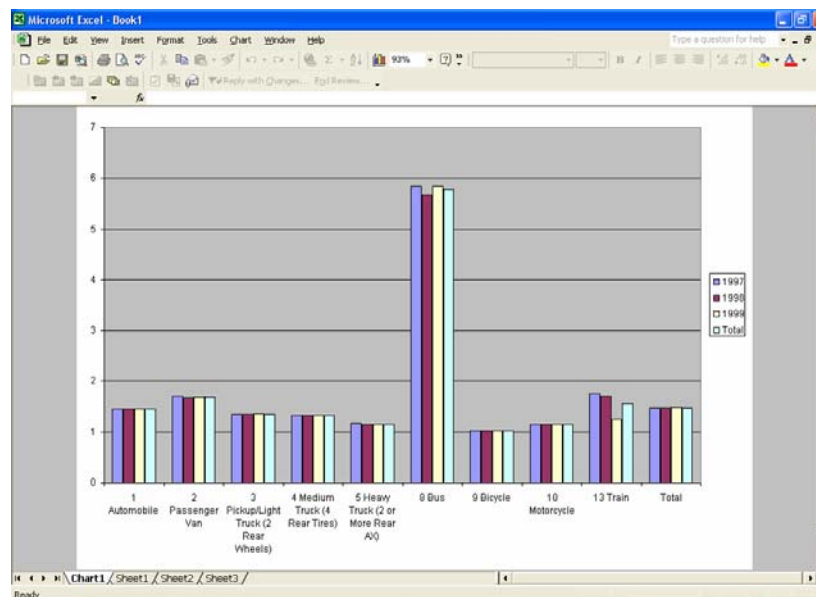


Figure 5-14. AVO Trend by Vehicle Type

5.6.2. Field Data Comparisons

Table 5-1 compares the AVO estimates from the field data collected in the 1998 study conducted by TEI Engineers and Planners and RSH, Inc. with those estimated from FAVORITE for the same locations. The list includes only those locations with at least 100 counted vehicles. In computing the AVOs in FAVORITE, accident from one mile upstream and one mile downstream of the field collection location for years 1998 and 1999 were included.

Table 5-1. Comparisons of AVOs from Field Data and FAVORITE

County	Roadway	Milepost	Total Vehicles	Total Occupants	AVO from Field Data	AVO from FAVORITE
Hillsborough	10190000	1.340	269	348	1.27	1.29
Hillsborough	10320000	0.683	393	542	1.22	1.38
Pinellas	10310000	2.847	765	1165	1.52	1.52
Hillsborough	10320000	7.871	136	186	1.27	1.37
Sarasota	10075000	24.081	110	137	1.20	1.25
Duval	72280000	14.16	137	192	1.20	1.40
Orange	75280000	7.500	149	249	1.35	1.67
Orange	75280000	17.059	455	593	1.23	1.30
Volusia	79110000	4.668	102	148	1.27	1.45
Broward	86070000	22.595	452	573	1.23	1.27
Broward	86095000	6.151	220	285	1.22	1.30
Broward	86070000	1.108	606	836	1.25	1.38
Dade	87004000	2.108	195	247	1.27	1.27
Palm Beach	93220000	27.939	324	441	1.26	1.36
Palm Beach	93220000	13.279	330	473	1.23	1.43

Figure 5-15 plots the field estimates for the 15 locations against the FAVORITE estimates for the same locations. While the regression equation shows a positive relationship between the two sets of AVO estimates, the R^2 value is not high. Overall, the AVO estimates from the crash records tend to be higher than the field estimates. This is expected since the windshield method used in the field data collection was not able to capture all passengers, especially infants and children.

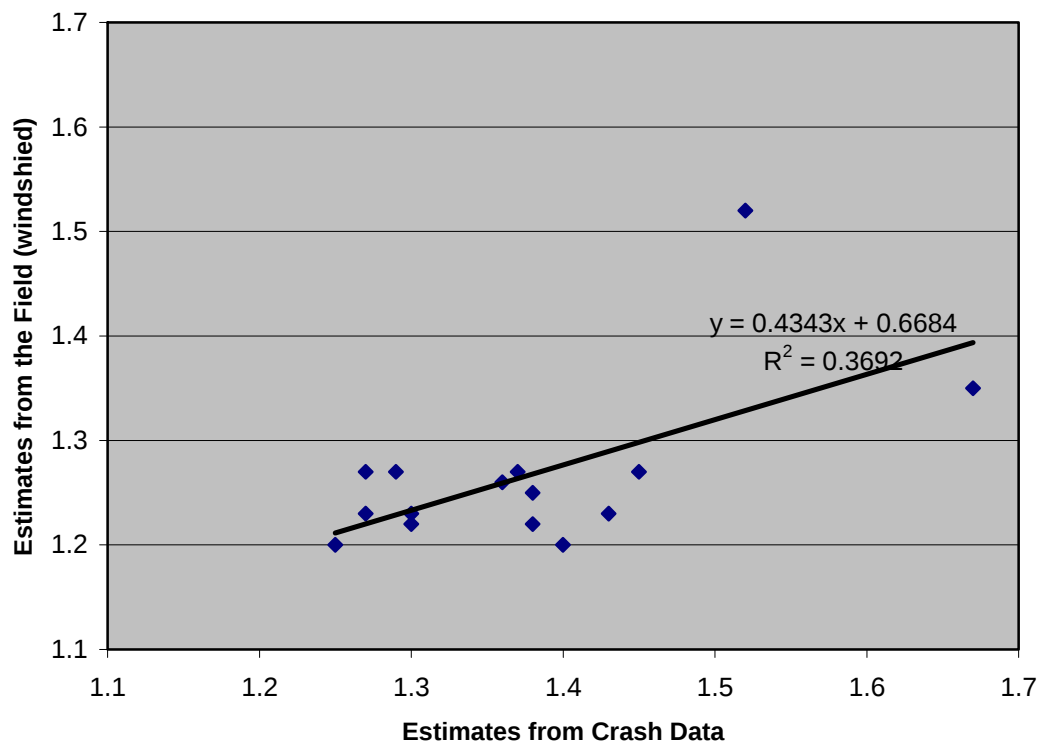


Figure 5-15. Field versus FAVORITE Estimates

CHAPTER 6

AUTOMATED FIELD DATA COLLECTION SYSTEM

This chapter describes an automated field data collection system designed for use with a handheld Pocket PC. In addition, a post-processing program that can calculate the average occupancy rates from the data collected from the automated data collection tool is described. It is noted that the research team investigated the possibility of applying voice recognition technology in lieu of screen input on a Pocket PC. This was found to be impractical for two reasons: (1) Lack of appropriate software: the research team was not able to find a suitable voice recognition system that can work with the Windows CE operating system; and (2) unreliability: interference from traffic noise in the field can easily result in the recording of incorrect data.

6.1. Overview

The field data collection tool provides a touch-screen interface on a Pocket PC to record the number of passengers in vehicles classified by vehicle type and/or lane number. The field collected data are then processed by a desktop program to produce average occupancy rates and related statistics.

The field data collection tool is a typical Window installation designed to run on all Microsoft Windows CE operating systems. For successful installation and application of the tool, your system must have:

- Minimum screen size: 3.5"
- A minimum of 48 MB memory storage space for a full installation

For general help on using Windows CE system such as managing the environment and the file system, refer to the Microsoft Windows CE and Microsoft ActiveSync User's Guide.

6.2. Installation

To install the field data collection tool, follow the following steps:

1. Set up the connection between the Pocket PC and the desktop computer.
2. Insert the field data collection tool CD into the CD-ROM drive.
3. Use Windows Explorer to find **Setup.exe** in the CD1 folder on your CD-ROM drive. Double click the **Setup.exe** icon.
4. Choose a destination location.
5. Continue to follow on-screen instructions until the shown screen in Figure 6-1 pops up. Clicking **Yes** will install files to the default storage in the Pocket PC directly. Clicking **No** will invoke the screen shown in Figure 6-2, which allows you to select a storage option from the **Save In** list box. Click **OK** button to complete the installation.

Note: *If the **Confirm File Replace** box is shown on the screen of Pocket PC at the same time, select the **No to All** button to keep the existing files of the Pocket PC system.*

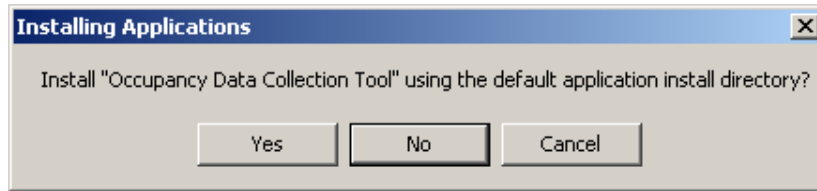


Figure 6-1. Determine Application Storage

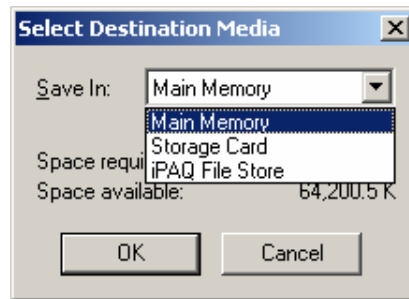


Figure 6-2. Select Storage Option

6.3. Main Screen

The field data collection tool can be started by clicking the **Start|Programs** menu from the Pocket PC. The main screen of the program is shown in Figure 6-3. The screen allows you to specify whether to differentiate your occupancy data by vehicle type, lane number, or both. These options are specified by checking the appropriate checkboxes on the screen.

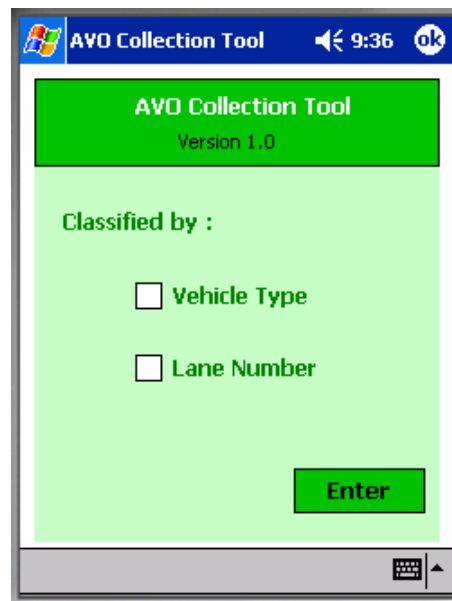


Figure 6-3. Input Screen for Data Classification

6.4. General Screen

After making the choices on classification, you can click the **Enter** button to bring up the screen shown in Figure 6-4, which lists three buttons on top of the screen: General and Data Entry. By default, the screen will display the **General** form. This allows you to select whether to create a new file or open an existing file, and to input the general roadway location information. The **Data Entry** button allows you to input the number of passengers and, if applicable, vehicle type and/or lane number.

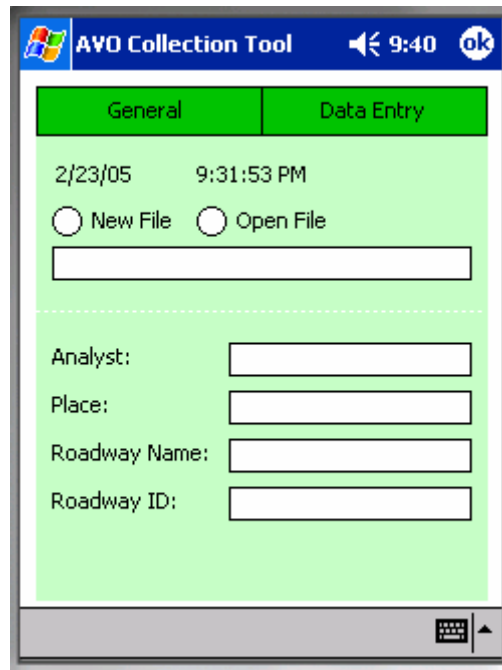
The screenshot shows a Windows-style application window titled "AVO Collection Tool". The window has a blue title bar with a Windows logo, a speaker icon, the time "9:40", and an "ok" button. Below the title bar is a green header bar with two tabs: "General" (selected) and "Data Entry". The main area has a light green background. It displays the date "2/23/05" and time "9:31:53 PM". Below this are two radio buttons: "New File" (selected) and "Open File". Under the radio buttons is a white text input box. A dashed horizontal line separates this section from the location information section below. This section contains four labels with corresponding white text input boxes: "Analyst:", "Place:", "Roadway Name:", and "Roadway ID:". At the bottom right of the window is a small icon and a scroll bar.

Figure 6-4. Input Screen for General Information

6.4.1 Create a New File

To create a new file, follow the steps below:

1. On the **General** form screen, click the **New File** radio button to bring up the screen shown in Figure 6-5.
2. Enter the new file name in the text box.
3. Choose the folder for saving the new file from the **Folder** list box.
4. Choose a storage location from the **Location** list box.
5. Click **OK** to finish creating a new file. The file path and name will be shown in the text box under the radio buttons as shown in Figure 6-4.

6.4.2. Open a Existing File

To open an existing file, follow the steps below:

1. Click the **Open File** radio button to bring up the screen shown in Figure 6-6.

2. Select a folder name from the **Folder** list box. All the files with the selected file type (*.cdb) in the folder will be displayed.
3. Double click a file name to open. The file path and name will be shown in the text box under the radio buttons on Figure 6-4.

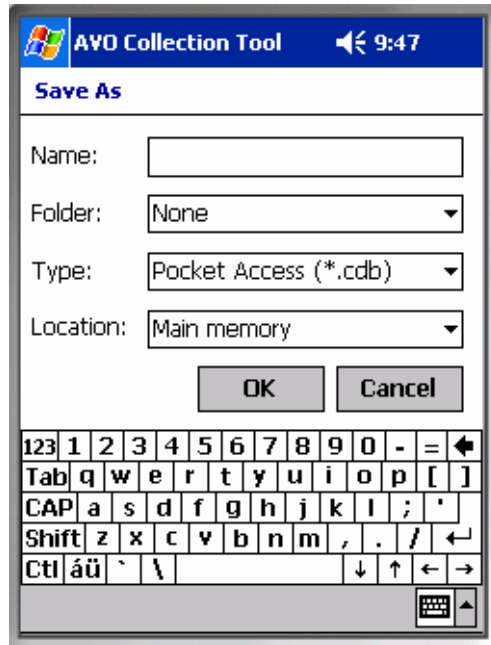


Figure 6-5. Create a New File

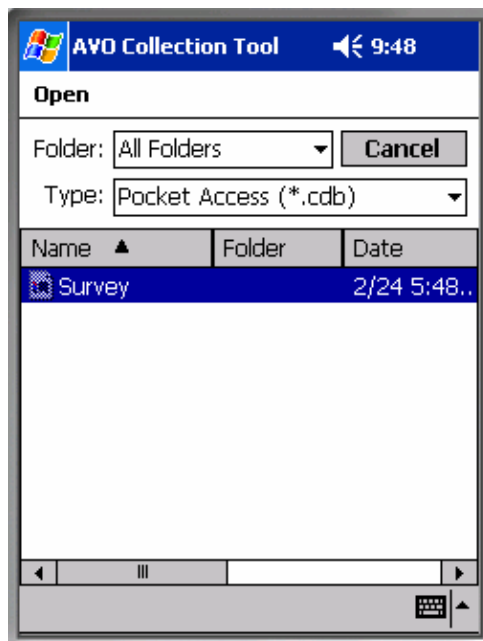


Figure 6-6. Open an Existing File

6.4.3. Input Location Information

The general location information can be added or edited after creating a new file or opening an existing file.

1. Click the **Name** text box.
2. Click the  icon on the right corner to bring up a keyboard to key in the name.
3. Repeat Steps 1 and 2 for Place, Roadway Name, and Roadway ID.
4. Click the **Data Entry** button to start collecting occupancy data.

6.5. Data Entry Screen

Depending on the classification chosen, one of the following four screens will be displayed:

1. Without vehicle type and lane number
2. With vehicle type only
3. With lane number only
4. With both vehicle type and lane number

The respective screens for each of these four classifications are shown in Figures 6-7 to 6-10, respectively.

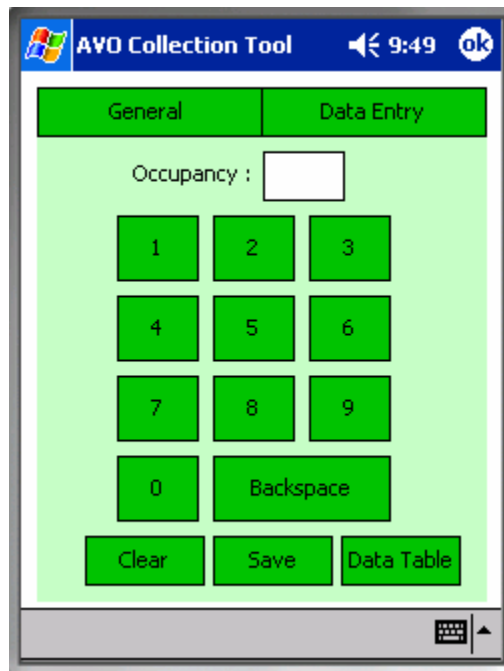


Figure 6-7. Data Entry Screen: Without Vehicle Type and Lane Number

The screenshot shows the 'AVO Collection Tool' interface. At the top, there is a status bar with a Windows logo, the text 'AVO Collection Tool', a speaker icon, the time '9:51', and an 'ok' button. Below this is a green header bar with two tabs: 'General' and 'Data Entry'. The 'Data Entry' tab is active. The main area has a light green background. It starts with two input fields: 'Type :' and 'Occupancy :'. Below these is a grid of buttons. The first column contains buttons labeled 'Car', 'Truck', 'Bus', and 'Other'. The second column contains buttons labeled '1', '4', '7', and '0'. The third column contains buttons labeled '2', '5', '8', and 'Backspace'. The fourth column contains buttons labeled '3', '6', '9', and 'Backspace'. At the bottom of the grid are three buttons: 'Clear', 'Save', and 'Data Table'. A small keyboard icon and a cursor are visible at the bottom right of the screen.

Figure 6-8. Data Entry Screen: with Vehicle Type Only

The screenshot shows the 'AVO Collection Tool' interface. At the top, there is a status bar with a Windows logo, the text 'AVO Collection Tool', a speaker icon, the time '9:53', and an 'ok' button. Below this is a green header bar with two tabs: 'General' and 'Data Entry'. The 'Data Entry' tab is active. The main area has a light green background. It starts with two input fields: 'Lane :' and 'Occupancy :'. Below these is a grid of buttons. The first column contains buttons labeled 'Lane 1', 'Lane 2', 'Lane 3', and 'Lane 4'. The second column contains buttons labeled '1', '4', '7', and '0'. The third column contains buttons labeled '2', '5', '8', and 'Backspace'. The fourth column contains buttons labeled '3', '6', '9', and 'Backspace'. At the bottom of the grid are three buttons: 'Clear', 'Save', and 'Data Table'. A small keyboard icon and a cursor are visible at the bottom right of the screen.

Figure 6-9. Data Entry Form with Lane Number

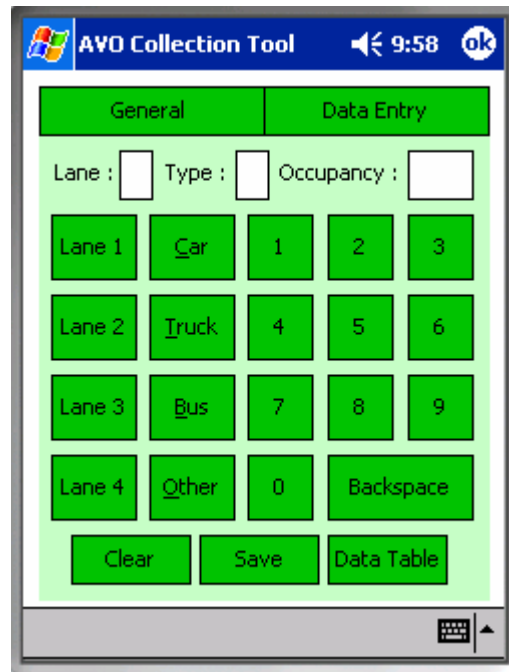


Figure 6-10. Data Entry Screen: with Both Vehicle Type and Lane Number

The operations of these four screens are similar:

- Click the number buttons to enter the vehicle occupancy. The number will be shown on the **Occupancy** text box.
- Click **Save** to save a new entry (for each vehicle).
- Click the **Backspace** to delete an entered occupancy number.
- Click the **Lane** (if applicable) to specify the lane number. You may set your own lane numbering scheme and then use it consistently. Click a new lane number to replace an entered lane number.
- Click **Vehicle Type** (if applicable) to specify the vehicle type. Click a new vehicle type to replace an entered vehicle type.
- Click **Clear** to clear all entries, including occupancy, and if applicable, lane number and vehicle type.
- Click **Data Table** to bring up screen shown in Figure 6-11. This is further described in the next section.

6.6. Editing Data Table

The data table allows you to make changes to the entered records. An example is given in Figure 6-11.

The screenshot shows a software window titled "AVO Collection Tool" with a Windows logo on the left and a speaker icon and "1:33" on the right. The main area contains a form with the following fields and values:

Analyst:	Xin Li
Roadway Name:	West Flagler Street
Place:	Miami-Dade
Roadway ID:	87001
Lane Number:	1
Vehicle Type:	C
Occupancy:	2
Date & Time:	2/05/05 5:43:52 PM

Below the form are two rows of buttons: "Edit", "Delete", "Previous", "Next" in the first row, and "Move First", "Move Last", "Main Menu" in the second row. At the bottom right of the window is a small keyboard icon and a scroll bar.

Figure 6-11. Data Table

- Click **Edit** to edit a record. The **Edit** button will change to the **Save** button. Click this button to save changes to the record.
- Click **Delete** to delete the current record.
- Click **Previous** to retrieve the previous record.
- Click **Next** to retrieve the next record.
- Click **Move First** to retrieve the first record.
- Click **Move Last** to retrieve the last record.
- Click **Main Menu** to go back to the **Data Entry** screen.

6.7. Import and Export Database

ActiveSync 3.1 uses file filters to automatically handle conversions between desktop computer file formats and Windows CE–based device file formats. The ADOCE control provides the file filter, Adofiltr.dll, to handle the conversion from the Microsoft Access file format (.mdb) to the Windows CE database file format (.cdb). The procedures for manually importing and exporting databases to and from the desktop and Windows CE devices are described below.

6.7.1. Import from Desktop to Windows CE

The steps for importing a database file from a desktop computer to a Pocket PC are as follows:

1. Connect the Windows CE–based device to the desktop computer.
2. Open ActiveSync™ 3.1 and choose **Import Database Tables** from the **Tools** menu.
3. Type the database path and filename in the **Open** dialog box; click **OK**.

4. In the **Import** dialog box that appears, specify the location on the target device where the database will be placed. The default folder is "My Documents."
5. Check a table in the database view window to copy the table to the device. Each table can be expanded and individual fields from that table selected.
6. Selecting the **Default** button will reset the selections to their original state, which includes all tables and fields in a database.
7. Click **OK** to copy the entire database to the **Databases** folder on the device.

6.7.2. Export an ADOCE Database to a Desktop

The steps for exporting an ADOCE database file from a Pocket PC to a desktop computer are as follows:

1. Select **Export Database Tables** from the **Tools** menu in the **Mobile Devices** window.
2. Type the database path and filename of the database to be placed on the desktop computer in the **Location** box.
3. Select all the tables to be copied in the **Select the tables to copy** box.
4. Select the box labeled **Overwrite existing tables and/or data**; click **OK**. Selecting this option causes the converted table from the device to replace the table in the file.
5. Open the database file in Microsoft Access to see the exported database.

6.8. Post-Processing Program

This section describes a post-processing program that serves as a companion program to the AVO field data collection tool described above. Developed as a simple Microsoft Access application, the program can compute the average vehicle occupancy rates and the related statistics from vehicle occupancy data collected by the AVO field data collection tool. The output from the application is a simple one-page report that summarizes the number of vehicles, number of occupants, average occupancy rates, standard deviation of average occupancy rates, and 95% confidence intervals for average occupancy rates. Figure 6-12 shows an example of the summary report. As can be seen, the output summary is stratified by vehicle type and lane number. Each summary table also includes a Total column and a Total row. Each of the total values provides the aggregate statistics for each vehicle type and each lane. The common cell for the Total row and Total column gives the overall statistics for a data collection location.

The name of the application file is **AVO REPORT.MDE**, which is a Microsoft Access executable file. To generate a summary report:

1. Find the **AVO REPORT.MDE** file and double click it to bring up the screen shown in Figure 6-13.
2. Click the **Select Database File** button to bring up the screen shown in Figure 6-14.
3. Specify a MDB data file that contains data tables generated by the AVO field data collection tool. As soon as an MDB data file is selected, all the data tables in the file will be listed on the list box shown Figure 6-13.

- Highlight a data table on this list and click **OK** to generate a summary report for the occupancy data in the selected data table.

Result : Form

Vehicle Occupancy Summary Report

General Information

Analyst: Xin Li Date: 2/05/05 Time: 5:43:52 PM
 Place: Miami-Dade Roadway Name: West Flagler Street Roadway ID: 87001

Number of Vehicles

	Car	Truck	Bus	Other	Total
Lane 1	12	1	1	2	16
Lane 2	16	7	0	0	23
Lane 3	11	3	1	1	16
Lane 4	11	5	0	2	18
Total	50	16	2	5	73

Number of Occupants

	Car	Truck	Bus	Other	Total
Lane 1	20	1	11	3	35
Lane 2	20	16	0	0	36
Lane 3	15	3	9	1	28
Lane 4	22	8	0	5	35
Total	77	28	20	9	134

Average Vehicle Occupancy Rates

	Car	Truck	Bus	Other	Total
Lane 1	1.67	1.00	11.00	1.50	2.19
Lane 2	1.25	2.29	0.00	0.00	1.57
Lane 3	1.36	1.00	9.00	1.00	1.75
Lane 4	2.00	1.60	0.00	2.50	1.94
Total	1.54	1.75	10.00	1.80	1.84

Standard Deviation for Average Vehicle Occupancy Rates

	Car	Truck	Bus	Other	Total
Lane 1	1.03	0.00	0.00	0.50	2.46
Lane 2	0.56	1.28	0.00	0.00	0.97
Lane 3	0.64	0.00	0.00	0.00	1.95
Lane 4	0.95	0.49	0.00	1.50	2.32
Total	0.74	1.03	1.00	1.17	1.95

95% Confidence Interval for Average Vehicle Occupancy Rates

	Car	Truck	Bus	Other	Total
Lane 1	(1.09,2.25)	(1.00,1.00)	(11.00,11.00)	(0.81,2.19)	(0.98,3.40)
Lane 2	(0.98,1.52)	(1.34,3.24)	(0.00,0.00)	(0.00,0.00)	(1.17,1.97)
Lane 3	(0.98,1.74)	(1.00,1.00)	(9.00,9.00)	(1.00,1.00)	(0.79,2.71)
Lane 4	(1.44,2.56)	(1.17,2.03)	(0.00,0.00)	(0.42,4.58)	(0.87,3.01)
Total	(1.33,1.75)	(1.25,2.25)	(8.61,11.39)	(0.77,2.83)	(1.39,2.29)

Record: 1 of 1

Figure 6-12. Output Summary Table

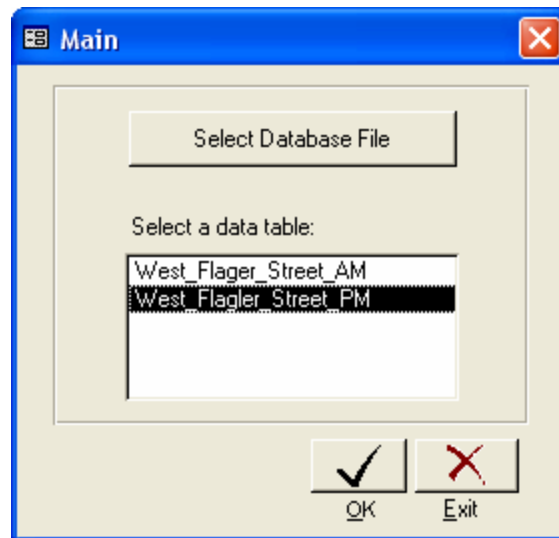


Figure 6-13. Screen for Selecting Database File and Data Tables

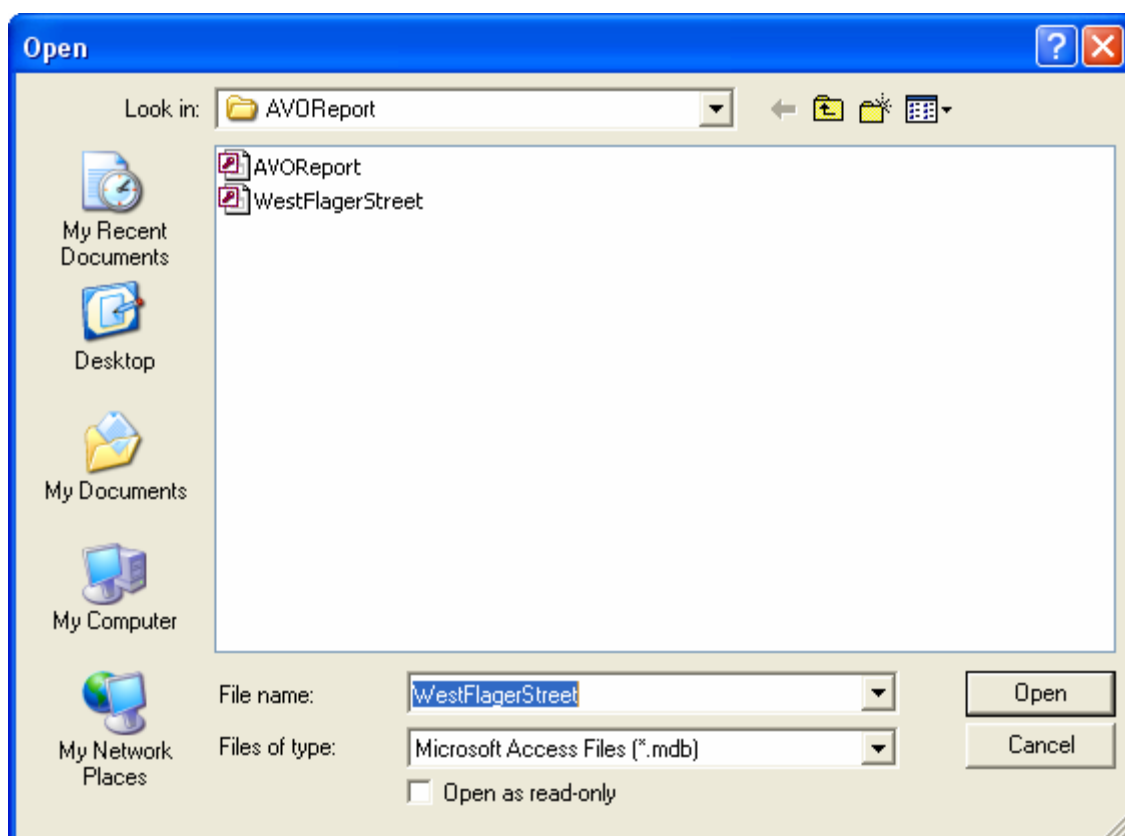


Figure 6-14. Screen for Specifying a Database File

CHAPTER 7

POTENTIAL DATA COLLECTION METHODS

The chapter summarizes several new data collection methods for estimating vehicle occupancy rates. Current research into new methods for collecting vehicle occupancy has mainly been motivated by the needs for automated enforcement of high occupancy and managed lanes. These methods can be divided into photographic and in-vehicle detection.

7.1. Photographic Detection

With advances in image processing and pattern recognition, a number of researchers have explored the use of photographic systems primarily for automated enforcement of HOV lanes. In a proof-of-concept project, the Georgia Tech Research Institute developed a prototype vehicle occupancy system in 1997 for Georgia DOT to help determine the number of persons in a moving vehicle. The prototype system uses digital infrared cameras and infrared strobe lighting to capture views of vehicle interiors. A non-intrusive vehicle detection unit is used to trigger vehicle image capture, as well as collect vehicle volume and classification data.

In 1998, the University of Minnesota performed a study for the Minnesota Department of Transportation to examine the prospect of using mid-infrared and near-infrared cameras to determine vehicle occupancy. It was found that mid-infrared cameras could not produce clear images at highway speeds, while near-infrared cameras worked significantly better. However, the team could not achieve the accuracy and reliability needed for real-world automated enforcement of HOV lanes (University of Minnesota, 1999).

A prototype high occupancy vehicle monitoring system called (HOVMON) was recently developed by Laser Optical Engineering (LOE) in England (Laser Optical Engineering, 2004). The system detects human faces in moving cars without distracting drivers using infra-red cameras. A breakthrough of this system involves the use of a tiny gap in the infra-red spectrum in which light is absorbed by human skin of any color but reflected by hair, clothing and upholstery. This allows the system to reject dummies, large objects and dogs, anything in a fast moving car that could be detected in error by a conventional camera. The system was reported to have been “successfully” tested on the A647 roadway in Leeds for automated HOV monitoring. However, no detailed results of the test have been published.

While photographic systems have achieved some success in counting vehicle occupants and have been shown to have some potential for further improvements, they are not cost-effective (e.g., infra-red cameras are expensive) and are unable to count children in car seats or persons who are lying down in an automobile. They may also be too unreliable during nighttimes or under adverse weather conditions.

7.2. In-Vehicle Detection

7.2.1. In-Vehicle Systems

With the use of new technologies in fields related to transportation, different ways for collecting distinct types of data have been made possible. As such technologies become more affordable and more complete, the capabilities for addressing data collection constraints grow significantly. Many of the new methodologies that can be used for the collection of traffic data have derived from the implementation of different aspects from a relatively recent development in Intelligent Transportation Systems (ITS). Due to the fact that ITS relies heavily on the use of electronic components that monitor and communicate the performance and other parameters from vehicles and traffic, the capacity of using these electronic devices for obtaining traffic related parameters has become viable. In the case of collecting vehicle occupancy data, the use of different types of in-vehicle electronics in combination with wireless communications can be a future source of occupancy data. Such use of in-vehicle sensing instruments along with communication capabilities are being used in ITS areas such as Commercial Vehicle Operations and Emergency Management Systems among others.

As vehicle manufacturers and transportation advocates have learned about of the benefits of ITS applications, more vehicles with different equipment of this nature are being manufactured. Currently, different types of electronics that address safety issues are being installed in new vehicles. One of the main safety-oriented applications for the electronics that have been developed is the secure deployment of air bags. Car manufacturers have identified new systems that enable air bag deployment through the detection of passengers in a vehicle. In this manner the vehicle's central computer is capable of sensing the presence of persons within the car in order to deploy certain air bags or not. By having the capacity of detecting the number of occupants in a motor vehicle, it is also possible to use such information for data collection purposes. These techniques include the use of in-vehicle cameras which are capable of producing three-dimensional images in order to detect the passengers' location and position.

Seat occupancy detection can also be achieved with the use of detecting mats that can perceive the presence and weight of a person. Additionally electrodes that are embedded into the back of each seat can yield occupancy since these electrodes emit an electric field that is altered by the presence of a person.

Aside from the technologies used for improved deployment of air bags, other sensing capabilities within vehicles can produce passenger occupancy information. For instance, some vehicle manufacturers are producing vehicles that indicate the number of seat belts that are fastened in order to produce a surrogate measure of passenger occupancy. Moreover, some vehicle electronics developers are also using infrared sensors to detect the body heat emitted by the passengers which can also be used for occupancy purposes.

7.2.2. Information Transmission

Once the number of vehicle occupants has been identified by an in-vehicle system, the information has to be transmitted to a collecting station that can relate the data to a given

roadway segment or location and so that these records can be stored in a database. One of the manners in which this can be achieved is through the use of transponders and detectors with dedicated short-range communication capabilities. In this manner, short-range communication stations can be located on the side of the road in order to collect the information from each vehicle without affecting travel speed. An example of the use of this kind of transponders is the SunPass system, which is currently employed for toll collection purposes.

Another way in which data transmission can take place is through the use of Mobile Mayday Systems, which is an application of ITS technologies. Mobile Mayday Systems make use of in-vehicle electronics that provide sensing capabilities and are supported by a customer center (known as a telematics service provider) that has access to the information retrieved by the vehicle's sensors. In the event of an emergency the support center can provide assistance to the passengers traveling in the vehicle and dispatch emergency care. The customer center can establish data transmission through a wireless network and is also capable of determining the vehicle's location. Such wireless networks are provided by the communication platforms of common wireless/cellular communications carriers. The communication link that is available between vehicles and support centers within the Mobile Mayday Systems frame can be used for collecting information that is detected by the sensing electronics. For instance, occupancy data could be collected by the telematics service provider throughout specific locations and at different times of the day. A database can be developed in order to establish occupancy rates at different roadway facilities.

7.3. Remarks

The main issue that arises from these potential methodologies for collecting vehicle occupancy data is that they rely on vehicle technologies that are relatively new and that have not widely spread out yet. In fact, it is not possible to determine when these technologies will have been implemented extensively enough as to produce reliable, unbiased data. However, traffic professionals and vehicle manufacturers expect a significant growth of the number of vehicles with the capabilities required for this sort of collection, due to the important role that they can develop in safety and other transportation issues that can be addressed.

CHAPTER 8

STUDY GUIDELINES

A major conclusion from the previous chapter is that fully automated methods of vehicle occupancy data collection are still largely a distant reality due to technological, cost, and institutional barriers. Methods with a lesser degree of automation that record and analyze occupancy data electronically remain the current methods of choice. The guideline presented in this chapter is thus focused on only the manual counting methods. As mentioned, vehicle occupancy data derived from existing databases are limited by their existing conditions and limitations.

8.1. General

- The application objective forms the basis for all study design considerations. Once the objective is set, inappropriate collection methods can be screened out.
- The vehicle types to be included should be clearly defined at the beginning of a study. For general vehicle occupancy studies, data should be collected for passenger vehicles (including pick-ups, light trucks, vans, and motorcycles), heavy trucks, and buses.
- Data should generally be collected on Tuesdays, Wednesdays, and Thursdays. Major holidays should be excluded.
- Confining to the ability to precisely count occupants in buses, existing ridership database or onboard survey can be used to obtain much more accurate data than could be gained from roadside observation for transit vehicle.
- If bus passengers are to be counted, the number of occupants can be estimated approximately: one-quarter full (or about 10 occupants), one-half full (or about 20 occupants), and full (or about 40 occupants). The approximate figure of occupants varies depending on typical bus sizes served in the area.

8.2. Determining Time Periods

- Given that vehicle occupancy rates vary by time of day, it is important to select different critical time periods for analysis. This would typically include the three peak periods: AM peak period (7:00 to 9:00 AM), midday period (11:00 AM to 1:00 PM), and PM peak period (4:00 to 6:00 PM).
- The weighted average of the three periods should give an approximation that is reasonably close to a full daytime estimate.
- Noticeable degradation in observer's performance appears after continuous counting goes beyond a certain time period. A maximum of two hours should be assigned for each observation time period.

8.3. Determining Sampling Plan

8.3.1. Number of Observation Sessions

- The sample size required is a function of the desired error tolerance level and the desired confidence level. Obviously, the available funds will be the bottom-line constraint on the sample sizes.
- The confidence level should generally be set at 95%.

Area-wide Studies

- The commonly accepted error tolerance for area-wide studies range from ± 0.01 to ± 0.05 , depending on the desired level of accuracy. A commonly accepted level of error tolerance for area-wide studies is ± 0.03 .
- If no previous similar survey has been conducted or there is not a strong reason to select another value, the standard deviations derived from the data collected by URS Corporation in 1996 and 1997 can be used to estimate the composite standard deviation as follows:

$$\begin{aligned}\sigma_h &= (\sigma_{Lh}^2 + \sigma_{Sh}^2 + \sigma_{Wh}^2 + \sigma_{Oh}^2)^{1/2} \\ &= (0.076^2 + 0.068^2 + 0.008^2 + 0.006^2)^{1/2} \\ &= 0.102\end{aligned}$$

where σ_h = composite standard deviation of AVO in stratum h ,

σ_{Lh} = standard deviation of AVO across link-days within a season in stratum h ,

σ_{Sh} = standard deviation of AVO across seasons in stratum h ,

σ_{Wh} = standard deviation of AVO across time periods within a day in stratum h , and

σ_{Oh} = standard deviation of AVO due to observer error in stratum h .

- Using this composite standard deviation at the ± 0.03 tolerance level and 95% confidence level, the number of link-days to be collected is:

$$N_h = \left(\frac{Z \times \sigma_h}{T_h} \right)^2 = \left(\frac{1.96 \times 0.102}{0.03} \right)^2 = 44$$

Thus, a minimum sample size of 44 link-days is required for AM peak, midday, and PM peak periods, respectively. The resulting total minimum sample size is 132 observation sessions within a year.

Corridor Studies

- The commonly accepted tolerance for corridor studies range from ± 0.03 to ± 0.09 , depending on the desired level of accuracy. A commonly accepted level of error tolerance is ± 0.06 .

- If no previous similar survey has been conducted or there is not a strong reason to select another value, the standard deviations derived from the data collected by URS Corporation in 1996 and 1997 are used to estimate the composite standard deviation as follows:

$$\begin{aligned}\sigma_h &= \left(\sigma_{Dh}^2 + \sigma_{Sh}^2 + \sigma_{Wh}^2 + \sigma_{Oh}^2 \right)^{1/2} \\ &= \left(0.028^2 + 0.068^2 + 0.008^2 + 0.006^2 \right)^{1/2} \\ &= 0.074\end{aligned}$$

where σ_h = composite standard deviation of AVO in stratum h ,

σ_{Dh} = standard deviation of AVO across days within a season in stratum h ,

σ_{Sh} = standard deviation of AVO across seasons in stratum h ,

σ_{Wh} = standard deviation of AVO across time periods within a day in stratum h , and

σ_{Oh} = standard deviation of AVO due to observer error in stratum h .

- Using this composite standard deviation at the ± 0.06 tolerance level and 95% confidence level, the number of days to be collected is:

$$N_h = \left(\frac{Z \times \sigma_h}{T_h} \right)^2 = \left(\frac{1.96 \times 0.074}{0.06} \right)^2 = 6$$

Thus, data should be collected on 6 randomly-selected dates on the corridor. This means a total number of 18 observation sessions will be needed for each location to cover the three peak periods of a day. If the study is to be conducted for only one season, the seasonal variation term can be ignored. A sample of one day of data collection (i.e., a total number of 3 observation sessions) per location is needed to achieve the same level of accuracy.

8.3.2. Sampling Procedure

After the number of observation sessions is determined, the observational locations and days can be selected from a list of link-days to produce an unbiased estimate of AVO. A random sampling technique on highway links and dates applied to allocate the required observation sessions. The link list includes all possible links for the transportation network in the study area.

Area-wide Studies

The basic highway network used to define the survey population can be a computer-coded regional highway network. The FSUTMS' computerized file for the highway network can be used to identify detailed roadway segments within the study area. The highway network file contains sufficient information to serve as a complete link list for determining the links to be observed.

To avoid potential bias resulting from failing to select high-volume segments and to take into account the variations in link length, the links should be designed to be sampled systematically

with a probability proportional to the VMT of each link. The following procedure is extended from a sampling scheme developed by Levine and Wachs (1994) for area-wide studies:

- The VMT for each link is derived by multiplying the daily traffic volume with the length of a link. The list of all links for the study area is sorted by the VMT in the descending order.
- The VMT are summed to yield the total VMT of a study network. A cumulative count of VMT for all links is taken by starting at the first link in the list. The cumulative VMT for the first link is the VMT for itself. For the second link on the list, the cumulative VMT is the sum of VMT for the first and second links. For the third link on the list, the cumulative VMT is the sum of VMT for the first, second, and third links. The process is continued until the last link of which the cumulative VMT is equal to the total VMT of the whole network.
- A link sampling interval (i) is computed by dividing the total VMT by 44. A random number (s) between 1 and i is generated to select the first observation site on the link with the cumulative VMT containing s within its range. The second observation site is selected on a link with a range of cumulative VMT containing a value of sum of i and s . This process is continued by adding i to the previous sum and selecting the link in which the number falls within the range. This continues until the required 44 links for a single temporal stratum are selected.
- Starting from the first day, each day within the year is given a unique and sequential index number starting at 1 and continuing to 365. Measurements are to be taken on Tuesdays through Thursdays while excluding holidays. The list of eligible dates will include the maximum days of 156, 157, or 158, if no holidays occur on the measurement dates.
- A three-digit number is randomly generated to select the date. If the number matches one of the index numbers of eligible dates, the date is selected and assigned to one of the selected links in the same order as the links are selected. If the number does not match any one on the date list, this number should be ignored and another random number is generated to identify the date to be sampled. This process is continued until each of the 44 links has been assigned a particular date.
- For all observation links, three measures a day are carried out in the morning between 7:00 and 9:00 AM, in the midday between 11:00 AM and 1:00 PM, and in the afternoon between 4:00 and 6:00 PM. Therefore, the 44 observation links yield a sample size of 132 individual observation observations.

Corridor Studies

For a corridor study, local knowledge is necessary to define the locations to be selected. Links with higher volumes are more likely to be the locations of primary interest. Since the locations are selected judgmentally, the studies require only a list of eligible dates to be randomly selected for each location. The sampling procedure for selecting the dates for area-wide studies is directly applicable to corridor-level survey programs. One exception is that a separate sample of

dates will be selected for each location in this case. The procedure for determining the observation sessions for corridor studies is as follows:

- Each day within the year is indexed successively from 1 to 365. Remove holidays and days other than Tuesdays through Thursdays to create an eligible list of dates.
- Randomly generate a three-digit number to select a date for each location in the same order as the locations are selected. If the number matches one of the index numbers of eligible dates, the date is assigned to one of the selected location. If the number does not match, another random number is generated to identify the date to be sampled. This process is continued until 8 days has been assigned to the location.
- The previous step is repeated for the next location until no more location can be selected.
- For each observation location, vehicle occupancy is measured during three time periods on a day (7:00 to 9:00 AM, 11:00 AM to 1:00 PM, and 4:00 to 6:00 PM) for 8 assigned dates. This will yield a sample size of $8N$ individual observation periods, where N is the number of selected locations.

8.4. Field Operation Design

8.4.1. Locate Counting Sites

A field check of each selected link is necessary to find an appropriate site to conduct the counts before the regular observations begin. Ideally, locations that can facilitate observation of approaching vehicles and provide adequate protection for the observers should be selected. Experience gained from the 1997 FDOT study by URS suggested that the following factors should be considered for roadside observations:

- Clear view of occupants in approaching vehicles on each lane of travel.
- Protection from errant or out-of-control vehicles.
- Minimize disruption to normal traffic flow due to the presence of observers.
- Availability of parking, shelter, and restroom.

Several tips for selecting counting sites with clear view of occupants were also suggested:

- The counting site should be close to the elevation of the road observed.
- Counting sites on overpasses, ramps, and weaving sections should be avoided.
- Mid-block location on the downstream from intersection should be selected.

After a counting site is selected, a simple sketch of the road segment that identifies the site location can be documented for the location-specific survey information. Other information such as location description/ID, facility type, number of lanes, date of observation, parking restriction, and other information relevant to the observation can also be included.

8.4.2. Work Plan

A work plan on the basis of the developed sampling plan should be prepared for collecting the occupancy data.

- Manual observations during the daylight hour will use the systematic short-count method. Observers could count on a single lane or multiple lanes at a time with scheduled shifts from lane(s) to lane(s), which is primarily a function of traffic volume but is also related to the degree of visibility of vehicle occupants. On most roadways, traffic flow is such that one person cannot correctly record data from more than one lane. Each observer assigned to collect data for a link for one lane at a time in one travel direction is thus preferred.
- For each lane, observation is to be taken for 10-minute interval. Lane is counted successively after each for two 10-minute intervals followed by a 5-minute interval reserved for reviewing the counts taken and for rest.
- To produce an adequate representation of vehicle occupancy at the location, counting both directions and all lanes is important. Hence, a minimum of two observers are needed for each observation session. However, the dates are sampled randomly so that an individual date could be assigned to multiple links, which require additional observers. Further, the availability of backup personnel is apparent if any one of the crew members is unable to participate in the observation.
- A work schedule for the entire year that indicates the locations and days are to be observed and with observers' name should be prepared. Before starting the vehicle occupancy data collection, the work schedule is given to each observer for the entire survey period with the location-specific survey information obtained from field check and the instructions summarized during the training session.

8.4.3. Personnel Training

- Prior to the actual data collection, field observers should be properly trained on the procedures to be used and given a thorough background on the study. The instruction is to enable observers to conduct the counts in a consistent and reliable manner.
- Vehicle types are discussed with the crews to clear the distinctions among them.
- The data collection equipment used for the study should be explained on how to record occupancy data.
- The crews should be instructed not to guess the number of occupants if they cannot see properly into a vehicle. With the typically large number of vehicles being sampling and relatively stable occupancy, the accuracy of a vehicle occupancy count is affected insignificantly by missing a few vehicles.

- Specific instructions for resolving potential problems of observation in the field should be made. It could guide the crews in making a proper decision in the event that an unusual situation arises.

8.4.4. Data Collection Tool

- The battery life presents a major concern for application. Either several backup standard batteries or one extended battery that could have sufficient battery life for a day count is necessary for one pocket PC. If multiple standard batteries are to be acquired, one battery should only be used for one observation session. Alternatively, one extended battery that is able to last long enough should be selected.
- Field crews should use care with the Pocket PCs and protect them from the weather.

8.5. Data Collection Plan

- Advance notice with the appropriate public agencies before data collection activities begin is recommended. This helps prevent unnecessary explanations to individual officers and reduce disruption of counting and loss of data.
- Upon arriving at a site, two observers assigned for different directions using the pocket PCs start with inputting roadway information and creating new files for the collection. The good file naming convention will prove to be a great help in the later stage of data analysis. A file name should include at least location, direction, lane number, date, and time period (AM, MD, and PM).
- When the specified observation time is hit, observers begin to count the number of different category of vehicles and record data by short-count technique for two hours per observation session. Each category of vehicles used for recording occupancy data is referred to the vehicle type/number of occupants combinations. Observers record a category for each vehicle by simply clicking the data entry screen.
- Traffic should be observed on each lane for both directions of a selected link. Depending on the number of lanes, time invested in collecting the data on each lane within one observation session could be different. If there is only one lane in the observed direction, the field crew count for two 10-minute intervals, rest for 5 minutes, and continue counting for the following two 10-minute intervals. If there are three lanes in the observed direction, the field crews could start counting, for example, in the inside (median) lane (lane #1) for one 10-minute interval, shift to the next lane (lane #2) to count for one 10-minute interval, and rest for 5 minutes. Then the count continues in the outside (curb) lane for one 10-minute interval, shifts back to lane #1 again for one 10-minute interval, and rest for 5 minutes. The process repeats until the count for the last interval of a two-hour period has been conducted, one observation session is completed. Table 8-1 illustrates the AM peak period schedule of observation taken place in each lane for one travel direction.

Table 8-1. Schedule of Lane Observation in AM Peak Period

Times	Number of Lanes per Direction					
	1	2	3	4	5	6
7:00 – 7:10	Lane #1	Lane #1	Lane #1	Lane #1	Lane #1	Lane #1
7:10 – 7:20	Lane #1	Lane #2	Lane #2	Lane #2	Lane #2	Lane #2
7:20 – 7:25	Break					
7:25 – 7:35	Lane #1	Lane #1	Lane #3	Lane #3	Lane #3	Lane #3
7:35 – 7:45	Lane #1	Lane #2	Lane #1	Lane #4	Lane #4	Lane #4
7:45 – 7:50	Break					
7:50 – 8:00	Lane #1	Lane #1	Lane #2	Lane #1	Lane #5	Lane #5
8:00 – 8:10	Lane #1	Lane #2	Lane #3	Lane #2	Lane #1	Lane #6
8:10 – 8:15	Break					
8:15 – 8:25	Lane #1	Lane #1	Lane #1	Lane #3	Lane #2	Lane #1
8:25 – 8:35	Lane #1	Lane #2	Lane #2	Lane #4	Lane #3	Lane #2
8:35 – 8:40	Break					
8:40 – 8:50	Lane #1	Lane #1	Lane #3	Lane #1	Lane #4	Lane #3
8:50 – 9:00	Lane #1	Lane #2	Lane #1	Lane #2	Lane #5	Lane #4

- During the observation period, observers should record data for as many eligible vehicles as they could observe. Recording errors can be made in the observation intervals. During the 5-minute break, observers should review the recorded data to preliminarily alleviate inaccuracy. Any unusual circumstances surrounding the counts involved should be recorded and filed for future reference during the analysis phase.
- The schedule of selected locations and days for data collection should be strictly followed. However, the counts will be sometimes missed for some reason. Data collection during periods of inclement weather increases the possibilities of resulting in counting errors and producing a biased sample. Data should not be collected during these periods. Occasionally, the observation interference arises due to sickness or otherwise unable to participate in the observation. As a result, observers will miss their designated observation sessions. Missed counts can then be made up on the same day in the succeeding week during the survey period.
- The supervisor should conduct random field checks during the period of observation to ensure adherence to the schedule and proper conformance to the collection procedures. A spot check of upstream count can also be used to assess the quality of the data being collected by the field crew. Observers who collect data improperly can be identified and corrected. Since the use of pocket PC could provide immediate feedback on the types of errors observers have made, it can relieve partial duty of supervisor's.

8.6. Data Analysis

- The data files from each location are transferred into a desktop or laptop PC. The data saved in a database format can be easily retrieved by any popular spreadsheet programs. A back-up files should be saved to prevent accidental loss of data.

- For each of observation time period, separate stratum-specific AVO estimate and composite standard deviation can be calculated by:

$$AVO_h = \frac{\sum_i P_i}{\sum_i V_i}$$

$$\sigma_h = \left[\frac{N_h \times \sum_i (P_i - AVO_h * V_i)^2}{\left(\sum_i V_i \right)^2} \right]^{1/2}$$

where P_i and V_i are the factored number of occupants and vehicles counted in session i , respectively; N_h is the number of sessions of data collection within each stratum. The factored number of occupants and vehicles for an observation session are two-hour estimated equivalents. The field data collected using short-count techniques should be expanded to a two-hour base. Based on the number of lanes for one direction, a separate expansion factor may be needed for each lane as shown in Table 8-2. For a roadway with multiple lanes in each direction, either the inside or outside lane could be assigned as Lane #1. However, the definition of Lane #1 should be used consistently for all selected multiple-lane facilities throughout entire the survey program.

Table 8-2. Expansion Factors for Each Lane

Lane #	Number of Lanes per Direction					
	1	2	3	4	5	6
1	12/10 = 1.2	12/5 = 2.4	12/4 = 3	12/3 = 4	12/2 = 6	12/2 = 6
2	—	12/5 = 2.4	12/3 = 4	12/3 = 4	12/2 = 6	12/2 = 6
3	—	—	12/3 = 4	12/2 = 6	12/2 = 6	12/2 = 6
4	—	—	—	12/2 = 6	12/2 = 6	12/2 = 6
5	—	—	—	—	12/2 = 6	12/1 = 12
6	—	—	—	—	—	12/1 = 12

- From above, three time period-specific AVO estimates and the corresponding composite standard deviations for a region or for each of the selected locations can be derived for the area-wide or corridor-level studies, respectively.
- The actual precision of the estimates of time period-specific AVO obtained from the survey can then be computed with the following formula as:

$$T_h = \left(\frac{Z^2 \times \sigma_h^2}{N_h} \right)^{1/2}$$

On account of the composite standard deviation obtained from survey is usually somewhat different from the estimate used to develop the sampling plan, the actual precision level will also be different from the desired precision for the study. For example, tolerances of ± 0.03 and ± 0.06 with 95 percent level of confidence for each observation time period are selected here for the desired precision of regional and locational estimates, respectively. Notwithstanding the actual precision is slightly higher or lower than the expected, the survey AVO estimates will still be valid if the composite standard deviation has been accurately estimated.

- An overall measure representing multiple strata can be developed by combining the estimates from the samples in each stratum through appropriate weighting techniques. For a daylight-hour measure, the AVO estimate, composite standard deviation, and actual level of precision are produced by weighting the respective time period-specific counterpart as follows:

$$AVO = \sum_h (W_h \times AVO_h)$$

$$\sigma = \left[\sum_h (W_h^2 \times \sigma_h^2) \right]^{1/2}$$

$$T = Z \times \left[\sum_h \left(W_h^2 \times \sigma_h^2 / N_h \right) \right]^{1/2}$$

where W_h is the relative proportion of the total VMT or traffic volumes. For all stratums, the sum of weight must add to 1.0.

- The weight is calculated for different occupancy study level as:

Area-wide Level:

$$W_h = \frac{VMT_h}{VMT} = \frac{\sum_i (V_{hi} \times L_i)}{\sum_h \sum_i (V_{hi} \times L_i)}$$

Corridor Level:

$$W_h = \frac{\sum_d V_{hd}}{\sum_h \sum_d V_{hd}}$$

where VMT_h = total VMT of all selected links for observation time period h ,
 VMT = total VMT of all selected links,
 V_{hi} = factored number of vehicles for observation time period h at location i ,
 L_i = link length of the selected link i , and
 V_{hd} = factored number of vehicles for observation time period h on day d at one location.

As a result, one daytime AVO along with the corresponding composite standard deviation and actual precision level for a region or for each of selected locations are derived for the area-wide or corridor-level studies, respectively.

- If an agency is interested in estimating the daytime AVO and its associated statistics across all selected locations for corridor studies, the formula used for computing the daytime counterparts for one location can be used again. In this case, each location instead of each observation time period stands for a different stratum. The weight is thus computed as:

$$W_i = \frac{\sum_h \sum_d V_{hd}}{\sum_i \sum_h \sum_d V_{ihd}}$$

where V_{ihd} is the factored number of vehicles for observation time period h on day d at location i .

CHAPTER 9

SUMMARY AND RECOMMENDATIONS

The Intermodal Surface Transportation Efficiency Act (ISTEA) of 1991 expanded the role of vehicle occupancy data by requiring continued monitoring of congestion management strategies, many of which emphasize more person movements than vehicle movements. With this increasing need for vehicle occupancy data comes the need to examine and reexamine the ways in which these data have been, and will be, collected. This report has reviewed the existing methods of vehicle occupancy data collection, examined issues related to geographic, temporal, and vehicle coverage design of occupancy data collection, and presented study guidelines for performing occupancy data collection as well as analyzing occupancy data.

Existing methods for vehicle occupancy data can be grouped into those that are collected in the field for the sole purpose of computing vehicle occupancy rates and those that are collected for other purposes, but may be used to estimate vehicle occupancy rates. In general, the field collection methods are more suitable for collecting data at the site-specific or corridor level, while the existing data methods are more suitable for area-wide or regional studies. Field data collection methods are more commonly used because they can be tailored to specific application needs in terms of location, sample size and accuracy, etc. Current field collection methods include roadway windshield, carousel, and video surveillance. Existing databases that can be used to generate vehicle occupancy include crash records, vehicle and household surveys, Census Transportation Planning Package (CTPP), Nationwide Personal Transportation Survey (NPTS), etc. The accuracy and scope of applications of existing data are constrained to what have been collected.

Important design considerations for vehicle occupancy data collection include geographic and temporal coverages, facility types, collection cycles, locations, and vehicle types. The general geographic units include site-specific, corridor, and area-wide. Temporal variability in average vehicle occupancy (AVO) estimates is a common issue across all data collection methods. Existing studies show significant variations in vehicle occupancy rates by time-of-day, day-of-week, and season-of-year. Different types of roads typically have different occupancy levels. For example, roadways of the higher functional hierarchy would typically be expected to have lower AVO. It is important to thus sample all roadway types in order to generate a representative estimate of regional vehicle occupancy. AVOs can also differ from one location to another. The spatial variations of AVO are related to the distribution of household types and work places. It is thus necessary to select many different locations in order to measure AVO variability adequately.

The vehicle types included in a data collection are determined by the purpose for which the data are to be collected. Different study purposes may utilize different criteria for interpreting AVO. In most vehicle occupancy studies, only data from passenger vehicles or light vehicles (private passenger automobiles, pickups, vans, recreational vehicles and motorcycles) are usually counted. Buses are typically excluded or counted separately because it is difficult to count all the occupants using the roadside windshield method or the carousel method. Trucks are generally excluded because they are used mainly for goods movement and have little to do with people mobility.

A typical procedure for vehicle occupancy studies include defining study objectives, selecting a data collection method, establishing a sampling plan, comparing costs, performing random sampling, and computing the AVO. The first step in conducting a vehicle occupancy study is to define the study objectives, which form the basis for further study planning and design. Once the objectives are defined, inappropriate collection methods can be screened out. The remaining methods can then be considered based on cost comparisons. After the data collection method is selected, the actual survey procedures and the sampling plan can be designed. Appropriate sampling plans for site-specific, corridor, and regional studies were presented and discussed in detail. Statistical sound collection techniques are a major concern and should be properly designed. A sound sampling procedure is needed to ensure that the AVO estimates meet the desired precision with a certain level of confidence. In this study, standard deviations for several factors identified by Ferlis (1981) were derived from vehicle occupancy data collected in a Florida statewide study from 1996 to 1999. In the absence of standard deviations from local data, these standard deviations are recommended for use in determining the appropriate sample sizes for corridor and area-wide studies.

As part of this study, a user-friendly software system called FAVORITE (**F**lorida **A**ccident **V**ehicle **O**ccupancy **R**ate **I**nforma**T**ion **E**stimator) was developed to estimate the occupancy rates from multiple years of crash records on the Florida state roadway system. The system can estimate occupancy rates for select roadway segments, corridors, or regions for specific time periods for different types of vehicles. FAVORITE comes with the 1990-2002 accident data and includes passengers for up to two vehicles for each accident. In addition, the database also includes a number of variables that can be used for various analyses, including district, county, hour of day, day of week, month of year, vehicle type, facility type, area type, and crash severity. Because the system makes use of a comprehensive statewide database, it can potentially be a highly cost-effective means for monitoring statewide, regional, and site-specific vehicle occupancy trends. While a preliminary assessment of the system show outputs that are consistent with the expected data trends, an enhanced version of the system would require additional research that takes into account over- and under-involvement of certain types of accidents, e.g., young drivers and higher-occupancy vehicles are more prone to traffic accidents. Failing to correct for these factors could produce biased AVO estimates.

To facilitate field data collection and processing, an automated field data collection tool designed for use with a handheld Pocket PC was developed as part of this study. The tool eliminates the need for manual data post-processing by allowing the user to make use of the touch-screen interface on a Pocket PC to record the number of passengers for different types of vehicles and different lane numbers. In addition, a companion program that can calculate the average occupancy rates from the data collected from the automated tool was also developed. The research team first investigated the possibility of applying voice recognition technology in lieu of screen input on a Pocket PC. This was found to be impractical due to the lack of an applicable commercial voice recognition system that could work well with the Windows CE operating system used by Pocket PCs. It was also found that traffic noise in the field could interfere with voice recording, resulting in erroneous data.

Current research into new methods for collecting vehicle occupancy has mainly been motivated by the needs for automated enforcement of high occupancy and managed lanes. These methods can be divided into photographic and in-vehicle detection. With advances in image processing and pattern recognition, a number of researchers have explored the use of photographic systems primarily for automated enforcement of HOV lanes. While photographic systems have achieved some success in counting vehicle occupants and have been shown to have some potential for further improvements, an operational, cost-effective system for occupancy data collection does not currently exist. The use of different types of in-vehicle electronics in combination with wireless communications can be a future source of occupancy data. These include systems that can detect the presence and weight of a passenger for safer deployment of air bags; in-vehicle cameras to detect passenger location and position (also for safer deployment of air bags); automated detection of seat belt usage by each passenger (already implemented for the driver); etc., all of which hold potential to provide more accurate vehicle occupancy information than any of the existing methods.

One finding of this study was that fully automated methods of vehicle occupancy data collection are still largely a distant reality due to technological, cost, and institutional barriers. Methods with a lesser degree of automation that record and analyze occupancy data electronically remain the current methods of choice. As part of this study, a set of study guidelines was developed for the manual counting methods for corridor and area-wide studies. The guidelines address design issues related to time periods, sampling plan, field operation, work plan, personnel training, use of data collection tool and equipment, data collection plan, and data analysis.

REFERENCES

1. Asante, S. A., L. H. Adams, J. J. Shufon, and J. P. McClean, "Estimating Average Automobile Occupancy from Accident Data in New York State," *Transportation Research Record* 1553, 1996, pp. 115-123.
2. Barton-Aschman Associates, Inc., *Vehicle Occupancy Determinators*, Final Report Prepared for Arizona Department of Transportation, August 1989.
3. Benke, R. J. and R. A. Sjoberg, "Auto Occupancy Parameter Variations," *Traffic Engineering*, 1977, pp. 21-27.
4. Chang, L. Y. and F. L. Mannering, "Predicting Vehicle Occupancies from Accident Data: An Accident Severity Approach," *Transportation Research Record* 1635, 1998, pp. 93-104.
5. Federal Highway Administration, *Estimating Auto Occupancy: A review of Methodology*, September 1972.
6. Federal Highway Administration, *1977 Nationwide Personal Transportation Study*, 1980.
7. Ferlis, R. A., "Field Data Collection and Sampling Procedures for Measuring Regional Vehicle Classification and Occupancy," *Transportation Research Record* 701, 1979, pp. 1-6.
8. Ferlis, R. A., *Guide for Estimating Urban Vehicle Classification and Occupancy*, Final Report Prepared for FHWA Office of Highway Planning, March 1981.
9. Fisher, R. F., G. J. Williams, and J. P. Boyd, "Atlanta Vehicle Occupancy Monitoring," *Transportation Research Record* 779, 1980, pp. 27-32.
10. Grush, W. and B. Gross, *Vehicle Occupancy: A comparison of Household Survey Data and Roadside Observation Data*, FHWA Office of Highway Information Management, September 1995.
11. Harrold, C. W., *The Carousel Data Collection Methodology: Comparison with Other Data Collection Methodologies*, Task Report, 1997.
12. Heidtman, K., B. Skarpness, and C. Tornow, *Improved Vehicle Occupancy Data Collection Methods*, FHWA Office of Highway Information Management, April 1997.
13. Hupp, C. and C. Palombo, "Evaluation of the FHWA Vehicle Classification and Auto Occupancy Sampling Manual," Presented at Transportation Research Board Annual Meeting, January 1980.
14. Kuzmyak, J. R., *Vehicle Occupancy*, 1981.

15. Laser Optical Engineering, Ltd, "Laser Optical Engineering Vehicular Occupancy Monitor VOM-10", <http://www.laseroptical.co.uk/documents/vehicular-occupancy.pdf>, August 9, 2004.
16. Law, C. S., *A Photographic System for Auto Occupancy Counting and Vehicle Identification Vol. I, Principal Finding and Conclusion*, Final Report Prepared for FHWA Offices of Research and Development, November 1981.
17. Levine, N. and M. Wachs, *Methodology for Vehicle Occupancy Measurement: Part I, Operations Manual for Measuring Vehicle Occupancy – Instructions and Procedures*, October 1994.
18. Levine, N. and M. Wachs, *Methodology for Vehicle Occupancy Measurement: Part II, Technical Issues in Vehicle Occupancy Measurement*, October 1994.
19. Levine, N. and M. Wachs, "Factors Affecting Vehicle Occupancy Measurement," *Transportation Research, Part A, Policy and Practice*, Vol. 32A, No. 3, 1998, pp.215-229.
20. Liu, Q. X. and H. Desai, "What do We Know about Vehicle Occupancy? A Statewide Pilot Study in Florida," *Compendium of Technical Papers*, ITE, 1998.
21. McElhaney, D. R., "An Update of Road Traffic Monitoring in the United States," *International Conference on Road Traffic Monitoring and Control*, London, 1994, pp. 111-114.
22. Metropolitan Council of the Twin Cities Area, *Screenline Auto Occupancy Study Data and Methodology*, 1992.
23. Mid-America Regional Council (MARC), *2002 Vehicle Occupancy Study for the Kansas City Metropolitan Area*, Final Report, 2002.
24. Ohstrm, E. G., P. R. Stopher, "Automobile Occupancy, Vehicle Trips, and Trip Purpose: Some Forecasting Program," *Transportation Research Record* 987, 1984, pp. 8-13.
25. Pamanikabud, P., "Morning Vehicle Occupancy Rate Entering the City Center," *Urban Transportation and Environment*, 2000, pp. 899-903.
26. Pavlidis, I., B. Fritz, P. Symosek, N. P. Papanikolopoulos, V. Morellas, and R. Sfarzo, "Automatic Passenger Counting in the High Occupancy Vehicle (HOV) Lanes," *Proceedings of the 9th ITS America Meeting*, 1999.
27. Pietrzyk, M. C., *Demonstration of Video-Based Technology for Automation of Traffic Data Collection*, Center for Urban Transportation Research, University of South Florida, January 1996.

28. Reed, G. L., H. Desai, and L. B. Krueger, "Determination of Vehicle Occupancy for Small and Medium Sized Areas," Sixth National Conference on Transportation Planning for Small and Medium-Sized Communities, Spokane, Washington, 1998.
29. Roach, W. T. and M. Lester, *Techniques for Monitoring Auto Occupancy and Seattle Area Research Results*, Final Report Prepared for FHWA Offices of Research and Development, September 1978.
30. Rutherford, G. S., R. K. Kinchen, and L. N. Jacobson, "Agency Practice for Monitoring Violations of Highway-Occupancy-Vehicle Facilities," *Transportation Research Record 1280*, 1990, pp. 141-147.
31. Ryan, T. A., C. F. Davis, and B. G. Taylor, *Auto Occupancy Variations*, Prepared for the 63rd Annual Meeting of the Transportation Research board, August 1983.
32. Ulberg C. G., *Vehicle Occupancy Forecasting*, Final Report Submitted for Washington State Department of Transportation, February 1994.
33. Ulberg, C. G and E. McCormack, "Accuracy and Other Factors Affecting a Continuous Vehicle Occupancy Monitoring Program," *Transportation Research Record 1206*, 1988, pp. 35-47.
34. Ulberg, C. G. and E. McCormack, *Auto Occupancy Monitoring Study*, Final Report Submitted for Washington State Department of Transportation, June 1988.
35. University of Minnesota Department of Computer Science, *Automatic Passenger Counting in the HOV Lane*, Prepared for the Minnesota Department of Transportation, June 1999.
36. URS Consultants, Inc., *Vehicle Occupancy Data Collection and Analysis*, Final Report Prepared for the Florida Department of Transportation, Systems Planning Office, December 1997.
37. Versave, J. and B. Boucke, "Video for Traffic Data and Incident Detection by Traficon," Proceedings of the 5th ASCE International Conference, 1998, pp. 33-40.