

## INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

**UMI<sup>®</sup>**

Bell & Howell Information and Learning  
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA  
800-521-0600



**New Technologies for Data Collection and their  
Application for Empirical Investigation of  
Travel Time Measurement Issues**

Scott Stuart Washburn

A dissertation submitted in partial fulfillment of the  
requirements for the degree of

Doctor of Philosophy

University of Washington

1999

Program Authorized to Offer Degree: Department of Civil and  
Environmental Engineering

UMI Number: 9936502

---

UMI Microform 9936502  
Copyright 1999, by UMI Company. All rights reserved.

This microform edition is protected against unauthorized  
copying under Title 17, United States Code.

---

**UMI**  
300 North Zeeb Road  
Ann Arbor, MI 48103

In presenting this dissertation in partial fulfillment of the requirements for the Doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to University Microfilms, 300 North Zeeb Road, Ann Arbor, MI 48106-1346, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature Scott A. Wasl

Date 6/8/99

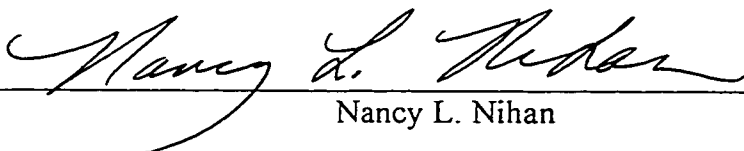
University of Washington  
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Scott Stuart Washburn

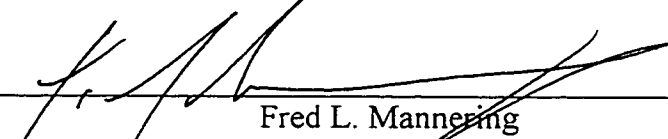
and have found that it is complete and satisfactory in all respects,  
and that any and all revisions required by the final  
examining committee have been made.

Chair of Supervisory Committee:

  
\_\_\_\_\_  
Nancy L. Nihan

Reading Committee:

\_\_\_\_\_  
Nancy L. Nihan

  
\_\_\_\_\_  
Fred L. Mannering

  
\_\_\_\_\_  
G. Scott Rutherford

Date: 6/7/99

University of Washington

Abstract

**New Technologies for Data Collection and their  
Application for Empirical Investigation of  
Travel Time Measurement Issues**

Scott Stuart Washburn

Chairperson of the Supervisory Committee:  
Professor Nancy L. Nihan, Ph.D.  
Department of Civil and Environmental Engineering

Travel time is emerging as the preferred measure of congestion and mobility. From both the planning and operational perspectives, travel time is considered to be a key and desirable variable for several reasons. Travel time is a great indicator of the quality of flow and operation on a transportation facility, its definition is straightforward and therefore also easy to understand by the commuting public, it is easy to apply to the evaluation of existing and future conditions analyses, it is easy to apply to the comparison of different travel modes, and it lends itself to efficient statistical evaluation.

This research investigated two new technologies for the collection of travel time data—video image tracking and voice recognition. Video image tracking appears to currently be one of the more promising technologies for offering comprehensive data collection, automatically, and in real-time. Unlike many of the other available technologies, video image tracking systems have the potential to provide all standard traffic measurements (e.g., speed, volume, lane occupancy, headway), as well as travel time. Additionally, video image tracking systems are not as susceptible to the sampling deficiencies and privacy concerns of some of the other tested technologies.

Given that portable methods of travel time data collection will continue to be commonly used, the implementation of a new technology, voice recognition, was also

investigated as a new tool for collecting data in the field. Voice recognition offers the potential of providing hands-free automated entry of data. While most previous applications of this technology have focused on dictation type uses, field data collection has very different demands for a voice recognition “engine” and offers many more challenges than the typical benign dictation environment (e.g., office).

As part of the process of investigating these technologies for their applicability to collecting travel time data was the examination of significant sampling and bias issues. These issues, as well as important data collection experiences by the author, were synthesized into a collection of practical guidelines that will supplement the existing literature and assist the transportation professional planning a travel time study.

## TABLE OF CONTENTS

---

|  |            |
|--|------------|
| <i>List of Figures</i> .....   | <i>iii</i> |
| <i>List of Tables</i> .....  | <i>v</i>   |
| <b>Chapter 1 Introduction</b> .....  | <b>1</b>   |
| 1.1 Problem Statement.....   | 4          |
| 1.2 Research Objectives .....  | 9          |
| 1.3 Organization of Document .....   | 9          |
| <b>Chapter 2 State of the Art of Travel Time Data Collection</b> .....                                     | <b>11</b>  |
| 2.1 Introduction .....   | 11         |
| 2.2 Probe Vehicle Techniques .....   | 12         |
| 2.3 License Plate Matching (LPM).....  | 15         |
| 2.4 Inductance Loop Detectors .....  | 18         |
| 2.5 Video Image Processing .....   | 20         |
| <b>Chapter 3 Speech Recognition Technology as a Tool for Data Collection</b> .....                         | <b>23</b>  |
| 3.1 Introduction .....   | 23         |
| 3.2 Voice Recognition Technology Overview .....  | 23         |
| 3.3 Key Issues.....  | 25         |
| 3.4 Potential Applications .....   | 26         |
| <b>Chapter 4 16-Bit Speech Recognition Enabled Program</b> .....   | <b>29</b>  |
| 4.1 Program Development.....   | 29         |
| 4.2 Research Design .....  | 36         |
| 4.3 Results and Interpretation.....  | 38         |
| 4.4 Conclusions .....  | 45         |
| <b>Chapter 5 32-Bit Speech Recognition Enabled Program</b> .....   | <b>47</b>  |
| 5.1 Introduction .....   | 47         |
| 5.2 IBM 32-bit VoiceType Speech Recognition Engine.....  | 47         |
| 5.3 Program Development.....   | 47         |
| 5.4 Research Design .....  | 53         |
| 5.5 Results and Interpretation.....  | 56         |
| 5.6 Conclusions and Recommendations.....   | 64         |
| <b>Chapter 6 Video Image Tracking as a Tool for Data Collection</b> .....                                  | <b>66</b>  |
| 6.1 Introduction .....   | 66         |
| 6.2 Mobilizer System.....  | 67         |
| 6.3 Research Design .....  | 67         |
| 6.4 Results and Interpretation.....  | 71         |
| 6.5 Conclusions and Recommendations.....   | 83         |
| <b>Chapter 7 Sampling Considerations for Travel Time Measurement</b> .....                                 | <b>88</b>  |
| 7.1 Introduction .....   | 88         |
| 7.2 Properties of Traffic Flow that Affect Point Sampling and Statistical Analysis of<br>Travel Times..... | 89         |

|  |            |
|--|------------|
| 7.3 Data Description .....   | 95         |
| 7.4 Travel Time Results.....   | 100        |
| 7.5 Sampling Capabilities.....   | 120        |
| 7.6 Conclusions and Recommendations.....   | 135        |
| <b>Chapter 8 Travel Time Measurement Bias Issues.....</b>  | <b>139</b> |
| 8.1 Introduction .....   | 139        |
| 8.2 Selection Bias for Point Sampling.....   | 141        |
| 8.3 Measurement Error Bias (Reference Point Time-Stamping).....  | 154        |
| 8.4 Aggregation Bias .....   | 167        |
| 8.5 Error Due to Outliers .....  | 169        |
| 8.6 Conclusions and Recommendations.....   | 173        |
| <b>Chapter 9 Guidelines For Travel Time Data Collection.....</b>   | <b>176</b> |
| 9.1 Introduction .....   | 176        |
| 9.2 Planning the Travel Time Study.....  | 177        |
| 9.3 Developing the Data Collection Plan .....  | 182        |
| 9.4 Conducting the Study .....   | 190        |
| 9.5 Data Analysis and Reporting.....   | 193        |
| <b>Bibliography.....</b>   | <b>196</b> |
| <b>Appendix A Voice Recognition Program Functionality .....</b>  | <b>202</b> |
| <b>Appendix B Equipment and Data Collection Considerations for Travel Time<br/>Measurement with the Mobilizer Video Image Tracking System.....</b> | <b>203</b> |
| <b>Appendix C License Plate Matching and Simulation Program Description.....</b>   | <b>215</b> |
| <b>Appendix D License Plate Considerations.....</b>  | <b>218</b> |

## LIST OF FIGURES

---

|   |     |
|---|-----|
| Figure 1. Example Grammar File .....  | 31  |
| Figure 2. Background Noise Threshold Adjustment Utility .....                     | 32  |
| Figure 3. Software Program User Interface (16-bit Version) .....                  | 33  |
| Figure 4. Microphone Volume Level Adjustment.....                                 | 50  |
| Figure 5. Software Program User Interface (32-bit Version) .....                  | 51  |
| Figure 6. Program Operation in Test Mode .....                                    | 52  |
| Figure 7. Time Stamp Entry Scheme Dialog Box .....                                | 53  |
| Figure 8. Number of License Plate Characters Input Dialog Box .....               | 53  |
| Figure 9. Montlake Blvd. (Upstream Station) .....                                 | 68  |
| Figure 10. Montlake Blvd. (Downstream Station).....                               | 68  |
| Figure 11. I-5/NE 50 <sup>th</sup> (Upstream Station).....                        | 68  |
| Figure 12. I-5/NE 45 <sup>th</sup> (Downstream Station).....                      | 68  |
| Figure 13. Montlake Travel Time Results (first 5-minutes) .....                   | 76  |
| Figure 14. Montlake Travel Time Results (second 5-minutes) .....                  | 77  |
| Figure 15. Montlake Travel Time Results (third 5-minutes) .....                   | 78  |
| Figure 16. Aerial Photo of 45th Data Collection Site.....                         | 97  |
| Figure 17. Schematic of I-5 Data Collection Site (North to the right).....        | 99  |
| Figure 18. Vehicle Matches versus Data Collection Time (Session 1) .....          | 107 |
| Figure 19. Arterial Travel Time Distribution (Session 1).....                     | 108 |
| Figure 20. Vehicle Matches versus Data Collection Time (Session 2) .....          | 109 |
| Figure 21. Arterial Travel Time Distribution (Session 2).....                     | 110 |
| Figure 22. Overall Travel Time Results (GP Lane).....                             | 113 |
| Figure 23. Overall Travel Time Results (HOV Lane) .....                           | 114 |
| Figure 24. Travel Time Distribution (free-flow conditions).....                   | 119 |
| Figure 25. Travel Time Distribution (stop-and-go conditions) .....                | 119 |
| Figure 26. Average Speed Distribution (free-flow conditions) .....                | 119 |
| Figure 27. Average Speed Distribution (stop-and-go conditions) .....              | 119 |
| Figure 28. Modeled Manual Sampling Distributions for Freeway Traffic.....         | 125 |
| Figure 29. Sampling Probability Functions .....                                   | 128 |
| Figure 30. Travel Time Sampling Comparisons.....                                  | 132 |
| Figure 31. Travel Time Tracking by Floating Cars at Various Headways (GP Lane) .. | 134 |
| Figure 32. Travel Time Results by Vehicle Classification (GP Lane) .....          | 149 |
| Figure 33. Travel Time Results by Vehicle Classification (HOV Lane).....          | 150 |
| Figure 34. Timestamp Accuracy Comparison (Session 1) .....                        | 163 |
| Figure 35. Timestamp Accuracy Comparison (Session 2) .....                        | 163 |
| Figure 36. Timestamp Error Distributions (Timestamp Accuracy Technique).....      | 164 |
| Figure 37. Timestamp Error Distributions (Maximum Sample Size Technique).....     | 165 |
| Figure 38. Timestamp Error Distributions (Session 1).....                         | 166 |
| Figure 39. Timestamp Error Distributions (Session 2).....                         | 166 |

|   |     |
|---|-----|
| Figure 40. GP Lane Travel Time with various aggregation levels .....      | 168 |
| Figure 41. Travel Time Plot for Identifying Outliers .....                | 171 |
| Figure 42. Multi-layered architecture for linktime estimation system..... | 204 |
| Figure 43. Surveillance camera image looking south.....                   | 208 |
| Figure 44. Surveillance camera image looking south.....                   | 208 |
| Figure 45. Video image snapshot using 1/60 shutter speed.....             | 209 |
| Figure 46. Video image snapshot using 1/1000 shutter speed.....           | 209 |
| Figure 47. Example calibration line setup .....                           | 211 |
| Figure 48. FOV with very limited area for placing calibration marks.....  | 212 |
| Figure 49. Matching program main screen.....                              | 215 |
| Figure 50. License plate character entry/match scheme .....               | 216 |
| Figure 51. Sampling options dialog box.....                               | 217 |

## LIST OF TABLES

---

|  |     |
|--|-----|
| Table 1. Summary of Arterial Data Collection Statistics .....  | 42  |
| Table 2. Summary of Ambient Sound Levels at Data Collection Sites.....   | 43  |
| Table 3. Data Collection Summary.....  | 55  |
| Table 4. Summary of Arterial Data Collection Statistics .....  | 57  |
| Table 5. Summary of Freeway Data Collection Statistics .....   | 58  |
| Table 6. Results Summary by Method and Traffic Condition.....  | 60  |
| Table 7. Results Summary by Data Collector.....  | 61  |
| Table 8. Arterial Summary Statistics.....  | 72  |
| Table 9. Arterial Travel Time Results .....  | 72  |
| Table 10. Arterial Travel Time Accuracy Results .....  | 72  |
| Table 11. Travel Time Statistical Comparisons.....   | 74  |
| Table 12. Travel Time Interval Statistics.....   | 75  |
| Table 13. Freeway Summary Statistics.....  | 79  |
| Table 14. Freeway Travel Time Results .....  | 79  |
| Table 15. Freeway Travel Time Accuracy Results.....  | 80  |
| Table 16. Travel Time Statistical Comparisons.....   | 81  |
| Table 17. Freeway Summary Statistics.....  | 82  |
| Table 18. Freeway Travel Time Accuracy Results.....  | 82  |
| Table 19. NE 45 <sup>th</sup> Travel Time Statistics, 1 <sup>st</sup> and 2 <sup>nd</sup> Session .....                    | 101 |
| Table 20. NE 45 <sup>th</sup> Travel Time Statistics, 3 <sup>rd</sup> Session.....   | 102 |
| Table 21. ANOVA Results for Session 1 .....  | 102 |
| Table 22. ANOVA Results for Session 2 .....  | 103 |
| Table 23. Travel Time Comparisons for NE 45 <sup>th</sup> Street.....  | 111 |
| Table 24. Freeway Vehicle Matching Statistics.....   | 112 |
| Table 25. Sample Size Comparisons for Freeway Data.....  | 115 |
| Table 26. Distribution Skewness Comparisons .....  | 118 |
| Table 27. Sampling Rate Ranges for Arterials .....   | 121 |
| Table 28. Sampling Rate Ranges for Freeways .....  | 122 |
| Table 29. Manual Sampling Distributions for Freeway Traffic.....   | 126 |
| Table 30. Simulated Travel Time Results, GP Lane (I-5, 130th - 145th).....   | 129 |
| Table 31. Vehicle Type Percentages for First Arterial Data Collection Session.....   | 143 |
| Table 32. Vehicle Type Percentages for Second Arterial Data Collection Session.....  | 143 |
| Table 33. Vehicle Classification Frequencies for Origin/Destination Averaged and<br>for Matched Vehicles (Session 1) ..... | 144 |
| Table 34. Vehicle Classifications Key .....  | 147 |
| Table 35. Vehicle Classifications for Freeway Data.....  | 148 |
| Table 36. Comparison of Speeds under Non-congested Conditions .....  | 151 |
| Table 37. Time Stamp Accuracy Results.....   | 158 |
| Table 38. Timestamp Corrected Arterial Travel Times .....  | 161 |

|  |     |
|--|-----|
| Table 39. Comparison of Maximum Sample Size versus Time Stamp Accuracy .....   | 163 |
| Table 40. Freeway Simulation Spurious Matches .....  | 173 |
| Table 41. Common Travel Time Measurement Techniques.....   | 182 |
| Table 42. Unreadable Plate Percentages .....   | 219 |
| Table 43. Vehicle Classification Frequencies for Origin (NE 45 <sup>th</sup> St / 15 <sup>th</sup> Ave NE)....         | 226 |
| Table 44. Vehicle Classification Frequencies for Destination<br>(NE 45 <sup>th</sup> St / 7 <sup>th</sup> Ave NE)..... | 227 |

## **ACKNOWLEDGMENTS**

---

I would like to most sincerely thank my faculty advisor, Nancy Nihan, for her guidance and support during this work. I would also like to thank my other committee members, Fred Mannering, Scott Rutherford, and Les Jacobson, for their advice and assistance.

I would like to thank the members of the TransNow staff, particularly Lois Landgraf and Andrea Franklin, for their administrative assistance.

I would like to thank my parents, Joe and Jonell, for instilling in me the drive, motivation, and work ethic to pursue my most ambitious goals.

And finally, and mostly, I would like to thank my wife, Leslie, for her steadfast loving support throughout this entire journey.

## CHAPTER 1 INTRODUCTION

---

Transportation mobility impacts every facet of our society, from individuals' choices on residence location to business economic productivity. Transportation mobility is generally acknowledged by individuals as one of the most prized benefits of our society. However, the benefits that transportation mobility have created, have also inadvertently led to reductions in transportation mobility in major metropolitan areas in the form of traffic congestion. The negative effects of traffic congestion, such as lost productivity, accidents, energy inefficiency, and air pollution, have been well documented over the years. As a result of these recognized problems, the federal government enacted legislation (ISTEA and TEA-21) to address these issues. Additionally, the federal Clean Air Act Amendment (CAAA) has put emphasis on reducing mobile source emissions. Since a disproportionate amount of air pollution results from congested conditions, air quality is heavily influenced by congestion.

Given the increasing focus on traffic congestion and attempts to manage it, a recent research effort [Lomax, Turner, and Shunk, 1997] explored ways of trying to measure congestion. One of the recommendations from this research was that "travel time-based measures be used to estimate congestion levels". Also, one of the conclusions stated in the report was that "Direct collection of travel time data is a priority item for complete implementation of a suggested congestion measurement system." Other researchers have come to the same conclusion. Mulhall [1995] stated "After review and discussion of possible data types, travel time was identified as the most appropriate measure of congestion." Gallagher and Pagitsas [1995] stated "Travel time, measured in a variety of forms, has been identified as a variable which has the ability to express many of these mobility goals of the Massachusetts Congestion Management System." Smith [1995] stated "The development of congestion management systems is causing transportation agencies to more seriously consider travel time (and its derivatives such as speed and travel rate) as a basis for identifying and quantifying congestion and mobility.

One of the objectives of a congestion management system should be to develop performance measures that more directly address how the public views congestion and mobility. Travelers tend to think in terms of travel time, trip time, and delay.”

From both the planning and operational perspectives, travel time is considered to be a key and desirable variable for several reasons. Travel time is a great indicator of the quality of flow and operation on a transportation facility, its definition is straightforward and therefore also easy to understand by the commuting public, it is easy to apply to the evaluation of existing and future conditions analyses, it is easy to apply to the comparison of different travel modes, and it lends itself to efficient statistical evaluation.

There are four primary uses for travel time information. One is for general planning purposes, a second for evaluation of transportation system improvements, a third for real-time commuter information, and a fourth for incorporation into real-time operational control strategies (including incident detection). Some more specific areas where travel time information is very useful are now listed.

- Measuring the effectiveness of a treatment to a roadway facility; for example, signal timing changes on an arterial, capacity additions on a freeway, HOV lane additions, and transit signal priority treatments.
- Calibration of simulation models. Simulation has become an extremely popular tool, and the calibration of existing conditions is a critical step to the modeling of any situation. Travel time is recognized as excellent parameter to base transportation corridor simulation calibrations upon.
- Evaluation of transportation policies. For example, the Washington State Department of Transportation (WSDOT) has included travel time savings and reliable trip time as required objectives for the high occupancy vehicle (HOV) system in the Puget Sound region [WSDOT, 1996].
- Real-time traffic information. The commuting public readily understands travel time.

In response to the ISTEA legislation, IVHS (ITS) America proposed a strategic plan to meet its goals and objectives [IVHS America, 1992]. Two of the main technology areas that have been defined by IVHS (ITS) America are Advanced Traffic Management Systems (ATMS) and Advanced Traveler Information Systems (ATIS). ATMS is also identified as being the basic building block of IVHS/ITS, with all other functional areas utilizing information provided by ATMS. IVHS (ITS) America has defined several key roles for ATMS, including the following—“...predict traffic congestion and provide alternative routing instructions to vehicles over wide areas, in order to maximize the efficiency of the highway network and maintain priorities for high-occupancy vehicles (HOV's)...will collect, utilize, and disseminate real-time data on congestion on arterial streets and expressways...” One of the key roles it defined for ATIS was—“...provide a variety of information that assists travelers in reaching a desired destination”. The strategic plan also states, “In order to implement ATMS, real-time traffic monitoring and data management capabilities must be developed, including advanced detection technologies such as image processing systems, automated vehicle location and identification techniques, and the use of vehicles themselves as traffic probes.” These sentiments were expressed in another report [Hughes and JHK 1994]—“As arterial and freeway systems become more complex, methods of collecting more comprehensive and accurate traffic flow data are needed to support future ITS applications intended to better manage and maintain these systems.”

In today's information age, people expect to find the kind of information they need to help them make many of the decisions they face everyday. One major area in which information is greatly lacking is the kind of information that would help commuters decide which route they want to take to their destination. ATIS is aimed at meeting this challenge. However, it is ATMS that will be responsible for obtaining and providing the information that gets utilized by ATIS.

## 1.1 Problem Statement

### 1.1.1 Existing Condition

Despite the current emphasis on travel time measurement, it has yet to gain widespread acceptance or use by transportation engineers [NCHRP Report 398, 1997]. In addition, some transportation engineers have not fully embraced the idea of basing congestion management measures on travel time. The following passage was contained within the NCHRP report. “Some members of the NCHRP Project 7-13 Review Panel and other outside reviewers of this work have expressed concern about using travel time measures. Many of these concerns relate to data collection procedures and costs. The Study Team agrees that the state of the practice is not to the point where the widespread direct collection of travel time and speed data from the traffic (or travel) stream is a realistic possibility. However, when the needs of the variety of customers identified for congestion statistics are considered, there is no better method of analyzing improvement alternatives and presenting that congestion information than using travel time and speed measures.”

It is apparent that the use of travel time measurements will not become standard procedure until several significant issues in regard to this variable are resolved. One issue is the cost, in terms of both labor and equipment. For manual methods, such as floating car or license plate matching studies, the labor costs can be substantial. For automatic methods, such as AVI or video imaging, the equipment costs can be prohibitive for many applications. Given the widespread use of inductance loop detectors (ILD's), and the maturity of the technology, it is not surprising that many transportation engineers would prefer to rely on data already available from this source. Thus, travel time may be viewed by some engineers as a last resort measurement if other, more readily available and obtainable, measures cannot be used to satisfy the project requirements.

From the directives offered in the IVHS (ITS) strategic plan, it is obvious that automatic travel time data collection is necessary for the needs of traffic management centers. Currently, where travel time information is gathered and/or reported, it is more

commonly through indirect estimation using older technologies like ILD's. Although many researchers have developed algorithms to utilize ILD data (e.g., volume, lane occupancy, and speed (paired detectors)) for travel time estimation, the reliability and statistical rigor of these estimates has generally been deemed inadequate for real-time operations input or other applications where very accurate and unbiased measurements are needed. Furthermore, the costs associated with ILD maintenance is a significant concern. Unreliable and/or failed ILD's can present significant problems to these algorithms. In the Puget Sound region, indirectly estimated travel times are currently only used for reporting to the commuting public (WSDOT) via a traffic conditions site on the internet. However, one will quickly learn that they should allow for a large margin of error when using this information for trip planning purposes. But until automatic travel time technology is pervasive throughout major metropolitan centers, there will continue to be efforts into research and application of surrogate estimation methods. But these efforts would undoubtedly benefit from proven direct travel time measurement technologies for model/algorithm validation.

Despite the research being done in the area of advanced technologies for automatically collecting travel time data, there will continue to be a need for manual methods of travel time data collection. For example, before-and-after study data are typically collected by field personnel for a very specific, limited number of occasions. And planning data are often obtained by carrying out special studies for limited locations and time periods. But manual methods, while not usually expensive in terms of equipment, can be considerably expensive in terms of labor. The major portion of labor cost comes from the number of field personnel required. With the advent of the affordable portable computer, post processing has become a much less significant factor than it used to be for most manual methods. Therefore, there is still the need for improved manual methods of travel time data collection that can reduce the number of field personnel required.

Another issue for travel time is that the method(s) for measuring it is not as straightforward as the methods for many other commonly collected traffic measurements.

For example, when one wants to measure volume, one just simply counts the vehicles passing by. When one wants to measure spot speed, one uses a radar gun and “shoots” vehicles. But with travel time, it is not always intuitive to the transportation engineer practitioner what method should be used to collect travel time data. The appropriate method (e.g., floating car, license plate matching) depends on several factors, such as study purpose, required data accuracy, and others. Thus, a travel time study often requires additional research on the part of the practitioner to determine the most appropriate method.

Interrelated to this issue, once a transportation engineer practitioner has chosen a method for collecting travel time data, they may not be sure of exactly how to apply that technology/method to ensure that accurate and statistically valid results are obtained. This issue is certainly not trivial, although one might not get that impression from many current practitioners. As finely detailed and exhaustive as many studies are that examine transportation facility improvements, the application of a chosen travel time data collection method is often performed in an ad-hoc manner. In these cases, the limitation is not so much with the technology, but with the procedures followed to ensure statistically valid results. Some researchers that have investigated different methods/technologies for travel time data collection have acknowledged the importance of a technically sound data collection plan, regardless of the technology/method. A Volpe Center study [Liu and Gaines, 1996] stated “...a sampling plan would be an integral part of any survey plan to ensure confidence in the results...”. However, documentation on guidelines to follow to ensure a technically sound data collection are limited, and are usually very generalized. As the Volpe study continued to state “...the experimental nature of these surveys did not allow for up-front assessment of the appropriate statistical samples to be used to ensure valid representation”. While some various examples can be found scattered though out the literature, the Institute of Transportation Engineers’ ‘Manual of Traffic Engineering Studies’ [ITE, 1994] provides little guidance to the practitioner beyond a sample size formula. Thus, it is easy to see why many ad-hoc procedures have been used for collecting travel time data in the past.

Likewise, it is also not difficult to believe that various sources of bias often get incorporated into the collected data as a result of these ad-hoc procedures. With a sound data collection plan it should be possible to control for potential sources of bias.

A sampling plan is certainly a key aspect of any data collection plan. But oftentimes, investigators send data collectors into the field without a good idea of how many travel time samples need to be obtained. Usually, the researcher will end up with either an inadequate number of samples, or considerably more than needed. Although, the latter is more desirable from a statistical validity standpoint, it is usually not desirable from a resource allocation standpoint. Although, various studies have recommended guidelines for adequate sample sizes for statistically valid travel times, little discussion has been given to the actual sampling strategies. It is possible that developing a sampling strategy is not considered a priority because practitioners just assume that potential sources of bias can be accounted for in the error term associated with the sample size formula. Technically, though, this is not the intent of the error term. Sources of bias should be eliminated such that the error term can be made as small as possible, and will only account for random perturbations that cannot be controlled. Thus, a higher level of statistical confidence will be obtained from the results. Manual methods typically offer the most flexibility in sampling strategies, given that they are controlled by data collection personnel in the field. Automated methods should be evaluated for potential sources of bias as well. For example, methods that rely on sampling a designated class or fleet (e.g., transit vehicles) of vehicles will potentially suffer from a selection bias problem.

### *1.1.2 Condition Being Sought*

As recommended by IVHS (ITS) America, there are many technologies deserving research attention. While some work has already been done in studying the application of technologies like AVI and AVL to travel time data collection, little work has been done in the area of applying video image analysis to travel time data collection. Video image tracking appears to currently be one of the more promising technologies for offering

comprehensive data collection, automatically, and in real-time. Unlike many of the other technologies, video image tracking systems have the potential to provide all standard traffic measurements (e.g., speed, volume, lane occupancy, headway), as well as travel time. Additionally, video image tracking systems are not as susceptible to the sampling deficiencies and privacy concerns of some of the other tested technologies. And relatively easy integration into a traffic management center with video camera surveillance infrastructure already in place is another major benefit of this technology. Thus, one of the main tasks of this research was to investigate a video image tracking system for its potential applicability to usage in ATMS.

Given that portable methods of travel time data collection will continue to be commonly used, the implementation of a new technology, voice recognition, was also investigated as a new tool for collecting data in the field. Voice recognition offers the potential of providing hands-free automated entry of data. While most previous applications of this technology have focused on dictation type uses, field data collection has different demands for a voice recognition “engine” and offers many more challenges than the typical benign dictation environment (e.g., office).

Being as travel time is a more sophisticated measurement than typical point measurements where a near total population sample size is possible, methods for the proper application of any technology to travel time data collection are necessary. Without accurate and unbiased measures of travel time, evaluation of transportation system improvements becomes convoluted. Additionally, travel time reporting to the commuting public will be of little use to commuters if it is frequently inaccurate. Accurate methods of measuring travel time are also needed for use as independent checks of developed models/algorithms that are based on other technologies (e.g., loop detector output) for indirect estimation of travel time (where cost may prohibit new technology). And, the results of any simulation modeling are always going to be only as good as the calibration effort and data. This research examined data obtained from application of both of the above discussed technologies for travel time measurement.. These data were

examined in detail for potential sampling deficiencies and susceptibility to all potential bias sources.

## **1.2 Research Objectives**

There were three major objectives of this research. The first objective was to determine the feasibility of two new technologies, namely voice recognition and video image tracking, as travel time data collection tools. Integral to this was an evaluation of the ability of each of these technologies to provide accurate, statistically unbiased travel time measurements. The first technology, voice recognition, is targeted at consultants or agencies that need to perform quick, portable travel time studies for varying facilities on a special need basis. These studies are done on an infrequent basis, usually for updating general transportation statistics for planning projects, or for evaluating before-and-after measures due to an improvement project. The second technology, video image tracking, is targeted for public agency traffic management centers. This technology looks to fill the need for providing real-time travel information for the commuting public and for support of ITS applications. This technology is most suitable for permanent installations on key travel routes. The second objective was to use these technologies to investigate different travel time sampling issues and potential sources of bias that can become incorporated into travel time statistics without a sound data collection plan. The third objective consisted of the development of a set of practical data collection guidelines to supplement the existing literature that can be consulted by the transportation engineering practitioner looking to develop a data collection plan for travel time measurement.

## **1.3 Organization of Document**

Chapter 2 is a review of the current state of the art in travel time data collection methods and technologies. The remainder of the dissertation is comprised of four components. The first component (chapters 3, 4, and 5) consists of the description of a new technology (speech recognition) for use in manual data collection, its incorporation into a software application, and the testing of the application. The second component

(chapter 6) discusses a new technology (video image tracking) for use in the automatic measurement of travel time. The third component (chapters 7, 8, and 9) discusses sampling and bias issues that need to be considered in the development of a travel time data collection plan. The fourth and final component (chapter 9) consists of practical guidelines for the collection of travel time data that can be utilized by transportation engineer practitioners to help them develop and execute a travel time data collection plan.

Rather than a separate chapter for conclusions and recommendations, each main chapter contains conclusions and recommendations at the end of the chapter. Additionally, chapter 9, in essence, summarizes much of the material contained in the preceding chapters.

## CHAPTER 2

# STATE OF THE ART OF TRAVEL TIME DATA COLLECTION

---

### 2.1 Introduction

There are three primary areas for use of transportation data, whether for freeways or arterials. These areas can be classified as detection/operation, incident management, and planning. [Hughes and JHK 1994] Obtaining accurate and timely data for each of these areas is an important element in transportation professionals' strategies and plans for maximizing the efficiency of existing transportation facilities. Primary traffic measures for monitoring, operation, and management of transportation networks are flow/volume, occupancy/density, speed, and travel time.

Travel time measurement is the most important traffic parameter for congestion monitoring systems. The value of travel time measurements is that transportation facilities with differing operations (e.g., arterial vs. freeway, HOV lane vs. GP lane) can be effectively compared by utilizing this universal measure of effectiveness.

Many technologies currently exist for measuring traffic flow parameters. Of course, the most common is inductance loop detectors. This technology, like many others, is specifically designed for point measurements, such as speed and volume. However, very few technologies exist for automatically determining the travel time between two points on a roadway facility. The initial generation of video image systems were aimed at computing several traffic flow measurements, but these were still exclusively point measurements, and thus incapable of providing direct measurement of travel time.

Most current methods for evaluation of travel time are not automated. The more common methods include the floating car technique, license plate matching, and cellular telephone reporting [Hamm, 1993]. Current technologies for measuring travel time manually, as well as automated methods are described below.

## 2.2 Probe Vehicle Techniques

Probe vehicle techniques generally rely on time marks being recorded at specific reference points along a roadway network while driving a vehicle with the traffic stream along that travel route. This information can then be converted to travel time, speed, and delay for each segment within the route. Originally, this technique consisted of either the driver or a passenger manually recording times from a stopwatch onto paper while traveling the route. Since then, the development of in-vehicle electronic distance measuring equipment has made this process more automated and precise. Even more recently, the probe vehicle technique has been extended to more passive methods (e.g., AVI, AVL). These methods make use of vehicles already in the traffic stream as part of their normal travels. These techniques are discussed in more detail below.

### 2.2.1 Average Vehicle or Floating Car Method

These two techniques are essentially the same, with both consisting of driving a vehicle within the traffic stream on the roadway facility of interest. The implementation just differs slightly between the two methods. For the average vehicle technique, the driver of the vehicle attempts to travel at a speed they judge to be the average speed of all vehicles within the traffic stream. In the floating car technique, the driver drives such that they pass as many vehicles as pass them. The driver, or often a passenger, records the times at which the vehicle passes predetermined travel route reference points, and any other desired information, such as stopped delay. Travel times, average speeds, and stopped delays can be easily be determined from these methods. These techniques are relatively straightforward to implement given that the main piece of “equipment”, an automobile, is usually readily available. The main source of variation is usually the data recording method. As previously mentioned, pencil and paper is one method, and an electronic distance measuring device connected to the car is another method. Additionally, the use of voice recorder is common. With this method, it is usually not necessary for an additional passenger to record the information, as the driver can usually handle driving and speaking into a microphone at the same time.

The primary disadvantage of this technique is that it often suffers from sample size problems. Because of flow conditions and routing logistics, it is difficult for a single vehicle to make more than a few travel time runs (on any segment of significant length) during a peak period. Thus, it is usually necessary to use many vehicles to achieve any respectable level of statistical certainty with regards to the average travel time. Additionally, individual driver's can affect significant differences in measurement results due to individual perceptions of the prevailing traffic speed. This method is generally regarded as adequate only for planning purposes and not for before-and-after evaluation of a roadway facility. Hundreds of data collection efforts have been performed using this technique. The sample size issue is the most consistently reported issue of difficulty for this technique. For example, a study by Sisiopiku, Roupail, and Santiago [1994] found that there was a great potential for measurement bias due to the driver's perception of the average speed. Additionally, in this study, the sample sizes were so small using the average-car method that they supplemented the travel time data collection effort by performing license plate matching.

### *2.2.2 Automatic Vehicle Location (AVL)*

AVL systems rely on the use of an in-vehicle device to notify a traffic systems management center of the times at which it reaches certain positions. A vehicle operator manually reporting location information via a cellular phone is one example, but most current research is focusing on the potential of global positioning systems (GPS) to automatically report location information. GPS technology is based on signals sent from multiple satellites orbiting the earth, upon which an in-vehicle electronic device can be used to determine its precise position on the earth's surface. With the size and cost of portable GPS units decreasing substantially in the recent past, this technology is becoming more feasible. Since the sample size associated with this method is restricted to the size of the fleet for which vehicles are equipped with the GPS units, cost is a very significant issue. To obtain adequate coverage and statistically valid sample sizes, a very large number of vehicles need to be equipped with GPS units for most major

metropolitan areas. This technology, however, is really not mature enough yet to be used practically in this environment. A couple of major issues, such as hardware and software standards, and the current practice of the Department of Defense of intermittently degrading the satellite signals, still need to be resolved. The major advantage of this technology is that location and time information for each GPS unit equipped vehicle can be very precise. The major disadvantage is the limited sample size, which in fact could have practical limits due to the general motoring public being reluctant to use these devices because of privacy concerns. Reliance on commercial vehicles, or other specific fleet vehicles (e.g., taxis), may not only be insufficient for sample size, but may also lead to selection bias in the travel time estimates.

### *2.2.3 Automatic Vehicle Identification (AVI)*

Automatic Vehicle Identification (AVI) systems are similar in nature to AVL systems except that, rather than being able to determine a vehicle's position at any point in time, vehicle positions can only be reported at specific locations equipped with sensors. An AVI system consists of some sort of in-vehicle device, or transponder, that emits signals that are received by roadside devices. Thus, only when a vehicle passes a point with a receiver can that vehicle's position be determined. These types of systems are already fairly popular in electronic toll collection applications, where vehicles equipped with special transponders get uniquely identified by equipment stationed at the toll facility and their toll accounts are automatically debited. Again, AVI systems are capable of providing very accurate measurements of travel time as the measurement data comes directly from uniquely identified vehicles. The main disadvantages are basically the same as with the AVL systems: limited sample size (except in areas with electronic toll collection facilities) and personal privacy concerns. Although certain types of agency and/or fleet vehicles (e.g., buses, taxis) could easily be equipped with these devices (for non-toll facilities), it is still likely that the sample size obtainable from any given traffic stream would be small, and lead to selection bias problems as well unless the distribution of AVI tags is widespread. Experience with the Houston AVI system (which has several

automated toll booths), where approximately 40,000 vehicles were equipped with AVI tags, found that adequate sample sizes for statistical validity were generally obtained during the peak periods on the major freeways [Turner and Holdener, 1995].

### **2.3 License Plate Matching (LPM)**

A common method for determining travel time is to record license plates, and corresponding time-stamps, of passing vehicles at specific locations along a designated route. After the data collection, license plates are matched between all the observation points to determine travel times. The license plate matching technique is most commonly implemented by just a few methods. Most of these methods rely on placing people in the field who observe the license plates and then record them along with their corresponding time-stamp. A more recent method uses video to record passing license plates and corresponding software to extract the license plate characters. The license plate matching technique generally overcomes the limited sample size issue associated with the probe vehicle technique. However, data is limited to the point observations along a facility. The more common techniques are described below.

#### *2.3.1 Voice Recorder*

Voice recorders offer a convenient way to collect data, as speaking is the only continuous task that is required. Additionally, voice recorders are relatively easy to operate. For license plate data collection, they offer the advantage of allowing the user to keep his/her eyes on the traffic/license plates; thus, potentially increasing accuracy and sample size. However, for doing travel time studies, associating a time-stamp with each plate entry is necessary, and this can prove to be difficult with this method. Additionally, it is necessary to replay the tape(s) after the data collection effort and transcribe the data into an electronic format. This can be very time consuming and expensive.

Researchers [Miller, et al., 1993] explored this technique for license plate data collection for an origin-destination study. They found that it was possible to record up to 1,000 license plates per hour, but in practice this rate was difficult to sustain. They

concluded that this technique was not suitable for high volume locations. A very significant problem they encountered was the number of usable plates transcribed from the tape, sometimes as low as 80 plates per hour for difficult tapes. Additionally, they found that the plate matching rate was significantly lower for either the pencil/paper or the manual video method.

### *2.3.2 Portable Computer Keyboard Entry*

With recent advances in computing power and reductions in price, portable computers (e.g., laptops) have become a popular tool for field studies. Portable computers are fairly well suited to the collection of license plate data for travel time studies. The user can type license plate characters into a special program which will automatically time-stamp the entries, based upon the internal computer clock. The use of special software programs also eliminate the need for any further data transcribing after the data collection effort. A quality piece of software will allow the data collectors to immediately post-process (e.g., plate matching) the license plate data without any further modification(s).

There are still some difficulties with this method, however. The accuracy of the license plate entries is often directly correlated with the typing ability of the data collector. A data collector not familiar with the keyboard not only will have difficulty with accurate entries, but likely will gather a smaller sample because they spend a lot of time looking at the keyboard instead of the traffic stream. Another difficulty with this method can arise when license plates are being collected from a freeway overpass, where binoculars are necessary to read the plates. This situation usually requires two people at the data collection station (one to read plates aloud, the other to enter them) because it is difficult for one person to read plates with binoculars and still type the entries, while getting an adequate sample size. Mounting high-powered binoculars on a tripod might be of assistance to a single data collector, but for high speed traffic, it is usually necessary to follow the vehicle for a short time with the binoculars to get a long enough look at the plate to read it. Also, for the situation where one person reads the license plates aloud

while the other person records them, another potential source of error is introduced—the potential misinterpretation of spoken letters and/or numbers by the recording person.

A study [Rickman, Hallenbeck, and Schroeder, 1990] compared license plate matching using portable computers (keyboard entry) to floating car runs for travel time measurement. Using similar personnel numbers for both techniques, they found that the license plate matching technique resulted in anywhere from 2 to 15 times the number of travel time measurements than the floating car method. The upper value associated with facilities with a high proportion of through traffic. Despite the greater sample sizes of the license plate matching technique, they still found the travel time measurements between both techniques to be statistically comparable. This research did reveal another potential problem with floating car studies. They found the variance of the travel times to be greater for the keyboard entry sample than for the floating cars runs. They believed this was a result of drivers driving in a very consistent manner from run to run, drivers driving an average speed and avoiding very high or very low travel times, and the loop pattern of the runs resulted in drivers entering the arterial during the same part of the progression cycle from run to run.

### *2.3.3 Video Optical Character Recognition*

Automatic license plate matching consists of processing video-taped images of license plates using machine vision. More specifically, cameras are placed that focus on a small area of either the front or rear end of a vehicle so as to capture the license plate within the field-of-view. The video tapes are then processed using computer algorithms designed to translate the license plate image into a unique set of characters. The system can then use these unique sets of characters to try to match these license plates in successive fields-of-view. In a nutshell, the machine vision system simply automates the previously laborious task of using personnel to manually review the video tapes and enter license plates into the appropriate software program. The advantages of a video license plate reading system are that it is capable of obtaining a high sample rate of license plates and dramatically reduces the post-processing time experienced with manual methods.

The disadvantages of this type of system are that it is limited in the other kinds of information it can collect because of the very narrow camera field-of-view and a separate camera is needed for each lane of traffic. Because of the labor involved (setting up cameras, etc.) and the cost associated with dedicated cameras for every lane, this method is better suited for long-term evaluation of a designated facility.

Many different factors affect the success of this type of system, such as vehicle speeds, vehicle headways, vehicle volume, license plate mounting location, plate character scheme, plate occlusion, and weather and lighting conditions. Correct license plate recognition rates can vary dramatically depending on these factors. These rates reportedly can vary between 15 and 90 percent. A study in Tampa, Florida reported license plate capture rates of 63 to 85 percent of the total number of license plates viewed by the cameras from freeway sites. A study in Seattle, Washington reported a plate capture rate of about 75 percent. This capture rate was augmented by some manual processing however. For real-time processing, it is unlikely that as high a correct match rate will be obtained as off-line processing. Studies have also shown favorable statistical comparisons of travel time results using this method compared to more traditional methods.

The most popular system currently used in the United States is marketed by Transformation Systems/Computer Recognition Systems (CRS) [Shuldiner et al., 1996].

## **2.4 Inductance Loop Detectors**

While loop detectors are strictly a point processing device, some researchers have investigated ways to use loops to give passing vehicles a unique signature that can be re-identified at subsequent loop detector locations. Essentially, three approaches have been attempted in this regard.

The most common approach involves using the standard loop measures of volume and occupancy to compute speeds (or in the case of dual loop stations, measure speed directly) and then computing travel time based on known link distances and the assumption of constant speeds between detector stations. For single loop stations, which

are far more common than dual loop stations, the main drawback with this method is that a conversion from the detector measured occupancy to density is required. This generally involves a conversion factor based upon an assumption of vehicle length. Many researchers have explored this technique [Nam & Drew (1996), Sen et al (1997), Sisiopiku et al (1994), Petty et al (1997), Dailey et al, (1991)] with varying degrees of success, but none with the kind of statistical reliability required by ATMS for all flow conditions.

In another approach, Coifman [1998] investigated using vehicle lengths measured from dual loop stations (a.k.a. speed traps) to uniquely identify platoons of vehicles, generally three vehicles in length. Essentially, this algorithm identifies a group of vehicles at a downstream station that is likely the same group that passed an upstream station. This algorithm is dependent on the vehicle length resolution of the speed trap, which is a function of detector spacing, vehicle speed, and detector sampling frequency, and the platooning characteristics of the traffic. Preliminary testing was done using freeway data for congested conditions. These conditions are more conducive to the success of the algorithm for two reasons: 1) vehicle speeds are low, thus resulting in a higher degree of vehicle length resolution from the loop detectors, and 2) under congested conditions, vehicles have less opportunity to overtake other vehicles and platoon spreading is less likely, thus there will be more intact platoons observed at both stations. The challenge in applying this algorithm to non-congested conditions becomes apparent—higher vehicle speeds result in larger length resolutions, making it more difficult to distinguish between vehicles of similar sizes, and there is less likelihood of identical platoons passing both an upstream and downstream station. Additionally, this algorithm is likely to be ineffective for areas with a large amount of weaving or vehicle mixing at entrances and exits. Real-time field tests had yet to be conducted.

The other approach involves using the electronic “signatures” measured by loop detector as a vehicle passes over it. Due to the varying concentration of metal along a vehicle’s length axis, the magnetic flux of the detector will vary accordingly. A plot of the variation of the inductance versus time will yield a “unique” signature for a vehicle.

This method is heavily dependent on the sensitivity of the loop detector. Results of applications of this method to travel time measurement have yet to be published.

## **2.5 Video Image Processing**

Video image processing consists of converting video images gathered by closed-circuit cameras into digital representations. These digital images are then processed by computer hardware and software to extract traffic flow measurements, such as speed, volume, and headway. The basic principle behind these systems is that they can detect changes in the video image, such as a vehicle traversing a point, through rapid monitoring of the luminance levels of pixels on the video image. Since the pavement is generally a certain color (or gray shade) and luminance level, the passing of a vehicle over a certain spot will change the color and luminance levels of that spot, thus identifying a moving object. Sophisticated algorithms filter out shadows and other items that might be confused for a moving vehicle.

One of the key advantages a video image system has over methods like AVI and AVL is that it is a passive system; that is, the measurements can be taken without requiring any vehicles to be equipped with any special equipment. Thus, the issue of obtaining willing participants to allow their vehicles to be monitored through the use of special devices is avoided.

There currently exist three different forms of video image processing technology. These are described below.

### *2.5.1 Point Processing*

Previous generation video image processing systems were based on the "tripwire" approach, which determines vehicle passage through a video image when a vehicle passes a pre-selected band of pixels within the image. Vehicle detection is based on changes in image intensity at user-defined detection regions. In effect, this technology provides a video emulation of loop detectors. The user can establish numerous discrete detection regions within the field-of-view (FOV), but the system is not capable of identifying

unique vehicle characteristics for later re-identification at a downstream FOV. This video image processing technology is the oldest and most established of the three approaches.

Available systems using this approach include AUTOSCOPE, CCATS, TAS, IMPACTS, and TraffiCam [Chatziioanou, et al, 1994, Hockaday, 1991, Hoose, 1992, Klein & Kelley, 1996, and MNDOT, 1997].

### *2.5.2 Within Field-of-View Tracking*

The newer generation of video image processing systems are generally referred to as vehicle tracking systems. These systems are characterized by utilization of the entire video frame, focusing on the movement regions within each frame [Hockaday, 1991]. Vehicle trackers utilize one of two processing schemes: identification of differences in successive video frames or analysis of differences between the entire video frame and a background frame (no vehicles present) [Hockaday, 1991].

Unlike previous generation video image systems that were capable of only processing specific points within the field-of-view (FOV), vehicle tracking systems can process individual vehicles as they traverse a FOV. The previous generation video image systems were only capable of providing rough classifications of vehicle length as the only unique vehicle identifying parameter. Thus, these earlier generation video image systems were incapable of determining whether a particular vehicle passed through more than one FOV.

However, while these systems follow (or “track”) vehicles within an individual FOV, they do not obtain specific vehicle identifying information. Thus, they are still incapable of providing direct travel time measurements. The best these types of systems can do is to obtain accurate measurements of speed at frequent intervals and estimate travel times from that information.

### *2.5.3 Vehicle Re-identification Tracking*

A more advanced type of tracking system, on the other hand, actually obtains identification information from each vehicle as it passes through a FOV. This

information can then be used to re-identify vehicles at downstream FOV's. Thus, this type of system extracts details from vehicle images within each FOV and uses this information to match vehicles at successive FOV's. If enough correct matches are obtained, an average link (from upstream FOV to downstream FOV) travel time can be computed. Thus, the advantage of this type of tracking system is the potential to measure travel time, in addition to all the other standard traffic flow parameters. As will be discussed in much more detail later in this report, the Mobilizer system [developed by Condition Monitoring Systems (CMS)] uses this tracking methodology and was the focus of part of this research.

## **CHAPTER 3**

# **SPEECH RECOGNITION TECHNOLOGY AS A TOOL FOR DATA COLLECTION**

---

### **3.1 Introduction**

As discussed in the previous chapter, many techniques have been used, and are currently used, for collecting transportation data. For manual forms of data collection, methods range from using pencil and paper to, more recently, laptop computers running specialized software programs. However, with recent advances in computing technology, speech recognition technology has emerged as a potentially viable tool for transportation data collection, a tool with very distinct advantages over existing methods of data collection.

Most persons, when asked about voice recognition, probably think about dictation for word processing applications. This type of application for voice recognition is really only the tip of the iceberg, however.

Speech recognition is performed by a person every time they listen to another person talking. That person must decode the sounds coming from the other person's mouth and convert them into the known words of the alphabet. That is essentially what software for voice recognition does as well. As words and/or phrases are spoken by a person into a microphone, the sound wave is analyzed by the software to recognize sequences of sounds that comprise words and/or phrases in the predefined vocabulary.

### **3.2 Voice Recognition Technology Overview**

In-depth discussion on the technical details of speech recognition technology is beyond the scope of this document, but detailed information can be found in the IBM VoiceType [IBM, 1995] Technical Reference. Instead, the more significant components from a user's point of view will be discussed.

The state-of-the-art in speech recognition has advanced a significant amount in recent years. An earlier research paper [Dew and Bonsall, 1990] described the results of

an experiment that used an electronic device that transcribed spoken voice from a voice recorder. This study used a state-of-the-art (at the time) recording device to record data collectors' voices in the field. This recorder was then attached to the electronic transcription device. However, they concluded at the time that this method was not feasible because the accuracy rate (60-70%) was not acceptable. Additionally, this particular technique is less desirable than a method that can perform the transcription/recognition simultaneously with the data collection. This technique also suffered from difficulty of calibrating the time-stamp with each entry.

Recent improvements in computer hardware technology have now made computer speech recognition feasible for many applications. Numerous speech recognition software packages have appeared on the market recently. While most of these packages are aimed at a specific application, such as dictation into word processing programs, a few companies have developed voice recognition "engines" designed to be utilized by software developers in building custom voice recognition enabled applications. To utilize one of these engines, it is necessary to write a software application using an object-oriented programming language (e.g., Visual C++, Visual Basic) that links the programming code of the voice recognition "engine" with the code of the main software application. The key technological considerations for speech recognition are now discussed.

### *3.2.1 Dictation vs. Command-and-Control*

A dictation application is one which simply attempts to recognize each individual spoken word and enter it into a text editor (e.g., Microsoft Word). The voice recognition engine interacts only with the text editor. Due to the nature of this type of application, it must be capable of recognizing all words spoken in everyday language. Thus, the application must utilize an extremely large (tens of thousands of words) vocabulary.

Command-and-control applications are usually specific programs designed to interact with the user through voice commands. These applications are designed to recognize a very specific set of words and/or phrases. Each of these words/phrases

usually execute a particular command within the application. Command-and-control applications utilize a much smaller vocabulary set than dictation applications, on the order of tens to hundreds of words/phrases.

### *3.2.2 Continuous vs. Discrete Speech*

Because of the size of the vocabulary involved (usually tens of thousands of words) for general dictation purposes, it is necessary to use a discrete speech mode. This mode requires users to pause between individual words. Command-and-control type applications can utilize continuous speech because of the significantly smaller vocabularies. Thus, voice recognition “engines” function in either the continuous or non-continuous (discrete) speech recognition mode. Continuous speech “engines” can process long segments of spoken words without requiring the speaker to pause for unnaturally long periods of time between words. Conversely, non-continuous speech “engines” require significant pause lengths between spoken words.

### *3.2.3 Speaker Independent vs. Speaker Dependent*

Speaker-dependent engines must be trained for a specific user’s voice pattern—for each word in the vocabulary. This can be particularly cumbersome, especially for large vocabularies, and is not very practical for dictation type applications. Speaker-independent systems, on the other hand, do not require any voice training. Some systems offer speaker independence while also allowing speaker training if desired. Some systems offer a hybrid form of speaker training, where the user trains the system on a small subset of some very common words and phrases, and the system uses this abbreviated training to determine the speaker’s voice pattern.

## **3.3 Key Issues**

In using voice recognition for data collection, there are a few factors which play a significant role in the success of this technology for this purpose. These factors are now discussed.

### *3.3.1 Background Noise*

While dictation applications are aimed at the office environment, using voice recognition for data collection will often take place in a field setting. This presents a significant consideration for the performance of a voice recognition system under field noise conditions. Noise levels in a traffic setting can be as much as 40 dB higher than an office setting. This corresponds to volume conditions that can be 16 times louder than an office setting. Thus, a primary challenge for the voice recognition engine is separating out the user's voice from the background sounds.

### *3.3.2 Accuracy*

Obviously, the accuracy of recognized speech is an important concern. Frequent mistakes with a dictation application might be overcome with spelling and grammar checkers, but frequent mistakes with field data collection could result in a wasted, and costly, effort.

### *3.3.3 Processing Speed*

Processing human speech is a complex task. As such, a considerable amount of computing power is required to perform this task. While some data collection tasks may be regulated by the pace of the data collector, others are generally governed by the flow rate of the data collection item. In the case of license plate entry, vehicular flow rates obviously drive the data entry rate.

## **3.4 Potential Applications**

Voice recognition has the potential to be used for many applications in the transportation field. An introduction to the focus of the next two chapters is now described.

### *3.4.1 License Plate Data Collection for Travel Time Studies*

As previously discussed, two common methods for manually collecting travel time data are the voice recorder and the portable computer for recording license plates.

Voice recognition has the potential to combine the advantages of both of these methods while eliminating the disadvantages inherent to both.

To demonstrate the potential for voice recognition to be used as a tool for data collection, the job of developing a specific data collection application was undertaken. An application for collecting license plate data for travel time studies was developed. It was felt that this type of application would prove to be one of the most demanding applications of voice recognition for transportation data collection. Thus, if this application were successful, it would demonstrate that this technology is feasible for transportation data collection at this time. The reasons why this type of application is so demanding are discussed below.

- **Need for fast performance:** Given that data collection would occur under field conditions with sampling of vehicles from traffic streams, it would be necessary for the application to operate with sufficient speed to allow the potentially very frequent entering of license plates.
- **Loud background noise conditions:** Obviously, being positioned next to streams of vehicular traffic would impose some extreme background noise conditions. The noise levels would definitely far exceed those experienced under office conditions where voice recognition in the form of dictation is most common.
- **Complex recognition search space (high perplexity):** The use of 26 letters and 10 numbers, any of which could be equally likely on any given plate character entry, forms a very complicated, and computationally intensive, set of possible recognition phrases, especially when each plate entry can consist of up to six characters. Additionally, if other commands need to be considered, such as keywords for triggering the time-stamping, it becomes even more complex.

### 3.4.2 *Other Possible Applications*

Some other potential applications are listed below.

- Parking surveys—based on license plate entry

- Origin-destination studies—based on license plate entry
- Saturation flow studies—saying key words into the microphone when certain vehicles cross the stopbar
- Roadside driver interviews
- Speed recording—saying speeds displayed on a radar gun aloud, which are then entered automatically into a software program
- Probe vehicle studies—driver says specific word/phrase as he/she passes reference locations. Software program then time stamps entry for specific location. Again, combines advantages of both computer and voice recorder.

## CHAPTER 4

# 16-BIT SPEECH RECOGNITION ENABLED PROGRAM

---

### 4.1 Program Development

Considering the issues discussed previously under the speech recognition technology review section, it was necessary to choose a speech recognition engine that would meet the needs of the application. There are several vendors that offer speech recognition technology; however, most offer applications that are either dictation specific or aimed at very specialized tasks. What was needed for transportation data collection was a speech recognition engine that was aimed at the software application developer; one which incorporated the major necessary features (e.g., speaker-independence), yet allowed the software developer to have complete control of the speech recognition engine through program coding.

After a review of the available technology, it was determined that a speech recognition engine developed by IBM was most appropriate. The name of this engine at the time was Voice Type Application Factory (VTAF). The IBM<sup>®</sup> voice recognition “engine” is a continuous-speech, speaker-independent system. In other words, this speech engine can extract and process individual words spoken in a long sentence and the system does not need to be trained for individual voice types. However, a current constraint is that the user must speak English.

Before data collection and testing could occur, it was necessary to write the license plate data collection application. The initial version of this program was developed to run on a 16-bit Windows™ platform, as this was the available technology at the time. Once a functional version of the program using just keyboard input was developed, the next stage was to incorporate the voice recognition capability.

Program development utilized the Windows for Workgroups (or Windows 3.1) 16-bit operating system and the Visual Basic 3.0 16-bit development environment. Since the speech recognition engine interface code was based on the C++ language, a third-

party development tool was utilized that facilitated the conversion of the Visual Basic language calls to the speech engine.

#### *4.1.1 Vocabulary and Grammar Files*

A critical element to obtaining good speech recognition performance is the proper defining and building of grammar and vocabulary files. The vocabulary file defines what words or characters are acceptable input into the program. The grammar file defines the acceptable vocabulary input syntax for the application. For example, the grammar file can specify whether the user must speak numbers before letters or any other combination that is desired.

The purpose of the vocabulary and grammar files is to minimize voice input processing time and to maximize the accuracy of the recognized voice input. This is accomplished by: 1) limiting the vocabulary search space for acceptable words and 2) instructing the voice processing “engine” what types of words and their respective sequence are to be expected in a voice input. Figure 1 shows an example grammar file. This file is in Backus-Naur Form (BNF) format. The vocabulary file essentially consists of the words under the <action> and <alphanum> headings.

##### 4.1.1.1 Individual letters versus words

Currently, all voice recognition “engines” have difficulty distinguishing between individual letters of the alphabet, just as humans oftentimes do. In particular, distinguishing between the similar long ‘E’ sounds of several letters of the alphabet (e.g., B, C, D, E, G, P, T, V, Z) can be problematic. In light of this difficulty with individual letters, the first task was to test the simpler case of using words that represent the letters of the alphabet. The military alphabet was chosen for this purpose (e.g., Alpha for A, Bravo for B, Charlie for C, etc.).

```

<requests> ::= <plate> .
<plate>    ::= <alphanum> <alphanum> <alphanum> <alphanum> <alphanum>
           <alphanum> <action>
           | <alphanum> <alphanum> <alphanum> <alphanum> <alphanum>
           <action>
           | <alphanum> <alphanum> <alphanum> <alphanum> <action> .
<action>   ::= ENTER | CANCEL .
<alphanum> ::= ALFA    | BRAVO    | CHARLIE | DELTA |
              ECHO    | FOXTROT  | GOLF    | HOTEL |
              INDIA   | JULIETT | KILO    | LIMA  |
              MIKE    | NOVEMBER| OSCAR   | PAPA  |
              QUEBEC  | ROMEO   | SIERRA  | TANGO |
              UNIFORM| VICTOR  | WHISKEY| X_RAY |
              YANKEE | ZULU   |
              ZERO   | ONE    | TWO    | THREE | FOUR |
              FIVE   | SIX   | SEVEN  | EIGHT | NINE .

```

**Figure 1. Example Grammar File**

#### 4.1.2 Background Noise

One of the main factors that affects voice recognition performance is the background noise level. The technique incorporated into the voice recognition engine to account for this factor is to determine start-talking and end-talking threshold values. These values govern the sound level at which the speech processor will begin to listen for voice input and the level at which the speech processor will stop listening for voice input. For example, in a quiet office setting, the threshold values will be set relatively low because there is very little background noise that might be interpreted as voice input. This will allow the user to speak in a normal voice tone. On the other hand, in a noisy outdoor environment, the threshold values will be set relatively high. Although the user will be required to speak at a louder level (at least louder than the background noise level), this will prevent the speech processor from trying to interpret spurious background noises as speech. To allow the user to set the threshold levels to something appropriate for the particular noise environment they are working in, the threshold value setting

utility was been incorporated into this program. Figure 2 shows this utility's interface. This utility facilitates the setting of the threshold values by having the user speak a specified sentence into the microphone while in the desired location of data entry. The threshold values are relative and do not necessarily correspond to a particular sound pressure scale.

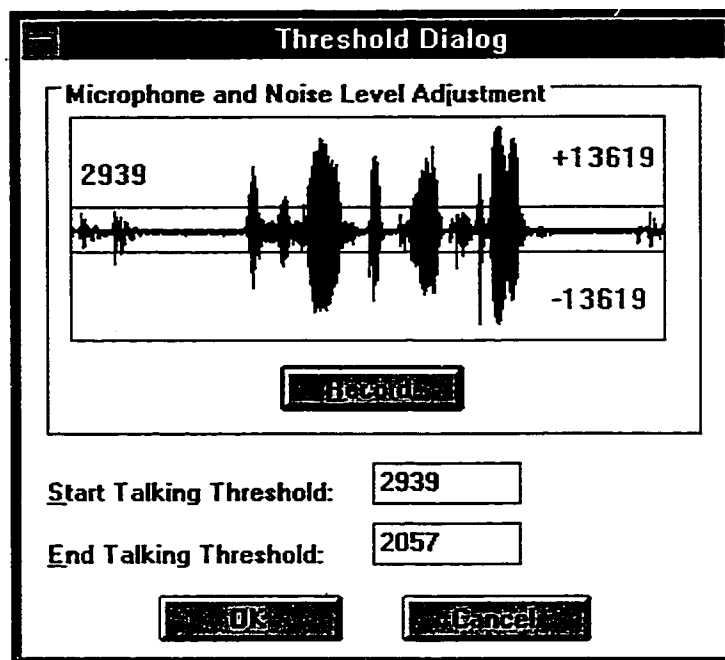


Figure 2. Background Noise Threshold Adjustment Utility

In the example shown in the figure, the voice recognition engine will not begin to process voice input until the sound level exceeds the 2939 level, and will stop processing voice input when the sound level drops below the 2057 level. The end-talking threshold value is less than the start-talking threshold value because peoples' voices tend to trail off in volume when finishing speaking.

#### 4.1.3 User Interface

This program was designed to be both functional in the field and easy to use. Figure 3 shows the main user interface. While most of the features of this user interface are self-explanatory, a few of the features deserve some discussion.

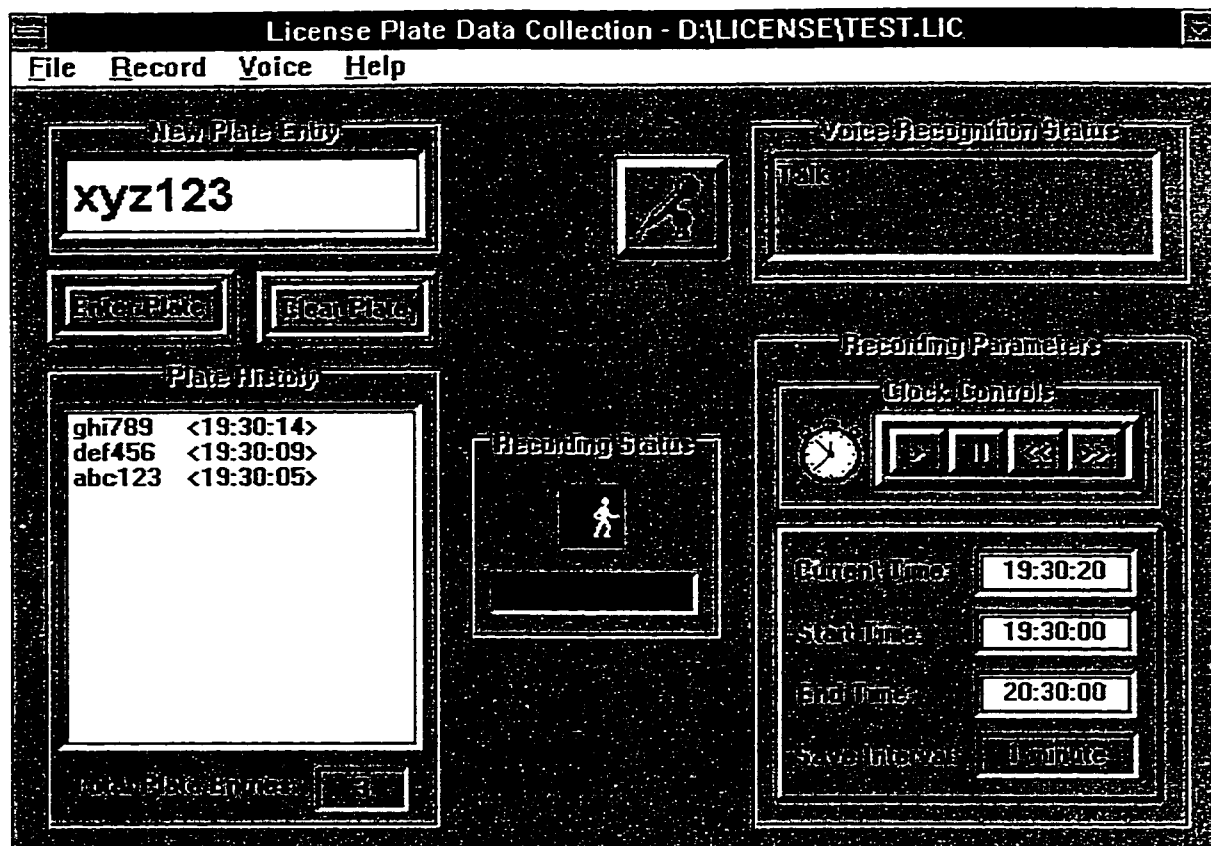


Figure 3. Software Program User Interface (16-bit Version)

#### 4.1.3.1 Voice Recognition Status

The Voice Recognition Status box simply displays the status of the voice recognition processor. Next to this display box is a button with a microphone image on it. Pressing this button once activates the voice recognition processor. If it is ready and waiting for speech input, it will display the message Talk. If the user has just spoken, it will display the spoken text in the status box.

#### 4.1.3.2 New Plate Entry

The current method for a user to indicate that they have finished with a voice input and are ready to make another one is to say the word 'enter' immediately following the plate characters that they would like to be entered and time-stamped. The word 'enter' instructs the voice "engine" that it should process the previously spoken words as

a unique entry. Once the voice “engine” has processed that entry, the program will display the recognized text in the Voice Recognition Status box briefly. The program will then convert the words into the appropriate characters (e.g., ‘alpha bravo charlie one two three’ will be converted to ‘abc123’) and display the result in the New Plate Entry box briefly before entering this text into the Plate History list box and time-stamping it. License plates can still be entered directly into the New Plate Entry box through the use of the keyboard. Buttons for entering and clearing the plate entry have been provided on the user interface for this purpose, although the keyboard Enter and Escape keys provide the same capability.

#### 4.1.3.3 Clock Controls

While originally intended for use in the field, this program was also designed to be used in the office/lab. The clock controls are provided for this purpose. For example, if the user post-processes a videotape of license plate data, the clock can be easily manipulated to correspond to the videotape time-stamp after pausing, fast-forwarding, rewinding, or other tape manipulations. There is a dialog box accessible through a menu that allows the user to specify real-time or non-real-time data collection mode and set the clock, start and end times as appropriate.

The program is also capable of providing audio feedback for all critical features, such as start and end recording mode changes, as well as when the current voice input processing is complete and the program is ready to accept another input.

#### *4.1.4 Equipment Considerations*

The intent behind this development effort was to develop an application that would not require any special equipment beyond a microphone. Therefore, an application that would run on a standard personal computer (PC) platform without the addition of non-standard hardware was the goal. The minimum requirements for using this application are discussed in the following sections.

#### 4.1.4.1 Computing Power

While keyboard based license plate surveys can be performed on rather modest portable computers, the complexity of the voice processing necessitates the use of a reasonably “high-powered” portable computer. Initial field testing was performed using a laptop computer configured with an 80486 processor, running at a 50 MHz processing speed, and 8 MB of RAM. More recent field tests have been performed with a laptop computer configured with a Pentium® processor running at a 120 MHz processing speed, and 32 MB of RAM.

#### 4.1.4.2 Sound Card

An obvious piece of required equipment is a sound card, or audio processor. A standard 16-bit, SoundBlaster™ compatible audio card is satisfactory. Most cards are capable of processing stereo inputs, but this application only requires monaural processing. It should be noted that there are certain models of sound card that do not amplify microphone input audio signals, so it may be necessary to install a signal amplifier inline between the microphone and sound card, or adjust the signal level through software control.

#### 4.1.4.3 Microphone

Microphone selection is a very important consideration, and probably the most critical link in the speech processing event. Preliminary testing has been performed with a relatively inexpensive, hand-held microphone, as well as with a very high quality headset style microphone. While the program can still function adequately under many conditions with the inexpensive type of microphone, for best results, it is recommended that a very high quality microphone be used for the more demanding situations (e.g., freeway overpass locations). At the very least, a microphone that is unidirectional or super cardioid in nature, incorporates noise canceling, and has an impedance of 600 ohms or less is recommended [IBM VTAF Technical Reference]. The high-quality microphone used during testing was a Sennheiser HMD 410-6 model. This is a headset style microphone with excellent voice reception characteristics. This headset style microphone

is probably ideal, as it ensures a fixed distance between the microphone and the user's mouth, and it also allows the most freedom for the user's hands, which could be important, such as in cases when the user needs binoculars to read the license plates. Additionally, the built-in headphones allow for direct audio feedback from the program to the user, which is especially convenient when the user is not directly viewing the screen or cannot hear the built-in computer speakers.

## **4.2 Research Design**

### *4.2.1 Introduction*

The main objectives in developing this application were to determine if speech recognition as a tool for transportation data collection was feasible at this time and to compare key measures of effectiveness of this method to the most currently comparable method of portable computer keyboard entry. In essence, this initial effort was a pilot study to determine if further research was warranted at this time.

### *4.2.2 Key Measures of Effectiveness*

When collecting license plate data for travel time studies, there are two key measures of effectiveness for the data collection process. One is the accuracy rate of the inputted license plates, and the other is the sample rate obtainable during certain traffic flow conditions over a specific time period.

There are two components to input accuracy. First is the accuracy of the data collector in correctly observing the license plate characters, and second is the accuracy of inputting those characters correctly. Accurate or inaccurate reading of the license plate characters is something both methods have in common, and during a normal data collection effort, it is usually not possible to discern what portion of the erroneous entries were a result of inaccurate plate reading or an error related to the input method. With the voice recognition method, it is plausible that the plate reading component of error would be less than that for the keyboard entry because it is never necessary for the data collector to take their eyes off of the traffic stream (i.e., license plates). Of course, for the

keyboard entry method, the data collector can take extra time reading the plate to be positive about its characters, but this will undoubtedly come at the expense of the sample rate relative to the voice entry method. For input accuracy, this is a function of typing accuracy for the keyboard entry method. For the voice recognition method, this is related to saying the characters in a manner consistent with positive recognition.

Sample rate is an important measure, because the more entries one can collect during a given time, the higher level of statistical confidence one will have in the obtained travel time results. The sample rate capability is composed of several components: 1) time to read and assimilate the license plate characters; 2) time to enter the license plate characters; and 3) time for computer to process the entry and become ready to accept another entry.

#### *4.2.3 Data Collection*

Preliminary testing in the field was conducted at a couple of arterial and freeway locations. For these initial tests, most of the developer adjustable parameters were left at their default values (the threshold settings, which are user adjustable, were of course modified for each specific situation). The first arterial and freeway test used an inexpensive microphone. The subsequent tests used the high quality Sennheiser microphone. The arterial tests were performed such that comprehensive descriptive statistics could be obtained, while the freeway tests were simply designed to be mainly qualitative in nature at this point. The main purpose of the freeway tests was to determine whether this technology was feasible at this time under the very high background noise conditions exhibited at these locations.

For these tests, data were collected for only a single location on the study site. The focus of these tests were to just evaluate sample and accuracy rates, not travel time measurements, which obviously require two or more data collection locations.

As mentioned previously, formal testing was performed using the military alphabet to represent the letters of the alphabet (e.g., Alpha for A, Bravo for B, Charlie for C, etc.). For these tests, four characters from each license plate were entered. The

choice of four characters was somewhat arbitrary at this point, but consistent with what has been cited in other literature [Rickman et al., 1990, Schaefer, 1988]. Chapter 7 provides more details on this issue.

Reference data were obtained by video recording all passing vehicles/license plates. These video tapes were then manually reviewed and the license plate data were entered into a software database. Keyboard and voice entries were compared to the video data for results calculations.

### **4.3 Results and Interpretation**

#### *4.3.1 Accuracy and Sample Size*

##### 4.3.1.1 Laboratory Conditions

Testing in the laboratory was conducted with an IBM<sup>®</sup>-compatible PC with an 80486, 66 MHz processor, and 16 MB of RAM.

Under relatively noise-free conditions, recognition accuracy rates for both the military alphabet letters and the numbers, when spoken individually, range from 95 to 99 percent. There is some variance in the recognition accuracy between the individual military alphabet words. Some words are recognized accurately virtually 100 percent of the time, while some are occasionally mis-recognized. Overall, however, the combined accuracy rate for the military alphabet words still averages at least 95 percent or better. The recognition accuracy rate for the numbers actually approaches 100 percent. The very rare mis-recognized number is usually a result of very poor enunciation on the part of the speaker.

Accuracy rates for different combinations of letters and numbers varied. As mentioned previously, accuracy rates for the numbers are inherently higher; thus, for combinations of characters input that include a larger proportion of numbers, the overall accuracy rate will be higher. For instance, for a combination of three numbers and one letter input, the accuracy rate varies between 95 and 99 percent, depending on the

distribution of the various letters input. In situations where a greater proportion of letters are used, one can expect the recognition accuracy rate to decrease slightly.

It should be noted that while the military alphabet was used for preliminary testing, any word can be defined for use with the program. Thus, for any particular words of the military alphabet that result in an unacceptable recognition accuracy rate for a user, he/she can replace them with more suitable words for their needs.

#### 4.3.1.2 Field Conditions

The first two tests, one arterial and one freeway, were performed with the following equipment:

- IBM<sup>®</sup>-compatible laptop computer—50 MHz, 80486 processor, 8 MB of RAM, and an inexpensive hand-held microphone.

**Arterial Test 1:** The first test site was a busy four-lane (two lanes each direction) arterial near the University of Washington. The data collection was performed from a pedestrian bridge that crosses the arterial. Data were collected from just the outside lane of the two lanes in the southbound direction for approximately one-half hour. The plate input method consisted of entering three numbers and one letter, except for small trucks, for which 4 numbers only were entered.

During this data collection, a total sample of 155 vehicle license plates were obtained from a total volume of 273 vehicles, for a sample frequency of 56.8 percent. Of the 155 plates collected during this study, 14 were entered incorrectly. Of this 14 that were entered incorrectly, three were by human error (i.e., misread plate) and the other 11 were by errors in the voice processing. This calculates to a 92.8 percent accuracy rate by the speech processing “engine”, and an overall accuracy rate 90.8 percent for the entire method. The data set was verified by careful review of a high-speed video recording of the license plates in this lane during this time period. Interestingly, 5 out of the 11 voice recognition specific errors were due to the “letter” alpha being mis-interpreted. This was likely a result of the effect of significant background noise on the lower quality

microphone, as this particular “letter” did not show any propensity for mis-interpretation in either the lab or in subsequent field tests using the high-quality microphone.

**Freeway Test 1:** Testing at the freeway on-ramp site, which exhibits even higher noise levels than the arterial location, also showed very good recognition accuracy. Again, this freeway test was primarily qualitative in nature; thus, precise statistics were not calculated for this test. This on-ramp consists of a metered lane and an HOV bypass lane. The data were collected off-and-on for approximately 45 minutes during the afternoon peak period. The license plate data were collected from the metered lane, and since the traffic was moving slowly, full license plate numbers were entered. While this data set was not verified with video, visual verification of the computer processed voice input against the spoken entry was performed. This process indicated an accuracy rate of approximately 90 percent. Thus, this test showed that even with a relatively low quality microphone, data collection at a noisy freeway site is feasible. It should be noted that while the noise levels at this site were certainly higher than at the arterial location, they were not as high as can be experienced at many freeway locations. This was mainly because the data collection was done at ground level and to the side of the main freeway lanes.

The following tests were performed with the following equipment:

- IBM®-compatible laptop computer—120 MHz, Pentium processor, 32 MB of RAM, and the high-quality headset microphone

Additionally, for these tests, a sound level meter was used to check the ambient noise levels at the specific test site. These sound level readings were taken for a short period of time, but encompassed the typical range of noise level that was prevalent for that site during the data collection period. Table 1 summarizes the sound levels present at the arterial locations and the freeway overpass location.

**Arterial Test 2:** The second arterial test was performed at a similar location as the first arterial test. This data collection was performed for approximately one-half hour during the afternoon peak period. This data collection effort took place during less than ideal environmental conditions. The light level was low, with many vehicles having headlights on, and the temperature was cold (around 45-50 degrees). For this test, data was collected via the voice input method, as well as keyboard input method, simultaneously (using the three letter and one number input scheme again). This provided two data sets that could be directly compared for accuracy and sample size values. Again, the reference data set was obtained through the use of high speed video. Despite the difficult working conditions, the results were still very good.

For the voice input method, a total sample of 199 vehicles was obtained from a total volume of 331 vehicles, for a sample frequency of 60.1 percent. From these 199 sampled vehicles, an accuracy rate of 96.3 percent was obtained. For the keyboard entry method, a total sample of 184 vehicles was obtained, again from a total volume of 331 vehicles, for a sample frequency of 55.6 percent. An accuracy rate of 94.0 percent was obtained from this method.

While the accuracy rates of both methods were quite good, the sample sizes were limited due to the nature of the traffic flow. At this particular location, the flow was heavily influenced by a traffic signal just a few hundred feet upstream. This resulted in heavily platooned traffic flow for 40 seconds out of every minute, and very sparse flow during the other 20 seconds. Thus, during the sparse traffic flow, it was easy to enter every passing vehicle plate, but during the heavily platooned flow, it was only possible to read about half of the vehicle plates passing by, regardless of the data entry method being used.

**Arterial Test 3:** The third arterial test was performed on the same arterial, and in a similar location, as the previous two arterial tests. For this test, however, data was collected from the opposite direction of traffic flow as the previous test. This flow condition was more uniform, but still exhibited some platooning effects due to a traffic

signal about a half-mile upstream. This test was performed under good lighting conditions, although the temperature was still around 50 degrees. This test was performed for just 12 minutes. Again, data was collected simultaneously by both the voice and keyboard input methods (using the three letter and one number input scheme). High speed video was used to verify the results again.

For the voice input method, a total sample of 103 vehicles was obtained from a total volume of 141 vehicles, for a sample frequency of 73.0 percent. From these 103 sampled vehicles, an accuracy rate of 97.0 percent was obtained. For the keyboard entry method, a total sample of 103 vehicles was also obtained, again from a total volume of 141 vehicles, for a sample frequency of 73.0 percent. An accuracy rate of 89.8 percent was obtained from this method.

Table 1 summarizes the important data collection statistics from the arterial tests.

**Table 1. Summary of Arterial Data Collection Statistics**

| Arterial            | Time Period<br>(minutes) | Flow Rate<br>(vphpl) | Voice Recognition |                  | Keyboard Entry |                  |
|---------------------|--------------------------|----------------------|-------------------|------------------|----------------|------------------|
|                     |                          |                      | Sample<br>Rate    | Accuracy<br>Rate | Sample<br>Rate | Accuracy<br>Rate |
| Test 1 <sup>a</sup> | 27                       | 618                  | 56.8%             | 90.8%            | — <sup>c</sup> | — <sup>c</sup>   |
| Test 2 <sup>b</sup> | 26                       | 771                  | 60.1%             | 96.3%            | 55.6%          | 94.0%            |
| Test 3 <sup>b</sup> | 12                       | 720                  | 73.0%             | 97.0%            | 73.0%          | 89.8%            |

<sup>a</sup> Test 1 used the lower quality microphone and the slower processing speed computer.

<sup>b</sup> Tests 2 and 3 used the higher quality microphone and the faster processing speed computer.

<sup>c</sup> Data not collected.

**Freeway Test 2:** The second freeway test was performed from a pedestrian bridge crossing a major Puget Sound Interstate (I-405). This particular freeway section is comprised of three general purpose lanes and one high-occupancy vehicle (HOV) lane in each direction. Additionally, there is an on-ramp in one direction and an off-ramp in the other direction. There are sound barrier walls along this particular stretch of freeway on both sides of the freeway. This generally has the effect of channeling the noise upward

and thus leading to even higher noise levels than might normally be encountered for freeway conditions.

A test sample of 40 vehicle plates was collected over approximately a 10 minute period. Binoculars were used to read the plates of oncoming vehicles in the southbound lanes of the freeway. Visual verification of the computer processed voice input against the spoken entry was performed. All 40 entries were correctly recognized and processed by the program.

The main purpose of this test was just to determine whether this technology would function sufficiently under these very high ambient noise levels. The plates were read at a relatively unhurried rate so as not to introduce accuracy errors due to the pressure of reading plates at a high rate to obtain a large entry sample size.

Table 2 below illustrates the range of background noise levels encountered at two typical traffic locations, along with noise levels experienced in an office setting for comparison.

**Table 2. Summary of Ambient Sound Levels at Data Collection Sites**

|             | Office<br>Setting          | Arterial<br>Sites <sup>a</sup> | Freeway<br>Overpass        |
|-------------|----------------------------|--------------------------------|----------------------------|
| Sound Level | dB(A) / dB(C) <sup>b</sup> | dB(A) / dB(C) <sup>b</sup>     | dB(A) / dB(C) <sup>b</sup> |
| Minimum     | < 50 / 52                  | 64 / 73                        | 77 / 79                    |
| Average     | < 50 / 57                  | 71 / 80                        | 81 / 84                    |
| Maximum     | 56 / 64                    | 78 / 88                        | 85 / 90                    |

<sup>a</sup> These values are representative of all three arterial tests due to the proximity of location for all three tests.

<sup>b</sup> While the 'A' scale is typically used to report noise levels because it is more indicative of what is perceived by the human ear, the 'C' scale is more appropriate for this research because we are interested in the noise level that interferes with the audio processing capabilities of a full-frequency spectrum microphone. The 'C' scale uses a broader frequency response, which is closer in frequency response to that of a microphone.

### *4.3.2 Data Collection Experiences*

#### 4.3.2.1 Data Collection Personnel

While it was beyond the scope of this preliminary study to address specific issues related to data collection personnel, it should suffice at this point to say that the data collectors in all cases were familiar and comfortable with the data collection methods being used and very responsible and diligent in their work. Specific mention of weather during the two latter arterial tests was made because it raised another potential issue about the differences in the two data collection methods. While not appearing to be a problem during these tests, it is conceivable that very cold weather could affect the entry accuracy rate of a keyboard user because of cold hands and fingers. Consequently, the data collector may decide to wear gloves in this situation; but given the compactness of portable computer keyboards, this in itself might contribute to a higher entry error rate. The cold weather issue is definitely not as problematic for the voice input method, unless it is so cold it interferes with one's ability to speak.

#### 4.3.2.2 Character Entry Scheme

While it has typically been the case that people can enter numbers faster than letters on the keyboard, this is not the case with the voice input. There is really no difference between speaking numbers and letters, as they both are actually words. However, if the data collector has not thoroughly memorized the military alphabet, it is possible that "letters" could take a little longer to enter than numbers because of the memory recall factor for the data collector. Additionally, the voice processor does have slightly higher accuracy rates with the numbers than the letters. Thus, if it is only necessary to enter three numbers for a certain arterial application, the accuracy of the voice input could be exceptionally high. With each letter added to the voice entry, one can expect the accuracy rate to diminish slightly.

#### 4.3.2.3 Sample Size Issue

It did become apparent during the testing that the sampling frequency under high flow conditions was somewhat constrained by the execution speed of the data collection program. That is, during periods of high traffic flow on the arterial, the program was not able to process the speech input at the same rate the data collector was able to speak the license plates. Unlike a tape recorder that can record as quickly as one can speak, the data collection program still required a couple of seconds between voice entries to complete all of the necessary processing. Of course, this constraint was more conspicuous with slower processing computers. For the first arterial test, in which the slower 80486 computer was used, it is estimated that the sample percentage could have been 5 to 10 percent higher with the faster Pentium® computer used in the subsequent arterial tests.

A contributing factor to the rate at which the program can accept input is the whether the audio feedback is enabled for plate entry. That is, if the user chooses to receive audio feedback through the headphones to acknowledge when the program is ready to accept another plate entry, that will inherently reduce the sampling rate due to the increased time required to execute that portion of program code, as well as the length of the audio signal. For the experienced user familiar with the sampling rate of the program, this audio feedback is probably not necessary; but for the inexperienced user, this audio feedback is recommended. This is an important point, because if the user is not looking at the screen, and they do not have the audio feedback enabled, they really do not know for sure whether their last entry was accepted and processed. If the user attempts to enter a plate while the program is still processing the previous entry, it will be ignored. For the tests reported in this chapter, the audio feedback capability was not enabled.

## **4.4 Conclusions**

From these preliminary tests, it was demonstrated that it is possible to obtain accuracy rates comparable to or even greater than those obtained through the keyboard

entry method. It was also demonstrated that it is possible to obtain sample sizes comparable to or even greater than those obtained through the keyboard entry method. While background noise can impact the recognition accuracy of this system, it was found the program still achieved acceptable levels of performance under the typical background noise levels experienced at traffic data collection locations. At the very least, this pilot study demonstrated the viability of voice recognition technology for at least one area of traffic/transportation data collection.

Other issues related to field data collection that intuitively appear to benefit from this technology, but were not specifically evaluated in this pilot study include:

- Cold weather induced errors less likely,
- Greater physical comfort during data collection, since user doesn't have to balance a computer in their lap, and
- Reduction in data collection personnel, particularly where binoculars might be required.

Since the overall results were promising, it was decided to pursue program development of an improved version of this application when the next generation of voice recognition technology became available, as this could possibly alleviate many of the aforementioned performance issues.

## **CHAPTER 5**

### **32-BIT SPEECH RECOGNITION ENABLED PROGRAM**

---

#### **5.1 Introduction**

As noted in the previous section, the 16-bit program performance was less than optimal due to the processing time involved for plate entries. However, since the technology showed good potential overall, it was decided to pursue the development of a 32-bit based application when the technology became available. It was hoped that the performance increases inherent in 32-bit architecture versus 16-bit architecture would improve the application performance to an acceptable level.

#### **5.2 IBM 32-bit VoiceType Speech Recognition Engine**

IBM introduced a new speech recognition engine based on 32-bit architecture in early 1997. Thus, the development of a 32-bit application was based on a completely new voice recognition engine. Additionally, the development environment included Windows 95 and Visual Basic 5.0, both supporting 32-bit architecture.

Other than the 32-bit architecture, there was another fundamental difference in the functioning of the new speech recognition engine relative to the previous one. Whereas the previous voice recognition engine could not process new speech inputs until it was done processing the previous speech input, the new voice engine accepts voice inputs on a continuous basis and places new inputs in a queue for processing if received before it is finished processing the previous input.

#### **5.3 Program Development**

Program development consisted of several tasks: 1) convert existing 16-bit application to a format compatible with the 32-bit version of Microsoft® Visual Basic (Version 5.0); 2) revise user interface based upon user feedback from field tests, 3) update program code that interfaces with newer version of IBM's voice recognition engine, and 4) revise grammar file for more efficient input scheme.

While the application code is proprietary, more detail about basic program functioning can be found in appendix A.

### *5.3.1 Character Entry and Time-Stamping*

One of the main purposes of the grammar file is to define the speech input syntax that is acceptable to the application. For this application, there were several possibilities: 1) entering of either four, five, or six characters from a license plate; 2) having the program automatically enter and time-stamp the plate entry upon completion of the voice entry; 3) manually time-stamping the plate entry (e.g., through the use of a voice keyword, such as 'enter' or a keyboard keystroke) after completion of the plate entry; and 4) manually time-stamping the plate entry before voice entry of the license plate.

Depending on the particular application, there are several different ways of structuring the grammar file. If it is important that a very accurate time-stamp be applied to each entry, then a scheme that associates a particular keyword or keystroke to trigger the time stamp may be most appropriate. In other words, entering a license plate into the database and time-stamping it is associated with a separate keyword or keyboard stroke, which could occur either before or after the actual speaking of the license plate characters. If time-stamp accuracy is not as critical, such as for an origin-destination type study in which approximate times are usually adequate, it may be unnecessary to use a keyword to trigger the time-stamp and entry. The license plate entry could just be automatically entered upon completion of speaking the plate characters. This situation offers another advantage over the keyboard input method because it saves one extra step for the user—the explicit indication of time-stamping.

An enhancement to the voice recognition engine as a result of the 32-bit architecture over the 16-bit version was the ability to “queue” voice entries. With the 16-bit version, voice entries could not be made to the program until the voice recognition engine was completely done processing the most recent entry. However, with the newer 32-bit version, if a voice entry is made before the voice recognition engine is done processing the previous entry, the entry will be temporarily stored to memory. When the

voice recognition engine is done processing the previous entry, it will retrieve the next stored entry and process it. This potentially gives the data collector the opportunity to collect more entries than he/she otherwise would be able to when having to wait for each entry to be processed in full before issuing another.

To eliminate the potential for spurious noises or incomplete entries being entered into the program and “muddying” the data file with invalid data, program logic was added to ignore entries of less than the user specified plate entry size. Thus, if an extremely loud vehicle bypasses the data collector, and the program returns an attempt at recognizing the sound input, it will not be entered into the database unless it meets the minimum user-specified length (in number of characters).

### *5.3.2 Use of Audio Feedback*

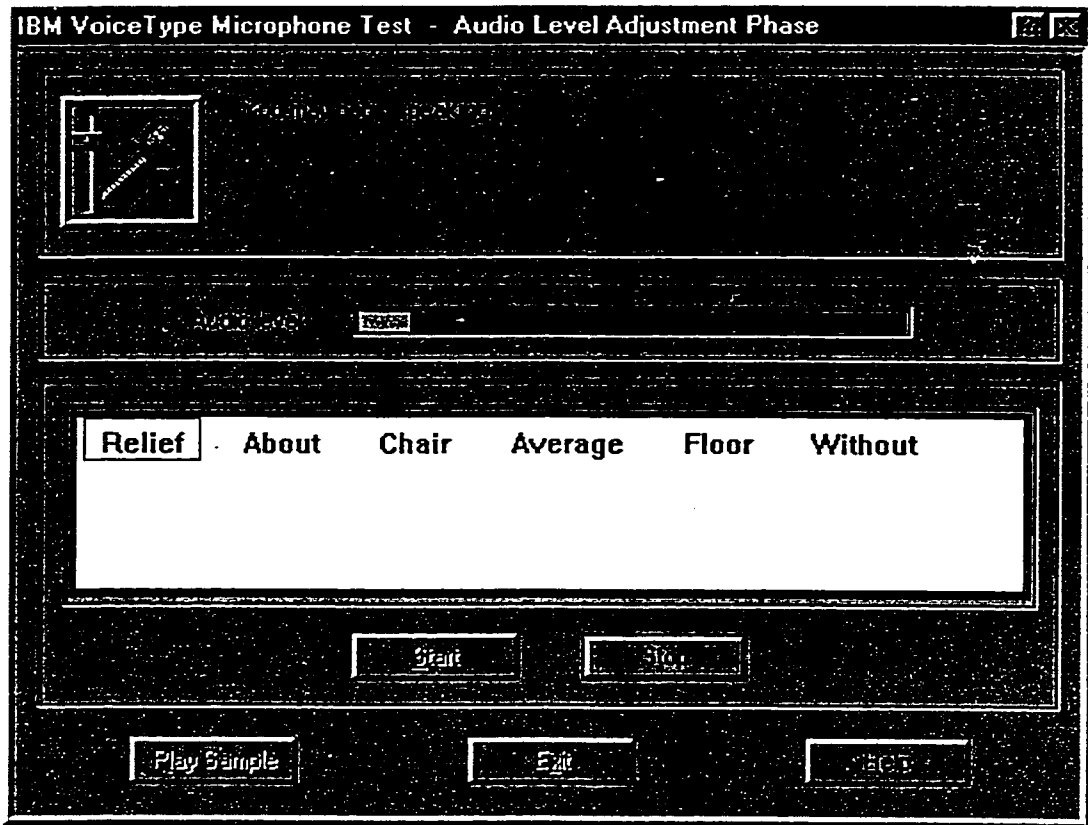
A significant issue in the program design was the issue of whether audio feedback was necessary for the user, and if so, how best to incorporate it. Since ideally with this application, the user would not need to look at the computer screen after data collection has started, using audio played through headset speakers might be necessary to keep the user informed of the program status during data collection. However, since computing performance is a significant issue, incorporating audio feedback into the program operation is not a trivial issue due to the computer processing time necessary to play audio files.

For the test application, very basic audio was incorporated in the form of simple beeps. This required the least processing “overhead” and still provided the user with program status feedback.

### *5.3.3 Background Noise*

As with the 16-bit version, determining and accommodating background noise conditions was necessary. Both the methodology and the process incorporated by IBM changed with the new speech recognition engine.

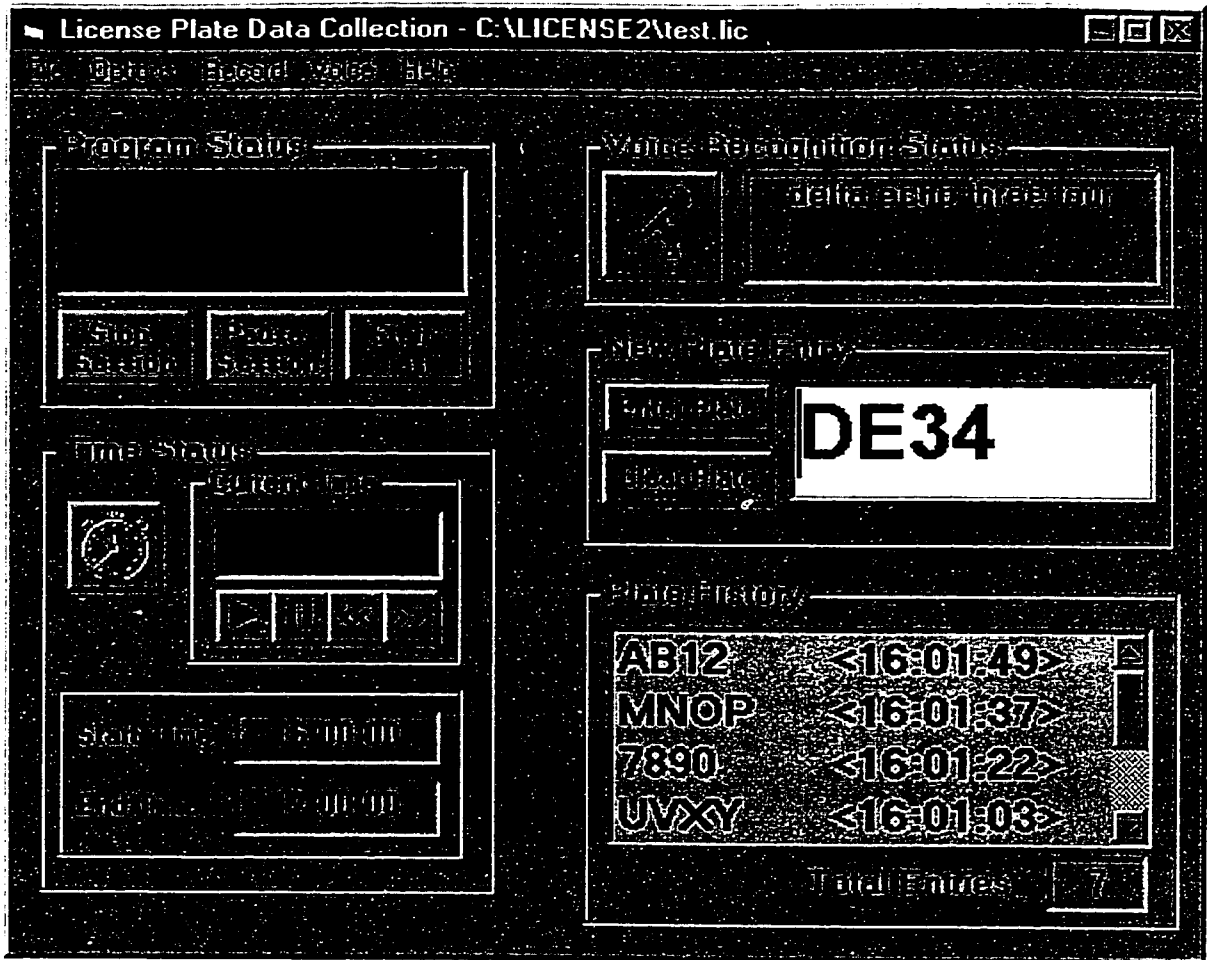
Basically, the microphone sensitivity level gets adjusted to a level which will still pick up sounds from a user's voice but not pick up background sounds. The new utility runs as a separate program as opposed to a utility incorporated directly into the data collection program. Figure 4 shows this utility's interface. Again, the user invokes this utility once in the desired location of data collection.



**Figure 4. Microphone Volume Level Adjustment**

#### 5.3.4 User Interface

Like the previous 16-bit version, the main intent of the program design was for it to be both functional in the field and easy to use. Based upon experiences from the previous version, the user interface (UI) was redesigned to make the program even easier and more intuitive to operate. Figure 5 shows the main user interface.



**Figure 5. Software Program User Interface (32-bit Version)**

Again, while most of the features of this user interface are self-explanatory, a few of the features deserve some discussion.

#### 5.3.4.1 Test Mode

Often times a user wants to test the speech recognition process before actually engaging in a data collection effort. Since there are several setup steps involved in preparing the program for a data collection effort, it was decided that there should be an easier and quicker way to perform strictly speech recognition testing. This was facilitated through inclusion of a Start Test button in Program Status portion of the UI. A single click of this button essentially puts the program into a normal operation mode. The

microphone button (in the Voice Recognition Status panel) must also be pressed as usual for engaging the speech recognition. The main difference in the functionality is that no data are saved to disk during test mode operation.

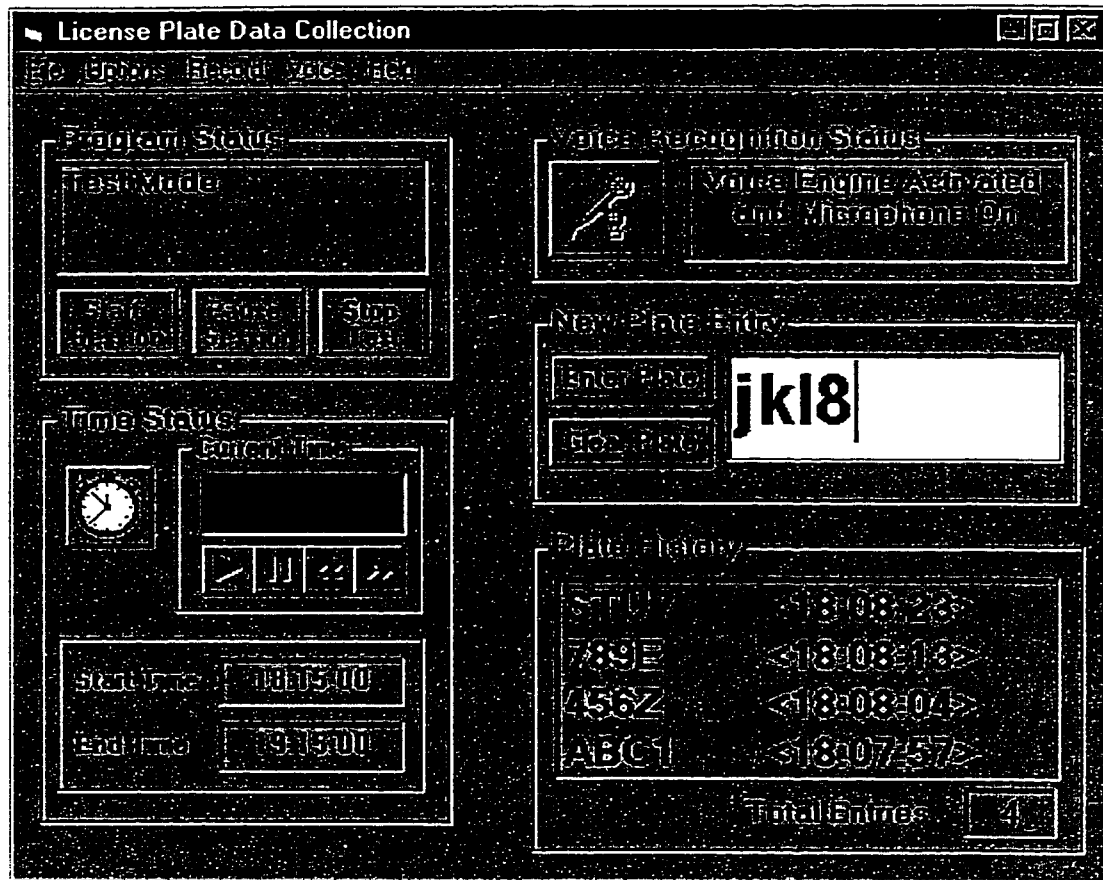
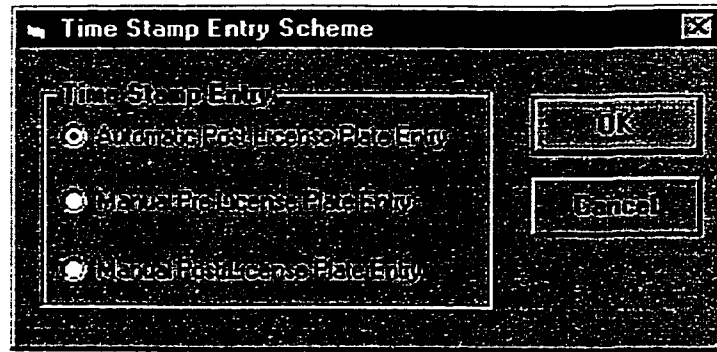


Figure 6. Program Operation in Test Mode

#### 5.3.4.2 Plate Entry Scheme

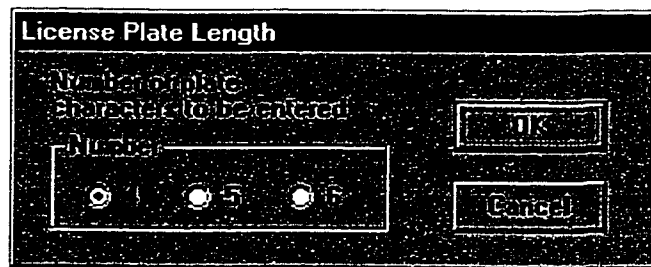
As discussed in section 5.3.1, the user may have differing needs for how the computer generated time stamp should be applied. The dialog box pictured below allows the user to select the desired time stamp entry scheme.



**Figure 7. Time Stamp Entry Scheme Dialog Box**

#### 5.3.4.3 Number of Plate Characters

As previously mentioned, the grammar file displayed in Figure 1 is of very high perplexity—meaning there are many (millions) of possible voice entry combinations. In order to reduce the perplexity, three different grammar files were defined. One file for each of the possible number of plate characters to be entered—four, five, or six. Thus, instead of one file accounting for each of these possible number of entered characters, the appropriate grammar file was invoked based upon the user selection in the dialog box shown below.



**Figure 8. Number of License Plate Characters Input Dialog Box**

## **5.4 Research Design**

### *5.4.1 Introduction*

Since the pilot study described in the preceding chapter had concluded that data collection by voice recognition was feasible, the main objective in developing this application was to attempt to improve the program performance and usability by

upgrading to a 32-bit architecture. After this program development, the key measures of effectiveness for this method would be compared to those of the computer keyboard entry method again. Another objective was to use this program for collecting travel time data, in order to evaluate the travel time results obtained through different sampling techniques with this technology.

#### *5.4.2 Key Measures of Effectiveness*

Again, the two key measures of effectiveness for comparison were accuracy rate of the inputted license plates and sample rate obtainable during certain traffic flow conditions over a specific time period. Another issue that was examined was the relative accuracy of timestamp entry values between the voice recognition and keyboard input methods. This issue is discussed in chapter 8.

#### *5.4.3 Data Collection*

Voice recognition equipment consisted of two high quality Sennheiser microphones and one 120 MHz Pentium laptop and one 133 MHz Pentium laptop.

Again, even with improvements to the 32-bit speech recognition engine, the use of individual letters for license plate entry was not feasible. Thus, testing continued with the use of the military alphabet (e.g., Alpha for A, Bravo for B, Charlie for C, etc.). Use of four characters for input was continued for these further tests.

Field testing was conducted mainly at several arterial locations, but also at a couple of freeway locations. Several tests were again conducted that consisted of collecting data at just a single location per time period. This allowed comparison of only the sampling and accuracy rates, as well as time-stamp accuracy. Two field tests were conducted in which data were collected simultaneously at two different locations along an arterial facility for the purpose of also evaluating the obtained travel time measurements. Table 3 below summarizes the field data collection experiments conducted for testing of the voice recognition application/technology.

Table 3. Data Collection Summary

| Date    | Method      | Start Time | End Time | Upstream Location                   | Downstream Location                | Dir.   | Traffic Conditions | Experiment Purpose                                  |
|---------|-------------|------------|----------|-------------------------------------|------------------------------------|--------|--------------------|---|
| 4/30/98 | voice       | 6:00 PM    | 6:40 PM  | NA                                  | I-5 / NE 145th                     | NB     | mostly free-flow   | sample rate, accuracy                               |
| 5/8/98  | kybd        | 5:15 PM    | 5:27 PM  | NA                                  | Montlake / N. Ped                  | NB     | free-flow          | time-stamp bias                                     |
| 5/19/98 | kybd        | 4:30 PM    | 5:00 PM  | Montlake / S. Ped                   | NA                                 | NB     | heavily platooned  | time-stamp accuracy vs. max sample size             |
| 7/1/98  | voice, kybd | 3:45 PM    | 4:30 PM  | Montlake / ped bridge               | NA                                 | NB     | heavily platooned  | voice rec vs. keyboard                              |
| 7/9/98  | voice, kybd | 3:45 PM    | 4:15 PM  | Montlake / ped bridge               | NA                                 | NB     | heavily platooned  | voice rec vs. keyboard                              |
| 7/23/98 | voice, kybd | 4:00 PM    | 5:00 PM  | 45 <sup>th</sup> / 15 <sup>th</sup> | 45 <sup>th</sup> / 7 <sup>th</sup> | WB     | Congested-signal   | voice rec vs. keyboard, and travel time measurement |
| 8/12/98 | voice, kybd | 2:45 PM    | 4:30 PM  | Pacific Way                         | NA                                 | EB     | Free-flow-signal   | voice rec vs. keyboard                              |
| 9/3/98  | voice, kybd | 4:00 PM    | 5:00 PM  | 45 <sup>th</sup> / 15 <sup>th</sup> | 45 <sup>th</sup> / 7 <sup>th</sup> | WB     | Congested-signal   | voice rec vs. keyboard, and travel time measurement |
| 9/29/98 | voice, kybd | 3:30 PM    | 6:00 PM  | SR-520 / Montlake                   | NA                                 | EB, WB | Mostly stop-and-go | sample rate, accuracy                               |

## 5.5 Results and Interpretation

Because of the shortcomings inherent in the 16-bit program version, when the technology became available, it was thought that a similar program written for a 32-bit platform would rectify some of these shortcomings.

Data were collected for maximum sample size, maximum entry accuracy, and maximum time-stamp accuracy. For the maximum sample size, a precise time-stamp was not a consideration, and voice entries were made as fast as the data collector could speak them, or as fast as the vehicles arrived, whichever was slower. For the maximum entry accuracy rate, the voice entry rate of license plates was such that that before the next entry was made, the user waited until they heard the beep signifying that the processing of the previous entry was complete. This is because it was discovered that sometimes the program can “hang” if entries were made too quickly and the queue of voice entries became large. For the maximum time-stamp accuracy, the objective was to time-stamp a license plate voice entry as close as possible to the actual time-stamp reference location. Thus, if the data collector felt that they would not get the vehicle plate entry time-stamped until after it crossed the reference point, they would ignore it.

Table 4 and Table 5 below provide a summary of the key statistics obtained from the data collection experiments that used the voice recognition method at the arterial and freeway locations, respectively. For table rows that contain both voice recognition and keyboard statistics, that means data were collected simultaneously using both methods, but with different data collectors of course.

Table 4. Summary of Arterial Data Collection Statistics

| Data Collection Date <sup>a</sup> | Time Period (min) | Flow Rate (vphpl) | Voice Recognition |                   |                       | Keyboard Entry  |                   |                       |
|-----------------------------------|-------------------|-------------------|-------------------|-------------------|-----------------------|-----------------|-------------------|-----------------------|
|                                   |                   |                   | Sample Rate (%)   | Accuracy Rate (%) | Total Correct Entries | Sample Rate (%) | Accuracy Rate (%) | Total Correct Entries |
| 7/1/98                            | 15                | 740               | 83.6              | 98.0              | 150                   | 85.8            | 96.8              | 150                   |
| 7/9/98 (1)                        | 15                | 848               | 83.3              | 89.9              | 152                   | 70.8            | 98.6              | 144                   |
| 7/9/98 (2)                        | 15                | 756               | 93.5              | 92.4              | 157                   | 58.6            | 96.3              | 105                   |
| 7/23/98 (1,1)                     | 25                | 485               | 83.3              | 87.4              | 125                   | 87.1            | 90.6              | 135                   |
| 7/23/98 (1,2)                     | 25                | 636               | 91.2              | 94.3              | 198                   | 79.0            | 97.3              | 182                   |
| 7/23/98 (2,1)                     | 25                | 442               | 63.0 <sup>b</sup> | 88.7              | 86                    | 88.9            | 96.9              | 127                   |
| 7/23/98 (2,2)                     | 25                | 629               | 69.0 <sup>b</sup> | 98.0              | 148                   | 79.2            | 99.4              | 174                   |
| 8/12/98 (1)                       | 15                | 432               | 96.3              | 97.1              | 101                   | NA              | NA                | NA                    |
| 8/12/98 (1)                       | 15                | 432               | 89.8              | 89.6              | 86                    | NA              | NA                | NA                    |
| 8/12/98 (2)                       | 15                | 388               | 90.7              | 86.2              | 75                    | 82.5            | 98.8              | 79                    |
| 8/12/98 (3)                       | 15                | 376               | 96.8              | 95.6              | 86                    | 77.7            | 93.0              | 66                    |
| 8/12/98 (4)                       | 15                | 432               | NA                | NA                | NA                    | 71.3            | 98.6              | 73                    |
| 8/12/98 (4)                       | 15                | 432               | NA                | NA                | NA                    | 75.9            | 100.0             | 78                    |
| 9/3/98 (1,1)                      | 23                | 389               | 89.3              | 86.9              | 93                    | 98.4            | 94.8              | 110                   |
| 9/3/98 (1,2)                      | 23                | 562               | 93.9              | 97.2              | 176                   | 68.2            | 94.1              | 127                   |
| 9/3/98 (2,2)                      | 15                | 612               | 63.8 <sup>c</sup> | 95.3              | 81                    | 63.0            | 94.2              | 81                    |

<sup>a</sup> The first number in parentheses is the data collection session. The second number in parentheses is the data collection station, 1 for upstream, 2 for downstream.

<sup>b</sup> This data collection used the post-license plate entry manual time stamp. Additionally, there was a power glitch during the middle of the session at the downstream station that resulted in 6 potential vehicles not being recorded.

<sup>c</sup> This data collection used the pre-license plate entry manual time stamp.

Table 5. Summary of Freeway Data Collection Statistics

| Data Collection Date <sup>a</sup> | Time Period (minutes) | Flow Rate (vphpl) | Voice Recognition |                   |                       |                 | Keyboard Entry    |                        |  |
|-----------------------------------|-----------------------|-------------------|-------------------|-------------------|-----------------------|-----------------|-------------------|------------------------|--|
|                                   |                       |                   | Sample Rate (%)   | Accuracy Rate (%) | Total Correct Entries | Sample Rate (%) | Accuracy Rate (%) | Total Correct Entries  |  |
| 4/30/98 <sup>b</sup>              | 15                    | 960               | 50.9              | 82.1              | 87                    | N/A             | N/A               | N/A                    |  |
| 9/29/98 <sup>c</sup> (1)          | 10                    | 1374              | 83.9              | 78.9              | 135                   | N/A             | N/A               | N/A                    |  |
| 9/29/98 <sup>c</sup> (2)          | 10                    | 1308              | N/A               | N/A               | N/A                   | 80.6            | 97.4              | 148                    |  |
| 9/29/98 <sup>c,e</sup> (3)        | 10                    | 1446              | N/A               | N/A               | N/A                   | 61.2            | 96.3              | 130                    |  |
| 9/29/98 <sup>c</sup> (4)          | 10 <sup>f</sup>       | 1466              | 77.8              | 97.8              | 186                   | N/A             | N/A               | N/A                    |  |
| 9/29/98 <sup>d</sup> (5)          | 10                    | 1452              | 80.3              | 90.8              | 167                   | N/A             | N/A               | N/A                    |  |
| 9/29/98 <sup>d</sup> (6)          | 10                    | 1356              | N/A               | N/A               | N/A                   | 57.6            | 93.0              | 119                    |  |
| 9/29/98 <sup>d,e</sup> (7)        | 6                     | 1480              | N/A               | N/A               | N/A                   | 78.1            | 93.6              | 102 / 170 <sup>g</sup> |  |

<sup>a</sup> The first number in parentheses is the data collection session.

<sup>b</sup> This data collection suffered from accuracy problems due to difficulty of reading license plates at high speed with low powered binoculars; thus, most of the accuracy errors were related to trouble reading the plate correctly, not the voice recognition application misinterpreting the spoken phrase. The sample rate, however, is still reasonably representative of what could be obtained. Since this was one of the first field tests of the new voice recognition application, a couple of program bugs were also discovered, which were subsequently fixed; thus, these were definitely not optimal results. It was obvious, however, that higher powered binoculars were still necessary.

<sup>c</sup> Since vehicles were moving very slowly (stop-and-go traffic) and the observation overpass was relatively low, the license plates were able to be read without the assistance of binoculars.

<sup>d</sup> Although the conditions were similar to those in footnote c (except other traffic direction was collected), the data collectors were instructed to use binoculars (low power) anyway.

<sup>e</sup> This data collection was performed by two people simultaneously—one person reading the license plates aloud, the other typing them into the program.

<sup>f</sup> Video camera died soon after data collection began, but data collection continued for another 6 minutes. Sample rate and accuracy rate are based on 2 1/4 minute video confirmed statistics. Total correct entries was extrapolated using first 2 1/4 minute statistics and additional 6 minutes of data collected.

<sup>g</sup> Extrapolated to 10 minutes.

In the following tables, Table 6 and Table 7, data collection statistics for both the keyboard entry method and voice entry method are aggregated by data collection scenario data collector, respectively. There are three basic data collection scenarios. Each scenario may be comprised of more than one data collection session, and more than one data entry scheme, but the site and flow conditions under which data were collected were similar for each scenario. Some of the scenarios are further subdivided (denoted by a following 1, 2, or 3) based upon the data entry scheme used (e.g., maximum time-stamp accuracy, manual post license plate entry time-stamp, etc.). In this table, any data corresponding with a particular scenario and subdivision number can be compared against any other data with the same scenario and subdivision number. For example, the “Keyboard 1-1” data collection statistics can be compared amongst data collectors 2, 3, and 4. Additionally, the “Keyboard 1-1” statistics for data collector 3 can be compared to the “VR 1-1” statistics for both data collectors 3 and 4.

Two natural questions arise:

- Is there a difference in the data collection statistics (sample rate, accuracy rate) between the keyboard entry method and the voice entry method?
- Is there a difference between the data collection statistics (sample rate, accuracy rate) for different data collectors?

Unfortunately, designing an experiment to account for all the various factors that affect the MOE statistics (data collection method, data collection personnel, data entry scheme, and traffic flow conditions) would require an extremely large number of runs, even without replication. Additionally, since personnel resources were limited, there were only four different data collectors utilized in all of these experiments. For a thorough investigation of the effect of different data collectors on the results of either method, a larger number of personnel would be required. Nonetheless, several comparisons can still be performed, with the knowledge that all of the various effects cannot be completely isolated from one another.

Table 6. Results Summary by Method and Traffic Condition

| Data Collector | Data Set       | Total Volume | Correct | Incorrect | Unknown | Sample Rate  | Accuracy Rate | Correct Entries per Hour |
|----------------|----------------|--------------|---------|-----------|---------|--------------|---------------|--------------------------|
| 2              | Keyboard 1-1   | 183          | 150     | 5         | 2       | 85.8%        | 96.8%         | 603                      |
| 3              | Keyboard 1-1   | 186          | 105     | 4         | 0       | 58.6%        | 96.3%         | 425                      |
| 4              | Keyboard 1-1   | 387          | 266     | 3         | 5       | 70.8%        | 98.9%         | 539                      |
|                | <b>Average</b> |              |         |           |         | <b>71.7%</b> | <b>97.3%</b>  | <b>522</b>               |
| 3              | VR 1-1         | 209          | 152     | 17        | 5       | 83.3%        | 89.9%         | 615                      |
| 4              | VR 1-1         | 369          | 307     | 16        | 4       | 88.6%        | 95.0%         | 628                      |
|                | <b>Average</b> |              |         |           |         | <b>85.9%</b> | <b>92.5%</b>  | <b>622</b>               |
| 2              | Keyboard 1-2   | 234          | 150     | 3         | 3       | 66.7%        | 98.0%         | 603                      |
| 4              | Keyboard 1-2   | 220          | 154     | 5         | 3       | 73.6%        | 96.9%         | 623                      |
|                | <b>Average</b> |              |         |           |         | <b>70.2%</b> | <b>97.4%</b>  | <b>613</b>               |
| 2              | Keyboard 1-3   | 220          | 105     | 4         | 1       | 50.0%        | 96.3%         | 425                      |
| 4              | Keyboard 1-3   | 415          | 206     | 4         | 5       | 51.8%        | 98.1%         | 462                      |
|                | <b>Average</b> |              |         |           |         | <b>50.9%</b> | <b>97.2%</b>  | <b>444</b>               |
| 3              | Keyboard 2     | 202          | 144     | 5         | 6       | 76.7%        | 96.6%         | 293                      |
| 4              | Keyboard 2     | 205          | 152     | 2         | 3       | 76.6%        | 98.7%         | 315                      |
|                | <b>Average</b> |              |         |           |         | <b>76.7%</b> | <b>97.7%</b>  | <b>304</b>               |
| 3              | VR 2           | 205          | 161     | 22        | 2       | 90.2%        | 88.0%         | 337                      |
| 4              | VR 2           | 202          | 187     | 7         | 1       | 96.5%        | 96.4%         | 384                      |
|                | <b>Average</b> |              |         |           |         | <b>93.4%</b> | <b>92.2%</b>  | <b>361</b>               |
| 1              | Keyboard 3-1   | 475          | 372     | 24        | 32      | 90.1%        | 93.9%         | 312                      |
| 2              | Keyboard 3-1   | 808          | 564     | 19        | 6       | 72.9%        | 96.7%         | 395                      |
|                | <b>Average</b> |              |         |           |         | <b>81.5%</b> | <b>95.3%</b>  | <b>354</b>               |
| 3              | VR 3-1         | 313          | 218     | 32        | 14      | 84.3%        | 87.2%         | 278                      |
| 4              | VR 3-1         | 439          | 374     | 17        | 12      | 91.8%        | 95.7%         | 472                      |
|                | <b>Average</b> |              |         |           |         | <b>88.1%</b> | <b>91.5%</b>  | <b>375</b>               |
| 3              | VR 3-2         | 162          | 86      | 11        | 5       | 63.0%        | 88.7%         | 211                      |
| 4              | VR 3-2         | 226          | 148     | 3         | 5       | 69.0%        | 98.0%         | 361                      |
|                | <b>Average</b> |              |         |           |         | <b>66.0%</b> | <b>93.3%</b>  | <b>286</b>               |
| 4              | VR 3-3         | 138          | 81      | 4         | 3       | 63.8%        | 95.3%         | 358                      |

Table 7. Results Summary by Data Collector

| Data Collector | Data Set     | Total Volume | Correct | Incorrect | Unknown | Sample Rate | Accuracy Rate |
|----------------|--------------|--------------|---------|-----------|---------|-------------|---------------|
| 1              | Keyboard 3-1 | 475          | 372     | 24        | 32      | 90.1%       | 93.9%         |
| 2              | Keyboard 1-1 | 183          | 150     | 5         | 2       | 85.8%       | 96.8%         |
|                | Keyboard 1-2 | 234          | 150     | 3         | 3       | 66.7%       | 98.0%         |
|                | Keyboard 1-3 | 220          | 105     | 4         | 1       | 50.0%       | 96.3%         |
|                | Keyboard 3-1 | 808          | 564     | 19        | 6       | 72.9%       | 96.7%         |
| 3              | Keyboard 1-1 | 186          | 105     | 4         | 0       | 58.6%       | 96.3%         |
|                | Keyboard 2   | 202          | 144     | 5         | 6       | 76.7%       | 96.6%         |
|                | VR 1-1       | 209          | 152     | 17        | 5       | 83.3%       | 89.9%         |
|                | VR 2         | 205          | 161     | 22        | 2       | 90.2%       | 88.0%         |
|                | VR 3-1       | 313          | 218     | 32        | 14      | 84.3%       | 87.2%         |
|                | VR 3-2       | 162          | 86      | 11        | 5       | 63.0%       | 88.7%         |
| 4              | Keyboard 1-1 | 387          | 266     | 3         | 5       | 70.8%       | 98.9%         |
|                | Keyboard 1-2 | 220          | 154     | 5         | 3       | 73.6%       | 96.9%         |
|                | Keyboard 1-3 | 415          | 206     | 4         | 5       | 51.8%       | 98.1%         |
|                | Keyboard 2   | 205          | 152     | 2         | 3       | 76.6%       | 98.7%         |
|                | VR 1-1       | 369          | 307     | 16        | 4       | 88.6%       | 95.0%         |
|                | VR 2         | 202          | 187     | 7         | 1       | 96.5%       | 96.4%         |
|                | VR 3-1       | 439          | 374     | 17        | 12      | 91.8%       | 95.7%         |
|                | VR 3-2       | 226          | 148     | 3         | 5       | 69.0%       | 98.0%         |
|                | VR 3-3       | 138          | 81      | 4         | 3       | 63.8%       | 95.3%         |

### 5.5.1 Performance of Keyboard Entry Versus Voice Recognition

For the '1-1' data set, the keyboard entry method had a lower average sample rate, but higher average accuracy rate. When considered together, the voice recognition method resulted in 19% more correct entries over the keyboard method. Comparing the '2' data set, again the keyboard entry method had a lower average sample rate and a higher average accuracy rate. But again, the net result was 19% more correct entries for the voice recognition method. Comparing the '3-1' data set, the keyboard entry method again had a lower average sample rate and a higher average accuracy rate. The net effect was still more correct entries for the voice recognition method, but only a 6% increase for

this data set. It should be noted this data set consists of simultaneous data collection at two different locations on an arterial. Although the general flow conditions were similar, there was a difference in volume between the two locations. However, there was one of each data collection method (i.e., keyboard entry and voice recognition) being used at each location.

Data sets 'keyboard 1-2' and 'keyboard 1-3' were done with different sampling schemes, namely maximum sample size and maximum time-stamp accuracy. Further discussion about these results is contained in section 8.3.1.

### *5.5.2 Effect of Different Time-Stamp Entry Schemes*

Data set '3-2' is based on a post-license plate entry timestamp technique for the voice recognition method. Data set '3-1' is based on a pre-license plate entry timestamp technique for the voice recognition method. The other voice recognition data sets were based on an automatic timestamp entry technique. For the post-license plate entry timestamp technique, the word 'enter' was used as the "trigger" word to notify the program when to make the timestamp. For the pre-license plate entry timestamp technique, the keyboard spacebar was used as the "trigger" for the program to make the timestamp. The difference in results from the different timestamp techniques is discussed in section 8.3.1.

### *5.5.3 Effect of Different User Voices*

Comparing all of the voice recognition data sets between data collector 3 and 4, it can be seen that data collector 4 consistently obtained higher sample rates and accuracy rates than data collector 3. Data collector 4 was the developer of the program and the author of this document. Given this author's more detailed knowledge of the voice recognition application, it is not unreasonable to expect better results when the author used this method versus somebody with less experience.

Comparing data set 'keyboard 1-1', it can be seen that data collectors 2 and 4 obtained better sample rate statistics for the first data set than data collector 3, but similar

accuracy rate results. Comparing data set 'keyboard 2', the results were comparable between data collectors 3 and 4. Thus, it is difficult to say whether data collector 3 was less proficient with just the voice recognition method, or less proficient with both methods.

Although there was only one set of traffic and site conditions under which data collector 1 collected data, it can be seen that this person obtained a much higher sample rate than data collector 2 when comparing the 'keyboard 3-1' data sets. However, data collector 2 did obtain a slightly higher accuracy rate than data collector 1. Again, it should be noted that these two data collectors were positioned at different locations on the same arterial, thus, the difference in traffic volume (higher at data collector 2's location) may have had an effect on the sample rate.

While it is not possible to completely isolate the effects of the technology versus those of the individual user with this data, it is clear that personnel training is certainly a key factor to the success of any method. It is not anticipated that good results will be obtained with either method by untrained users. It does appear, however, that the voice recognition method requires more user training at this point than the keyboard method, especially taking into consideration the need for data collection personnel to become familiar with the military alphabet, or some facsimile thereof. Despite adequate training, it is important to remember that when recruiting personnel for data collection, there will always be a range of capabilities for the personnel that will be translated in the effectiveness of the method employed. While the critical factor for the keyboard method is a person's familiarity with the keyboard and typing speed, the critical factor for the voice recognition method is a person's ability to speak loudly and clearly (i.e., enunciate well) and familiarity with the military alphabet.

It should also be noted that in a limited trial of the voices of two different women, recognition accuracy appeared markedly lower than those of two tested male voices. This quite possibly could have been a random result for these two women, or there may be something inherent in the female voice that led to these perceived results. The only

plausible explanation for this might be that the voice recognition engine utilizes a lot of audio information at the lower end of the frequency spectrum, in which case, the female voice would probably be more difficult to interpret because it generally is higher in frequency than a male's voice.

## 5.6 Conclusions and Recommendations

Although quite impressive at this time, voice recognition technology is still relatively in its infancy. As this technology becomes even more advanced, and likewise the computer hardware running this technology, the need for using a keyboard for input could become a thing of the past. Many researchers are currently experimenting with the application of voice recognition to many different fields, and there are certainly many areas just within the realm of traffic data acquisition that are readily adaptable to voice recognition technology. This research demonstrates the viability of voice recognition technology for just one of these areas of traffic/transportation data collection.

While entry accuracy is certainly an important measure, sample size in some cases can more than compensate for slightly lower accuracy rates. As demonstrated by many of the experiments, under certain flow conditions, a considerably larger sample was collected via voice recognition than with the keyboard, despite a slightly lower accuracy rate with the voice recognition. The net result was that there was a much larger pool of correct license plates from which to match using the voice recognition than the keyboard. This scenario is more typical under flow conditions which are fairly constant and not congested. Frequent gaps in the traffic stream, due to signals and/or platooning, or congested conditions will tend to result in more similar sample rates between the two methods; thus placing more importance on the accuracy entry.

A significant training factor for the voice recognition method at this time is the need to use the military alphabet. Although not difficult, it does require some time for memorization. Although the developed software application allows the user to bring up a dialog box ("cheat sheet") with the military alphabet words on the screen, best results are obtained when the user doesn't have to take their eyes off of traffic. Eventually, the

technology may get to the point where individual alphabet letters can be used, but this may be a while, given the difficulty even humans have interpreting individual letters spoken to one another.

Nonetheless, these studies have certainly demonstrated that the potential of this technology is very good at this time. With further refinement of the software in the future, the results can be expected to only get better. Generally, there is probably less variance in the sample rates and accuracy rates for the keyboard method from multiple data collectors at this time than the voice recognition method. However, with a little user training and practice, the voice recognition method has the capability for yielding greater sample rates and accuracy rates than the keyboard method. Under higher, uniform flow conditions, the voice recognition method can result in substantially higher sample rates, which will usually more than offset any potential difference in accuracy rate. Ultimately, with software and hardware improvements, the voice recognition input method should lead to both larger sample sizes and greater accuracy rates than the keyboard method due to the fact that the data collector never has to take his/her eyes off of the vehicles. This especially becomes more prevalent in situations where the observer may be required to use binoculars to see the license plates clearly.

The advantages of voice recognition technology would probably be most apparent with origin-destination studies, especially on freeways. With O-D studies, a higher sample rate will yield more precise results. Since very accurate time-stamping is not an issue for O-D studies, the voice recognition method will certainly yield greater sample sizes. In situations where binoculars are required to read the license plates, the voice recognition method will undoubtedly yield much greater sample rates.

## CHAPTER 6

# VIDEO IMAGE TRACKING AS A TOOL FOR DATA COLLECTION

---

### 6.1 Introduction

Whereas manual forms of travel time data collection are still in demand because of special studies, an automatic method is necessary to meet the demands of Intelligent Transportation Systems (ITS) applications. In particular, Advanced Traffic Management Systems (ATMS) will need to collect and report to the commuting public frequent travel time measurements for major routes, accurately and in real-time. A method that is integrated into the infrastructure of those major routes is the only feasible way to provide this information continually and consistently. Several current technologies for automatically providing travel time measurements were discussed in chapter two. One of these technologies, video image tracking, was preliminarily evaluated for its potential to provide real-time unbiased estimates of travel time. Of the technologies currently being investigated for this potential, video image tracking appears to be the most comprehensive solution for the needs of traffic management centers. Since many progressive traffic management centers have already heavily invested in video-based infrastructure for their major routes, it would be a natural transition to move to automatic video-based traffic parameters measurement. As previously mentioned, although other technologies, such as AVI, are being investigated for travel time measurement provision, these other technologies are more limited in the overall amount of traffic information they can provide. Video image tracking systems such as the Mobilizer have the capability to provide all standard traffic measurements (e.g., speed, volume, lane occupancy, headway), as well as travel time. From a traffic management center standpoint, relying on a single technology for its data collection needs instead of a combination of several would be more desirable, from both a cost efficiency and maintenance standpoint.

## 6.2 Mobilizer System

As mentioned in the State of the Art chapter, the Mobilizer system is a newer generation tracking system that offers the potential of uniquely identifying vehicles within a field-of-view (FOV). Additional system background information can be found in appendix C and a more comprehensive report [Nihan and Washburn, 1998].

Unlike other methods that might obtain its information by using license plates or other specific information that can be used to trace the vehicle back to the registered owner, video image tracking systems, such as the Mobilizer, use other vehicle features that still leave the registered owner anonymous. This is an important distinction in light of the public's concern over privacy issues.

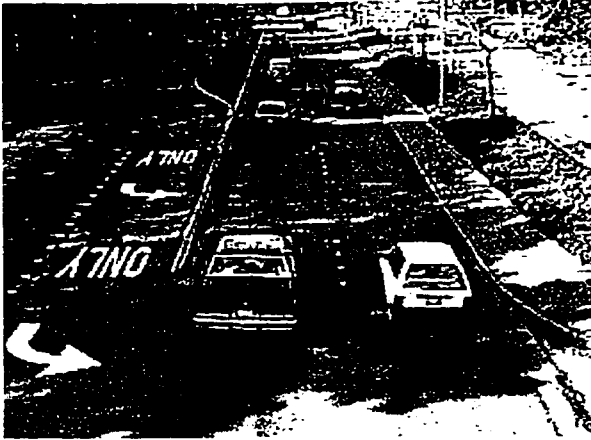
## 6.3 Research Design

The objective of this research was to evaluate the Mobilizer system for its ability to correctly match vehicles in successive FOV's and if so, its ability to provide statistically valid travel time estimates.

### 6.3.1 Site Description

Preliminary investigation was conducted utilizing data from two study sites. One location is a major arterial, the other a major freeway site. The initial studies consisted of one origin and destination point on each facility, with short spacing between them. This spacing was on the order of about 0.40 km (¼ mile) for each facility. This initial short spacing was used to limit the loss of vehicles between the two points and the number of lane changes; thus having a greater proportion of vehicles for which to test the matching algorithms of the Mobilizer system.

The arterial site consists of two lanes in each direction. There are two mid-section vehicle entry points (university parking lots) and one mid-section vehicle exit point within this study section. Figure 9 and Figure 10 show the upstream and downstream fields-of-views (FOV's), respectively.

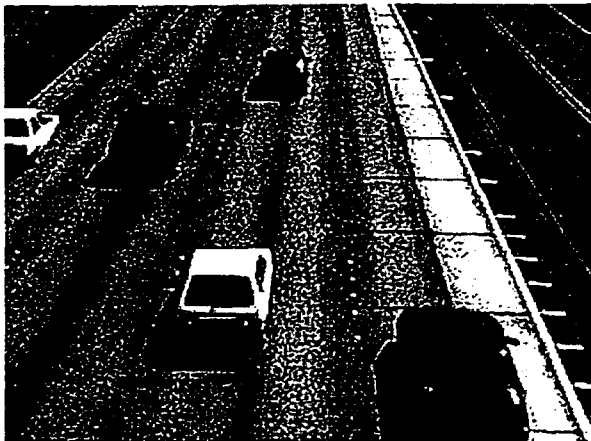


**Figure 9. Montlake Blvd.  
(Upstream Station)**

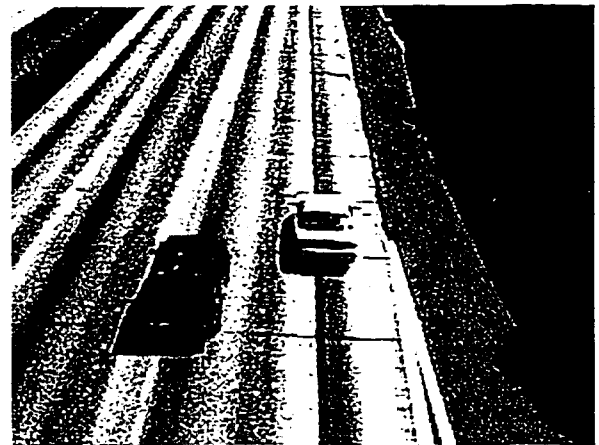


**Figure 10. Montlake Blvd.  
(Downstream Station)**

The freeway site consists of four general-purpose lanes in each direction. There is one mid-section vehicle entry point (NE 50<sup>th</sup> on-ramp) within this study section. Figure 11 and Figure 12 show the upstream and downstream FOV's, respectively.



**Figure 11. I-5/NE 50<sup>th</sup>  
(Upstream Station)**



**Figure 12. I-5/NE 45<sup>th</sup>  
(Downstream Station)**

### 6.3.2 Data Collection

For the arterial site, the video cameras were located directly over the middle of the two northbound lanes on pedestrian overpass bridges. The height of the camera on the

upstream overpass was just over 7 meters (23 feet). The downstream camera height was slightly higher. It was initially thought by the developer that the height of the camera on the pedestrian overpasses might not be high enough, but it turned out that this was not a problem. The camera height is important because if it is too low, it will be impossible to obtain a long enough FOV (for multiple looks at vehicles) while still obtaining a reasonable vehicle profile that contains not only the tops to of the vehicles, but also the back ends of the vehicles.

The video used for analysis was shot during the late afternoon and contained many shadows (both vehicle and street) at both stations (origin and destination) due to the angle of sun and many trees bordering the edge of the arterial. The presence of shadows was intermittent throughout the video because of clouds intermittently blocking the sun. The vehicle speeds during this time period varied from about 37 km/h (23 mph) to 64 km/h (40 mph). The speed limit for this roadway section is 56 km/h (35 mph).

For the freeway site, data were collected from just the two southbound outside lanes, with the video cameras being located directly over the middle of these two lanes on adjacent interchange overpass structures. The camera heights from these overpass locations were about 10½ meters (35 feet).

This video used for analysis was shot during the mid-afternoon and contained many significant vehicle shadows due to the angle of sun. The vehicle speeds during this time period varied from about 80 km/h (50 mph) to 109 km/h (68 mph). The speed limit for this roadway section is 96 km/h (60 mph).

This study utilized two high quality, 8mm video cameras (Canon L2), instead of standard WSDOT surveillance cameras. Using separate cameras allowed flexibility in positioning the cameras in optimal positions. A significant factor influencing vehicle tracking performance is camera positioning. Positioning the cameras directly over the travel lanes remedied two difficulties often encountered in video imaging analysis; 1) vehicle occlusion (i.e., one vehicle image overlapping another) due to offset-angle FOV perspectives, and 2) reduced vehicle image detail due to the physical distance of the

camera from the roadway and the necessary zoom levels. Again, see appendix C for more detail on these issues.

For both of these data collection efforts, a departing FOV was used. That is, the vehicles were moving away from the camera, rather than approaching it.

### 6.3.3 Reference Data

After processing the specified video streams, the Mobilizer system outputs a report that details every vehicle sighted in both FOV's. It also outputs graphics files that can be overlaid on the video image files, showing a unique identification (ID) number for each vehicle sighted in the traffic stream. With this information, the video image files can be visually reviewed and compared with the detailed output report to determine whether the correct vehicles in each FOV were matched. Due to the better placement/perspective of the cameras for this project and the better resolution provided by the Canon L2 cameras, visual confirmation of matched vehicles from the digitized image files was an easy process, albeit a time consuming one. The system allows the image files to be searched by frame or by vehicle ID; thus, each FOV image file can be searched for matching vehicle ID's based on the output file results, and a quick visual confirmation can be made as to whether or not the vehicles are a true match. This process was the primary basis for the verification of the Mobilizer reported results.

The video segments that were analyzed with the Mobilizer system were also observed manually for all possible matches. Since the original video quality was very good, it was relatively easy to visually match vehicles from two different fields-of-view. This was done for the both the arterial and freeway 15-minute analysis segments. For this "ground truthing", reference points were chosen in each FOV that corresponded to approximately the  $\frac{1}{4}$  distance point of the total length of the FOV.

For time-stamping the manual video entries, it was decided to round the times to the nearest one second. While it is possible to use a finer time resolution than this, one second was chosen for practical reasons. Because of the speed of vehicles and the camera shutter speed, it is not possible to stop the video at the point where each and every vehicle

is at the exact time-stamp reference point. The data collector just stopped the tape at the point where the vehicle was closest to the time-stamp reference point. Thus, most vehicles will either be a little past this point or a little before this point. Thus, it is very difficult to get the exact time-stamp at the exact reference location. However, since these time-stamp errors are essentially random, it is likely that these measurement errors will offset each other between the two stations, if not at each station independently. Certainly, if an extremely short section is considered, one block for example, a one second error could be substantial, considering that a vehicle traveling 60 mph covers 88 feet in one second. This could be further magnified if the errors at each recording location are in the opposite direction (e.g., one second too soon at the upstream location and one second too late at the downstream location) due to a time-stamp bias. But these roadway sections the roadway section lengths were much longer than a block, and the errors randomly alternated directions. One study [Bonneson and Fitts, 1995] indicated that video camera generated time-stamps can become inaccurate due to tape drag (on the order of  $-0.00020$  sec/sec for the type of camera employed in this research). This was not considered to be a problem for this data reduction since the amount of drag that could occur would likely be similar for both video tapes (upstream and downstream). Additionally, the ground truth data spanned a total of 15 minutes for both data sets, resulting in a possible time-stamp error of only 0.18 seconds at the end of the 15 minutes.

## **6.4 Results and Interpretation**

### **6.4.1 Arterial**

From the arterial video collected, a 15-minute piece of video was used for analysis. During the 15-minute period, 130 vehicles were matched from the traffic stream, from a population size of approximately 720 vehicles traveling through this section. This represents a sample rate of about 18 percent from this data set. It should be pointed out that the actual number of vehicles from which matches could be found would be less than 720, because some vehicles exit between the two stations, and some vehicles

enter between the two stations. Of these 130 vehicles matched, 99 of them were matched correctly. This corresponds to a match accuracy of about 76 percent. The following tables summarize these results, as well as the general descriptive statistics.

**Table 8. Arterial Summary Statistics**

| <b>Time Period (minutes)</b> | <b>Total Volume (vehicles)</b> | <b>Number of Matches</b> | <b>Percent Matches</b> | <b>Number Matches Correct</b> | <b>Percent Matches Correct</b> |
|------------------------------|--------------------------------|--------------------------|------------------------|-------------------------------|--------------------------------|
| 15                           | 720                            | 130                      | 18                     | 99                            | 76                             |

**Table 9. Arterial Travel Time Results**

|                              | <b>Minimum</b> | <b>Average</b> | <b>Maximum</b> |
|------------------------------|----------------|----------------|----------------|
| <b>Travel Time (seconds)</b> | 21.0           | 27.8           | 37.0           |
| <b>Travel Speed (km/h)</b>   | 37.2           | 49.6           | 65.3           |
| <b>Travel Speed (mph)</b>    | 23.1           | 30.8           | 40.6           |

A t-test was performed to determine if the average travel time for all 130 vehicle matches was significantly different from the average travel time obtained from just the 99 correct vehicle matches.

**Table 10. Arterial Travel Time Accuracy Results**

|                               | <b>All Matches</b> | <b>Correct Matches</b> |
|-------------------------------|--------------------|------------------------|
| Average Travel Time (seconds) | 27.41              | 27.99                  |
| Standard Deviation (seconds)  | 3.47               | 3.00                   |
| n (vehicles)                  | 130                | 99                     |
| t-stat                        | 1.33               |                        |
| t-ref (95%)                   | 1.97               |                        |
| p-value                       | 0.186              |                        |

As can be seen from the above table, at the 95 percent level of confidence, the difference in travel times between all matches and just the correct matches is not statistically significant.

The question also arises as to whether the match rate is adequate to estimate the true travel time between the two stations. After removing the incorrect matches, the

match rate becomes 14 percent. A general formula<sup>1</sup> can be applied to determine if this match rate is adequate.

$$N = \left( S \frac{K}{E} \right)^2 = \left( 3.00 \frac{1.96}{1} \right)^2 = 34.6 \rightarrow 35$$

where  $N$  = minimum number of matches  
 $S$  = estimated sample standard deviation, seconds  
 $K$  = constant corresponding to desired confidence level ( $1.96^2 = 95\%$ , two-tail)  
 $E$  = permitted error in the average travel time estimate, seconds

In this case, the 99 (14%) correct matches greatly exceeds the minimum number of 35 (5%) matches necessary for a 95 percent confidence level, allowing for a 1 second error in the average travel time estimate. Additionally, we can compute a 95 percent confidence interval for which the true average travel time is likely to fall, using the following formula.

$$\bar{x} \pm \frac{t_{\alpha/2, n-1} \times s}{\sqrt{n}} = 28.0 - \frac{1.99 \times 3.00}{\sqrt{99}} < \mu < 28.0 + \frac{1.99 \times 3.00}{\sqrt{99}} = 27.4 < \mu < 28.6$$

where  $\bar{x}$  = estimated sample mean, in seconds  
 $t$  = t-statistic reference value (for 98 d.f.)  
 $s$  = estimated sample standard deviation, seconds  
 $n$  = sample size  
 $\mu$  = mean of the population

While the basic statistical evaluations show that the Mobilizer system is capable of providing reliable travel time measurements, a more rigorous examination of the issue of statistical validity was warranted. Like any automated data collection system (as well as many manual methods), measurement bias is always a distinct possibility. This bias could be manifested in a vehicle selection bias and/or a time reference point bias. Thus, a better test involves comparing the Mobilizer travel time results with those obtained from

---

<sup>1</sup> More discussion on this formula can be found in chapter 7.

<sup>2</sup> This value actually assumes an infinite number of degrees of freedom (d.f.), but for sample sizes greater than 120 it is usually appropriate. Normally, an iterative procedure would be used to find a  $K$  value that corresponds with the sample size, but since there is

direct analysis of the original video. This analysis will yield more conclusive findings regarding the Mobilizer in providing unbiased accurate estimates of travel time under these conditions.

From the video ground truth data, 674 vehicle matches were identified for the 15-minute period. This was obtained from a volume of 704 vehicles at the origin and 731 vehicles at the destination. The travel time statistics for 674 vehicle matches yielded a mean of 27.8 seconds and a standard deviation of 2.8 seconds.

**Table 11. Travel Time Statistical Comparisons**

|                           | <b>Mobilizer<br/>All<br/>Matches</b> | <b>Mobilizer<br/>Correct<br/>Matches</b> | <b>Mobilizer<br/>Matches<br/>Ground Truth<sup>a</sup></b> | <b>Total<br/>Ground<br/>Truth</b> |
|---------------------------|--------------------------------------|--|---|-----------------------------------|
| Average Travel Time (sec) | 27.41                                | 27.99                                    | 28.13   | 27.75                             |
| Standard Deviation (sec)  | 3.47                                 | 3.00                                     | 2.99  | 2.83                              |
| n (vehicles)              | 130                                  | 99                                       | 99  | 674                               |
| F-stat                    |                                      |  |   | 1.32                              |
| F-crit (95%)              |                                      |  |   | 2.61                              |
| p-value                   |                                      |  |   | 0.267                             |

<sup>a</sup> These numbers are based on the identification of the 99 correct Mobilizer matches from the manual ground truthing of the original video.

These results indicate that there is no evidence that any of the three average travel times are statistically significantly different from one another.

---

such a large difference between the calculated sample size and number of samples in this case, the exercise would be purely academic.

Table 12. Travel Time Interval Statistics

| Time Interval  | Real Travel Time | Mobilizer Travel Time | # Mobilizer Matches | Upstream Volume | Percent Matches |
|----------------|------------------|-----------------------|---------------------|-----------------|-----------------|
| 1 <sup>a</sup> | 27.5             | 28.1                  | 2                   | 44              | 4.6             |
| 2              | 25.1             | 24.2                  | 7                   | 31              | 22.6            |
| 3              | 27.4             | 27.5                  | 5                   | 53              | 9.4             |
| 4              | 27.7             | 28.3                  | 6                   | 47              | 12.8            |
| 5              | 28.2             | 28.6                  | 9                   | 53              | 17.0            |
| 6              | 26.5             | 27.0                  | 5                   | 46              | 10.9            |
| 7              | 28.4             | 27.4                  | 4                   | 47              | 8.5             |
| 8              | 27.6             | 26.2                  | 4                   | 46              | 8.7             |
| 9              | 30.2             | 31.1                  | 12                  | 59              | 20.3            |
| 10             | 27.6             | 27.6                  | 4                   | 42              | 9.5             |
| 11             | 28.6             | 29.0                  | 11                  | 51              | 21.6            |
| 12             | 27.3             | 25.8                  | 5                   | 48              | 10.4            |
| 13             | 26.6             | 26.8                  | 5                   | 39              | 12.8            |
| 14             | 25.4             | 25.8                  | 7                   | 46              | 15.2            |
| 15             | 29.2             | 29.0                  | 10                  | 52              | 19.2            |
| <b>Average</b> |                  |                       | <b>6.4</b>          | <b>47</b>       | <b>13.6</b>     |

<sup>a</sup> There is always a startup lag during the first interval before the first match occurs. This is essentially a warm-up period, where the system is gathering a lot of information about the fields-of-view before the matching in earnest can begin.

This table shows that the Mobilizer matched from 4 to 12 vehicles per each one-minute interval (excluding the first interval). That translates to a range of about one in every 12 vehicles being matched to 1 in every 4 vehicles being matched. Under these low variance flow conditions, this range will yield valid travel time results.

It was also desired to see exactly where the Mobilizer-matched vehicles “fell” within the traffic stream. The following three figures (one for each 5 minutes of the total 15 minutes of data reduction) show a plot of the Mobilizer matches versus the population travel time matches.

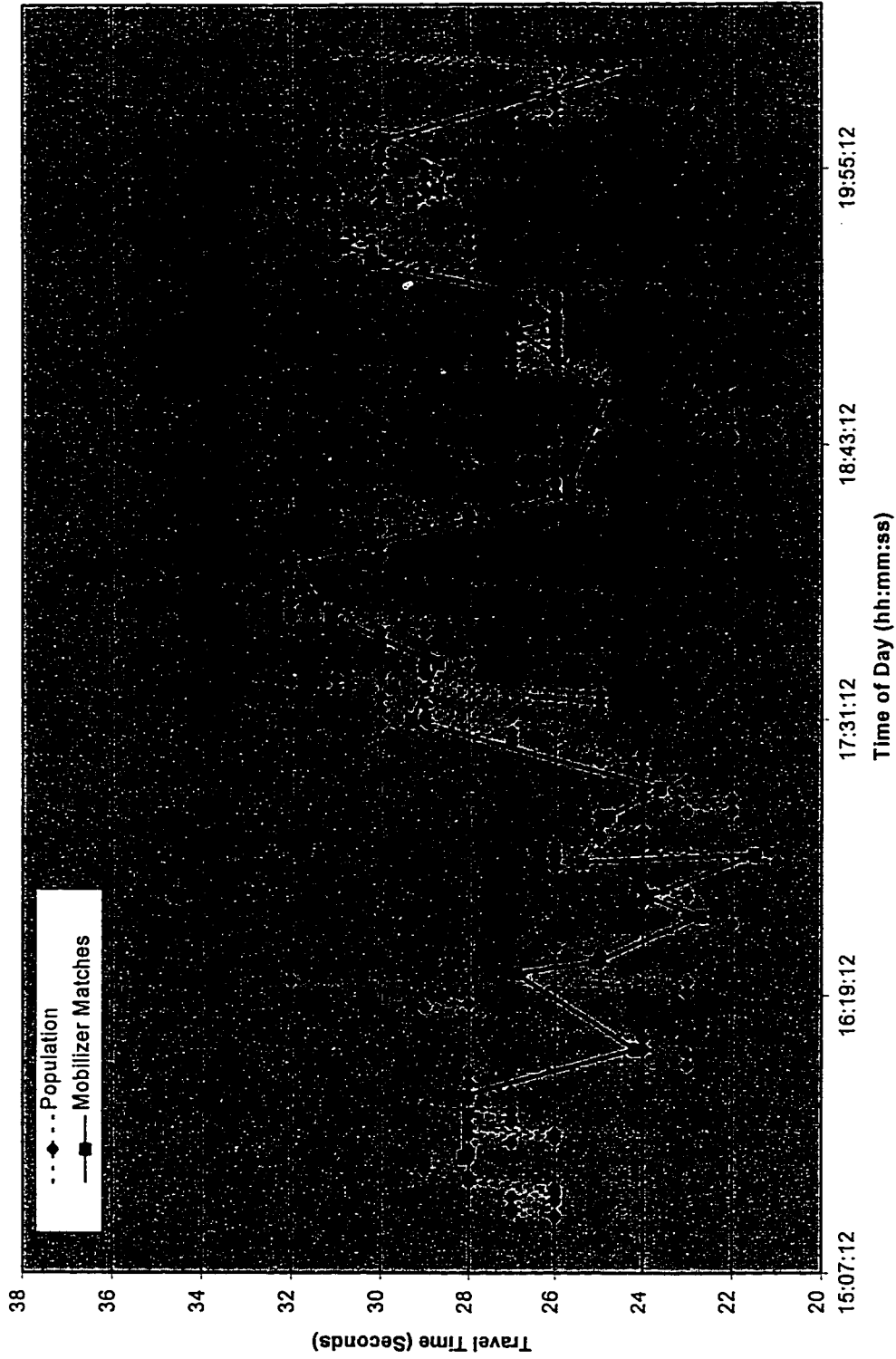


Figure 13. Montlake Travel Time Results (first 5-minutes)

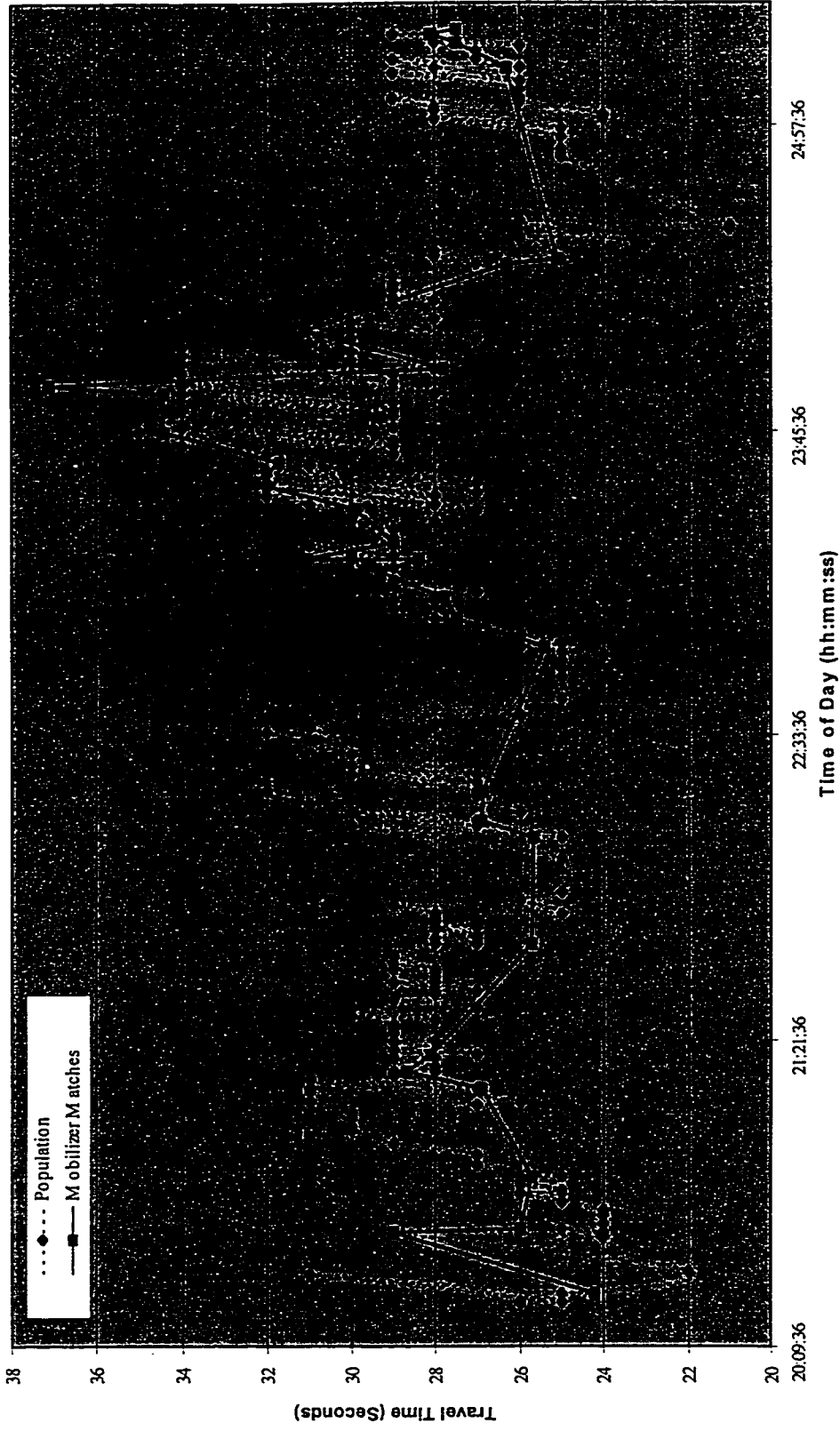


Figure 14. Montlake Travel Time Results (second 5-minutes)

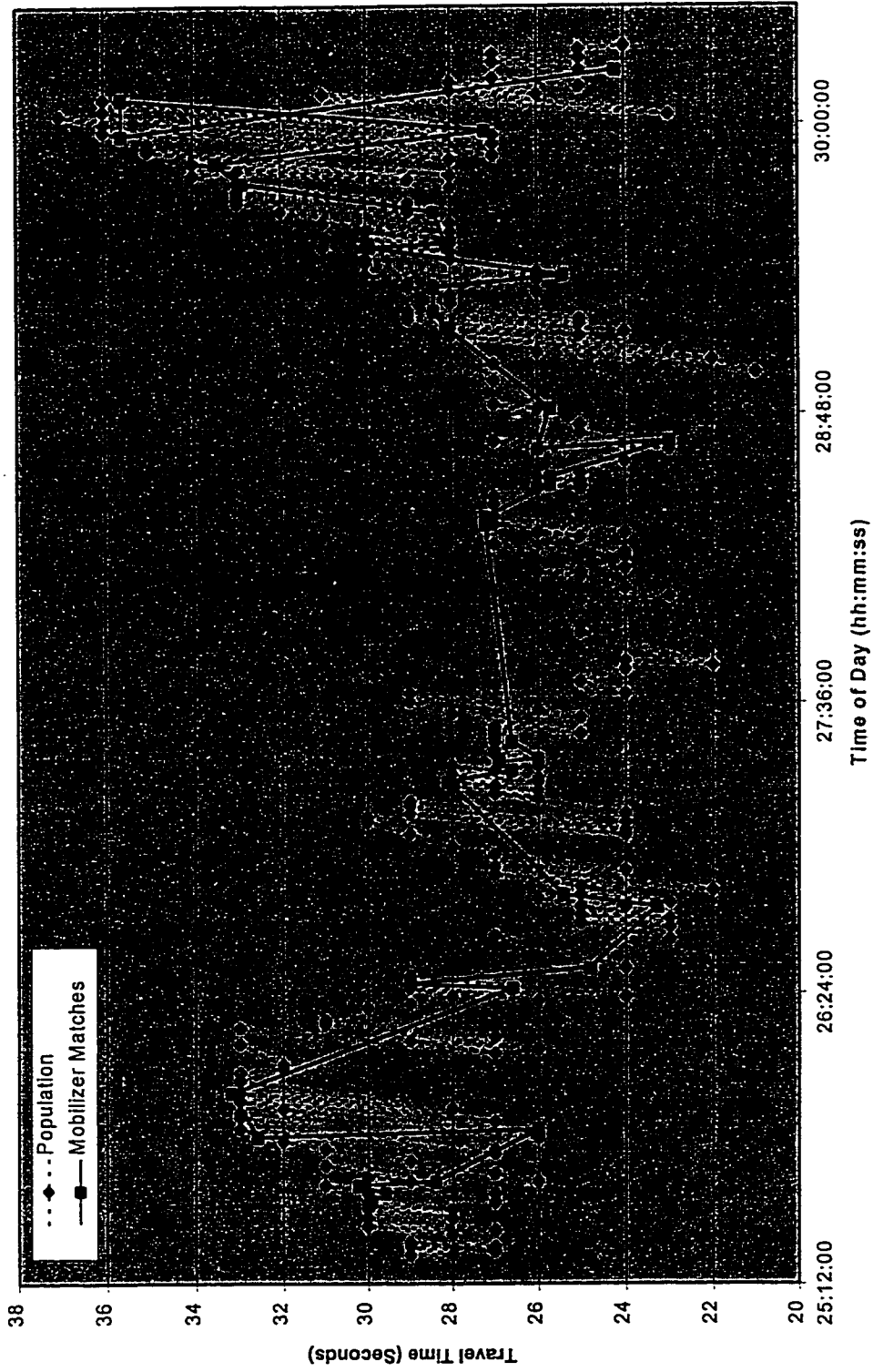


Figure 15. Montlake Travel Time Results (third 5-minutes)

From these plots, it appears that the Mobilizer matches seem to follow the trend of the total population match results fairly well. There are a couple of longer time gaps with no matches, but the overall travel time measurement coverage seems reasonable. While the Mobilizer matches generally avoid the extreme values (either really slow or really fast vehicles), there does not appear to be a propensity for matching either slower or faster vehicles with any greater frequency than the other.

#### 6.4.2 Freeway

From the freeway video collected, a 15-minute piece of video was used for analysis. During the 15-minute period, 46 vehicles were matched from the traffic stream, from a population size of approximately 420 vehicles traveling through this section. This represents a sample rate of about 11 percent from this data set.

Of these 46 vehicles matched, 36 of them were matched correctly. This corresponds to a match accuracy of about 78 percent. A t-test was performed to determine if the average travel time for all vehicle matches was significantly different from the average travel time obtained from just the correct vehicle matches. The following tables summarize these results, as well as the general descriptive statistics.

**Table 13. Freeway Summary Statistics**

| <b>Time Period (minutes)</b> | <b>Total Volume (vehicles)</b> | <b>Number of Matches</b> | <b>Percent Matches</b> | <b>Number Matches Correct</b> | <b>Percent Matches Correct</b> |
|------------------------------|--------------------------------|--------------------------|------------------------|-------------------------------|--------------------------------|
| 15                           | 420                            | 46                       | 11                     | 35                            | 76                             |

**Table 14. Freeway Travel Time Results**

|                              | <b>Minimum</b> | <b>Average</b> | <b>Maximum</b> |
|------------------------------|----------------|----------------|----------------|
| <b>Travel Time (seconds)</b> | 14.0           | 16.0           | 19.25          |
| <b>Travel Speed (km/h)</b>   | 79.8           | 96.2           | 109.7          |
| <b>Travel Speed (mph)</b>    | 49.6           | 59.8           | 68.2           |

**Table 15. Freeway Travel Time Accuracy Results**

|                               | All Matches | Correct Matches |
|-------------------------------|-------------|-----------------|
| Average Travel Time (seconds) | 16.19       | 15.96           |
| Standard Deviation (seconds)  | 1.46        | 1.14            |
| n (vehicles)                  | 46          | 35              |
| t-stat                        | 0.76        |                 |
| t-ref (95%)                   | 1.99        |                 |
| p-value                       | 0.451       |                 |

As can be seen from the above table, at the 95 percent level of confidence, the difference in travel times between all matches and just the correct matches is not statistically significant.

Again, to determine whether a sample size of 35 correct vehicle matches is sufficient for this data set, the following calculation is performed.

$$N = \left( S \frac{K}{E} \right)^2 = \left( 1.14 \frac{2.03}{0.5} \right)^2 = 21.4 \rightarrow 22$$

where  $K = 2.03$  for 34 d.f.

In this case, the 36 correct matches exceeds the minimum number of 20 matches necessary for a 95 percent confidence level, allowing for a 0.5 second error in the average travel time estimate. This calculation offers a similar level of precision to the previous data set as the vehicles on the freeway were moving about twice as fast as those vehicles on the arterial for a similar distance.

Given that the variance of travel times for this section was relatively low, it is apparent that this match rate would not have been adequate with a standard deviation closer to two. However, the match rate of the Mobilizer system is related to the speed of the vehicles. As demonstrated by the Montlake arterial data, the match rate was considerably higher with the slower traffic speed, thus overcoming the higher standard deviation of the measured travel times.

**Table 16. Travel Time Statistical Comparisons**

|                           | <b>Mobilizer<br/>All<br/>Matches</b> | <b>Mobilizer<br/>Correct<br/>Matches</b> | <b>Mobilizer<br/>Matches<br/>Ground Truth<sup>a</sup></b> | <b>Total<br/>Ground<br/>Truth</b> |
|---------------------------|--------------------------------------|--|---|-----------------------------------|
| Average Travel Time (sec) | 16.19                                | 15.96                                    | 16.09   | 15.86                             |
| Standard Deviation (sec)  | 1.46                                 | 1.14                                     | 1.31  | 1.28                              |
| n (vehicles)              | 46                                   | 35                                       | 35  | 432                               |
| F-stat                    | 1.15                                 |  |   |                                   |
| F-crit (95%)              | 2.62                                 |  |   |                                   |
| p-value                   | 0.328                                |  |   |                                   |

<sup>a</sup> These numbers are based on the identification of the 35 correct Mobilizer matches from the manual ground truthing of the original video.

These results indicate that there is no evidence to indicate that any of the four average travel times are statistically significantly different from one another.

Another six minute piece of video was analyzed from the same freeway video; but this portion of video was shot using the default shutter speed of the camera, 1/60 second. These results are presented to illustrate the difference due to the shutter speed. While the traffic stream may have been slightly different, for all practical purposes, everything was the same from this piece of video as the previous one shot with a shutter speed of 1/1000 second, as the two video segments were just minutes apart on the same videotape.

During the six-minute period, 58 vehicles were matched from the traffic stream, from a population size of approximately 183 vehicles traveling through this section. This represents a sample rate of about 32 percent from this data set.

Of these 58 vehicles matched, 42 of them were matched correctly. This corresponds to a match accuracy of about 72 percent. While the sample rate was considerably higher for this portion of video, the match accuracy was considerably lower. As discussed in the report [Nihan and Washburn, 1998] and appendix B, the vehicles in this portion of video are slightly blurred due to the much slower shutter speed. This results in a loss of image detail and subsequently results in more vehicles looking similar than they otherwise would; therefore resulting in the greater percentage of matches. But

as the match accuracy shows, a greater percentage of these matches are incorrect because, in fact, many of these blurry vehicles that look similar to the system are not the same vehicles. The following tables summarize these results.

**Table 17. Freeway Summary Statistics**

| <b>Time Period (minutes)</b> | <b>Total Volume (vehicles)</b> | <b>Number of Matches</b> | <b>Percent Matches</b> | <b>Number Matches Correct</b> | <b>Percent Matches Correct</b> |
|------------------------------|--------------------------------|--------------------------|------------------------|-------------------------------|--------------------------------|
| 6                            | 183                            | 58                       | 32                     | 42                            | 72                             |

**Table 18. Freeway Travel Time Accuracy Results**

|                               | <b>All Matches</b> | <b>Correct Matches</b> |
|-------------------------------|--------------------|------------------------|
| Average Travel Time (seconds) | 15.36              | 15.74                  |
| Standard Deviation (seconds)  | 1.89               | 1.02                   |
| n (vehicles)                  | 58                 | 42                     |
| t-stat                        | 1.18               |                        |
| t-ref (95%)                   | 1.99               |                        |
| p-value                       | 0.239              |                        |

However, despite the lower accuracy rate for this analysis segment, the travel time results were still found to be statistically adequate for this data set.

#### **6.4.3 Matching Parameter Settings**

There are several parameters that can be varied by the analyst that will affect the criteria used for determining vehicle matches, and subsequently the sample rate and accuracy rate results. These parameters are proprietary information of CMS and cannot be discussed in detail in this paper. These parameters can be adjusted for each unique pair of FOV's. For the purposes of the analyses in this report, it was decided to utilize the same set of parameter settings for each video data set. Thus, the parameter settings chosen for use in these analyses were general settings that will achieve good results under a wide variety of conditions. However, it is possible to improve upon the results discussed in this paper by tailoring the parameter settings specifically for each data set.

## 6.5 Conclusions and Recommendations

The test results described in this report show very good potential for this technology to be effective in matching vehicles in successive fields-of-view. For both the arterial and freeway tests with a high shutter speed, the match accuracy rates were in excess of 75 percent. Even factoring in the incorrect matches, however, the travel times were not found to be statistically significantly different from the travel times of just the correct matches. A key reason for this is because the system can be “told” to not consider matches that fall outside of the established travel time range (this is initially set by the user, but then is adjusted in real-time by the system). The most common type of mismatch error is a result of a very similar type of vehicle (e.g., sport-utility) being in the same vicinity (one or two vehicles away) of the correct vehicle. Very similar in this case generally means the same brand, model, year, and color intensity (light or dark), as it was confirmed on numerous occasions that the system is very good at distinguishing between similar vehicles with slight differences, such as license plates in different locations, different taillight locations, and other small differences. Even the presence of an item in view through the rear window can be used to distinguish between two otherwise identical vehicles. Unfortunately (for the Mobilizer system), there are a tremendous number of identical looking sport-utility vehicles (e.g., Honda Passport) and mid-sized sedans (e.g., Honda Accord) on the roadway. Given this fact, it will be very difficult to ever achieve match accuracy rates approaching 100 percent; however, for travel time purposes, it is unlikely that this high of an accuracy rate will ever be necessary.

In these tests, the weather conditions were fairly benign. However, significant shadows were present in several of the video segments. The system showed a good ability to adjust to shadow conditions—although some data are temporarily lost during the adjustment process, the transition is fairly quick (within 10 or 15 seconds). The presence of heavy shadow conditions did not appear to adversely affect the performance of the Mobilizer system. It is possible that prolonged rapidly fluctuating shadow conditions could have an adverse effect on system performance, but this condition was

not encountered during testing; thus no conclusions regarding this can be drawn at this time.

The Mobilizer system has very good potential to be an effective tool for comparing high-occupancy lane and general-purpose lane travel times. From the tests conducted for this report, the system is currently capable of matching vehicles in adjacent FOV's with reasonable success. Since this is the key element to determining travel time, comparisons in travel time can be made between any lanes captured within the FOV's. This is actually easily facilitated since the system currently does not consider matches of vehicles that have changed lanes; therefore, the travel time results are lane dependent.

It should be mentioned however, that a key consideration for collecting travel time data for lane comparison is the FOV definition. For data collection efforts to date, three is the maximum number of lanes that have been set up for data collection within a FOV. If wanting to compare an inside HOV lane travel time to three or four adjacent GP lane travel times, it is possible that the match and accuracy rates will be less than for a FOV with three lanes or fewer due to the level of zoom and perspective needed to capture that much width of roadway.

### *6.5.1 Additional Improvements*

The Mobilizer system does not currently utilize color information in its tracking algorithms. While the results are still very promising despite the lack of color information, adding color identification information to the algorithms will undoubtedly improve the results. This will become especially apparent in situations when two otherwise identical vehicles (e.g., same model of mini-van or sport-utility vehicle) can be distinguished based on their color. The current algorithms utilize shades of gray, so light colored vehicles can be distinguished from dark colored vehicles; however, a blue vehicle cannot be distinguished from a black vehicle, nor a white vehicle from a yellow vehicle.

The Mobilizer system currently digitizes the video stream at a resolution of 320 x 200 pixels. While this still affords good resolution, provided an adequate shutter speed was used, eventually the system will be capable of digitizing at a higher resolution. The

developers have confidence that the latest generation of Pentium® III processors will allow them to digitize at twice the resolution without any loss in processing performance currently realized from the previous generation of Pentium® processors. Certainly, the greater the resolution, the more detail that can be extracted from the vehicles; thus, even better results can be obtained, with both a higher match rate and greater match accuracy.

The addition of color recognition and increased resolution to the Mobilizer system will certainly improve the ability of the Mobilizer to correctly match vehicles in successive FOV's, particularly with respect to nearly identical looking vehicles, and thus improve overall match rates and accuracy rates. The only limitation to implementation of these items is processing power. It is desired by the developers to be able to run the system on standard PC platforms to keep the system as affordable as possible. Thus, the developer will work toward implementing these improvements as the state of standard PC processing power allows them to do so while continuing to keep the system capable of running in real-time.

### *6.5.2 Areas of Further Research*

As discussed in the report [Nihan and Washburn, 1998] and appendix B, there were various factors considered during the data collection process, such as overpass locations and distance calibration capabilities, to name just a couple. However, the predictability of traffic flow conditions on the study sections was not a primary consideration for initial data collection efforts. It was desired to capture video of traffic flow conditions that ranged from free-flow to stop-and-go. It was also desired to avoid inclement weather conditions (e.g., rain) since the data collection was being done by personnel in the field for these studies. Additionally, since the data collection periods were constrained by things such as battery power, videotape length, and personnel fatigue, data collection periods of more than three hours were not feasible. As it turned out, during scheduled data collections (excepting rain-outs), traffic flow conditions were either continuous free-flow or continuous stop-and-go for the entire data collection period.

For this study it was decided to focus on analyzing the theorized easier case of free-flow traffic conditions. Obviously, if the Mobilizer system were unsuccessful for these conditions, there would be little point in attempting to analyze stop-and-go conditions. Since the results demonstrate the viability of this technology under relatively non-congested flow conditions, it is recommended to further test the system under transitory and congested-flow conditions. Preliminary indications are that the matching results will actually improve under many less-than-free-flow conditions, short of stop-and-go, as the slower speed of traffic would allow the system to get even more “looks” at a vehicle within each FOV. However, stop-and-go conditions could prove to be the biggest challenge of all to the system. Since extremely small gaps between vehicles could be difficult for the system to detect, this could possibly result in the system considering adjacent vehicles to be a single vehicle or dismissing all closely spaced vehicles as invalid match candidates, thereby reducing the potential sample size. Additionally, future plans should include testing the Mobilizer system under a more inclusive set of environmental conditions typical to this region, particularly rain.

Although not a consideration for this research, it is worth mentioning that the same key concept (matching vehicles in successive FOV's) involved in measuring travel time with this technology is also directly applicable to determining origin-destination data, which is also considered highly desirable information for transportation planning and modeling purposes. While this technology could certainly be applied to collecting origin-destination data, it should also be pointed out that the data collection setup considerations would be substantially different if one was interested in obtaining this information by way of on-ramps and off-ramps. Thus, the application of the Mobilizer system to the collection of origin-destination data is another worthy area of investigation.

Finally, a significant area that has not been addressed to date is the relative costs associated with this video image tracking technology versus other kinds of automated technologies for collecting travel time data. Certainly cost effectiveness is a key consideration for agencies considering the deployment of technologies to support ITS applications. However, a cost effectiveness comparison is premature at this point, given

that the Mobilizer system is still undergoing development and no retail cost figures will be available for some time, let alone long-term maintenance and operation costs. Nonetheless, indications are that a cost effectiveness comparison would show favorable results for a system like this because of the additional capabilities the Mobilizer system provides over other data collection technologies. Namely, the Mobilizer is capable of providing all standard traffic measurements (e.g., speed, volume, headway, etc.) in addition to travel time, whereas technologies like video license plate recognition, automatic vehicle identification (AVI), and automatic vehicle location (AVL) are more limited in the kinds of traffic stream information they can provide beyond travel time. Thus, other data collection technologies, such as inductance loop detectors, may still be necessary in conjunction with those other technologies to supply other required traffic measurements for a traffic management center. Additionally, these other technologies still need to overcome some issues related to statistically unbiased travel time estimates, real-time operation, and personal privacy issues. In the current political state, the privacy issue is far from trivial. While the Mobilizer system attempts to uniquely identify vehicles, it does not associate that vehicle with a person, as might an AVI, AVL, or license plate recognition system. Therefore, although objective comparisons will certainly be in order eventually, several other subjective factors will need to be considered as well in any technology comparisons.

## CHAPTER 7

# SAMPLING CONSIDERATIONS FOR TRAVEL TIME MEASUREMENT

---

### 7.1 Introduction

Unlike measurements such as volume and speed, travel time requires measurements of vehicles at more than one point. Additionally, unique information must be gathered from each vehicle so that it can be re-identified at the other measurement locations. If unique identifying information could be obtained from every passing vehicle at each location of interest, then extremely high vehicle match rates would be obtained (but probably less than 100% due to intermediate entrance/exit points). However, because of traffic characteristics (high volume and/or high speed) for most facilities and technology/financial limitations, obtaining high sample rates of vehicles for uniquely identifying information is not currently possible. Thus, since we cannot sample 100% of the traffic stream, two questions naturally arise: 1) 'How many vehicles need to be sampled?', and 2) 'Is any particular sampling strategy more appropriate than others?' The discussion on this second question is extended into the following chapter (Travel Time Measurement Bias Issues). Both of these questions have been extensively investigated in relation to sampling populations of people and the sort, but their discussion as it relates to sampling streams of vehicles is not well documented.

The intent behind any sampling strategy should be to obtain a sample distribution that is representative of the total population distribution. If the sampling of vehicles from the traffic stream is done in a truly random fashion, and enough samples are taken, this should be the result. However, without prior knowledge of the distribution of the vehicle types and speeds within the traffic stream of interest, it is difficult to ensure a completely random sample. Certainly, even without specific knowledge of the population distribution, the goal should be to collect an adequate size sample and to avoid a systematic bias within the sampling scheme.

The objective of this chapter is to discuss the primary considerations for sampling traffic streams for travel time measurement, and to illustrate some of these considerations through field collected travel time data. The associated issue of sampling and/or measurement induced bias will also be addressed in the following chapter.

Before discussing specific sampling issues, it is appropriate to first discuss some of the basic principles of traffic flow characteristics as they affect sampling.

## **7.2 Properties of Traffic Flow that Affect Point Sampling and Statistical Analysis of Travel Times**

This section will discuss the relevant properties of traffic flow that influence point sampling capabilities and review the properties that affect statistical analysis of travel times.

### *7.2.1 Traffic Flow Properties that Affect Point Sampling Capabilities*

The sampling of vehicles from traffic streams is highly dependent not just not on the technology/method used, but also on the traffic flow itself. Since each technology generally has an inherent maximum sampling rate, the arrival rate of traffic flow past a data collection point will strongly influence the data obtained. For example, highly platooned flow regulated by a traffic signal will result in different sampling characteristics than the same total amount of flow spread uniformly throughout the same time period. Signals are the main contributor to vehicle platooning. The sampling characteristics of a traffic stream dominated by platoon flow are obviously going to be considerably different than those under steady flow conditions, for any given total volume.

For manual sampling techniques, the greatest sample sizes will usually be obtained from fairly uniform flow at low speeds. Obviously, stop-and-go type traffic will allow a sampling of nearly the entire population. As traffic speed increases, and as headways decrease, the sampling rate will inherently decrease with a manual sampling method. The additional difficulty with highly platooned flow is that groups of higher speed vehicles, with small headways pass a data collection point, and then there are gaps of time in which the flow between platoons is negligible. Thus, it is difficult to get a

large sample proportion from the platooned flow, and then there is almost no sampling done in-between platoons.

Very small headways, regardless of speed, have an adverse effect on a technology like video imaging analysis. Due to camera angle, closely spaced vehicles can appear as a single vehicle to a video imaging processing system. Additionally, due to processing requirements, sample rates will usually decline with increasing vehicle speeds. Sample rates for technologies such as AVI are more immune to variations in traffic flow. They are related to the traffic flow in the sense that the probe vehicles equipped with these devices are part of the traffic flow. The sample rate for this type of technology is more directly related to the density of probe vehicles dispersed in the traffic stream.

### *7.2.2 Review of Statistical Analysis of Travel Times*

A variable such as travel time will conform to a particular distributional form like any other variable. Whatever the distributional form, it will have a central tendency measure (mean, median, mode), and a dispersion measure (variance, standard deviation). In performing statistical testing of a measurement, such as hypothesis testing or calculating confidence intervals, it is necessary to know what the distributional form is. Past literature has assumed that travel times follow a normal distribution, or at least approximately normal (such as the t-distribution).

Breiman et al. [1977] concluded that both time mean speeds and space mean speeds are approximately normally distributed, despite that the fact that speeds are not independent over all flow conditions. The relationship between time mean speed and space mean speed is illustrated below.

#### Time Mean Speed vs. Space Mean Speed

Time mean (point) speed is the arithmetic mean of speeds measured at a specific point on the roadway over a specific time interval.

$$\bar{\mu}_t = \frac{\sum_{i=1}^n \mu_i}{n}$$

where:  $\mu_i$  = observed speed of  $i^{\text{th}}$  vehicle

$n$  = number of vehicles observed

Space mean speed is the arithmetic mean of speeds of vehicles occupying a section of roadway of specific length at a given instant. It is the average of vehicle speeds weighted according to how much time they spend traveling the section of roadway.

$$\bar{\mu}_s = \frac{nd}{\sum_{i=1}^n t_i}$$

where:  $d$  = length of roadway section

$t_i$  = observed time for  $i^{\text{th}}$  vehicle to travel distance  $d$

The relationship between time mean speed and space mean speed is given by the following formula.

$$\bar{\mu}_t = \bar{\mu}_s + \frac{\sigma_s^2}{\bar{\mu}_s}$$

where:  $\sigma_s^2$  = variance of the space distribution of speeds

While time mean speeds and space mean speeds are not directly related, space mean speed and travel time are. However, Berry [1952] found that while average travel speeds can be normally distributed, the resulting travel time distribution based on those same speeds may be skewed from normality under certain traffic conditions. It was generally observed that travel time distributions were skewed toward longer travel times under free-flow conditions. Berry recommended examining the distribution of both travel speeds and travel times and basing sample size statistics on the distribution that is more nearly normal. Berry evaluated distribution normality by use of a skewness index.

While it is commonly accepted that travel time is approximately normally distributed, the theoretical justification is still limited. The research by Breiman et al. is now over 20 years old. The research by Berry is over 45 years old. It is probable that travel times follow a normal distribution for many flow conditions, but not all. Full

investigation of this issue is beyond the scope of this work and is suggested for further research, but a short example will be given later in section 7.4.2.

While in practice, the distinction between time mean speed and space mean speed is often of minor importance, the distinction as it relates to the measurement of travel time is of greater significance. Algorithms that use inductance loop detectors to estimate speed and then convert to travel time are based upon point speeds. Although the use of dual loop stations (i.e., “speed traps”) are technically a space mean speed, since the speed is based on the time elapsed between the two loops, they are still essentially point speeds because the distance between the loops is very short. The main issue, though, as it relates to using loop detector estimated (or directly measured) speeds to estimate travel time is the significant distance between loop detector stations, usually on the order of ½ mile. During peak conditions, traffic conditions can be highly variable between loop detector stations, and this will not necessarily be reflected in the loop measurements.

Equations discussed in the literature as being applicable to determining sample sizes for travel time measurement are now reviewed.

#### Applicable Sample Size Formulas

There are several formulas that have been described in the literature as applicable to determining the necessary number of matches for travel time studies. These will be reviewed here, with a brief discussion of the advantages and disadvantages of each one.

#### ***Standard Deviation Formulation***

$$n = \left( \frac{t_{\alpha, n-1} \times s}{\epsilon} \right)^2$$

where: t = t-statistic from Student’s t distribution for specified confidence level and degrees of freedom;  
 s = standard deviation of travel time; and  
 ε = maximum specified allowable error.

T-statistic: t - value from the Student’s t distribution for (n-1) degrees of freedom. The t-statistic is based upon the specified confidence level (two-tailed test) in the

travel time estimate. Because the degrees of freedom for the t-statistic rely on a sample size,  $n$ , an initial sample size estimate must be assumed. Iterative calculations should be used to provide better estimates for the degrees of freedom.

If sample sizes are large enough, a z-statistic from the normal distribution may be substituted for the t-statistic. While the typical threshold for choosing a z-statistic over a t-statistic is a sample size of 30 or more, there could still be significant differences in your sample size estimate until your number of samples reaches over 100.

Allowable error:  $\varepsilon$  - the maximum permissible error in the travel time estimate, expressed as a percentage (%). The maximum error is specified by the study designer and will depend upon the uses of the travel time data.

The advantage to this formulation is that it is intuitive and simple to use. This formulation is widely regarded as providing the most reliable estimates of sample size. The disadvantage of course is that the standard deviation is usually not known with much degree of accuracy before the study begins. Thus, if the pre-study estimate of variance is inaccurate, a large enough sample for statistical validity may not be obtained, or a much larger sample than needed may be obtained, which is a waste of resources. Additionally, the value of 't' depends on the number of degrees of freedom, which is a function of the number of samples. Thus, an iterative procedure is required.

### ***Coefficient of Variation Formulation***

$$n = \left( \frac{t_{\alpha, n-1} \times \text{c.v.}}{e} \right)^2$$

$$\text{Coefficient of Variation, c.v.} = \frac{s}{\bar{x}}$$

$$\text{Relative Error, } e = \frac{\varepsilon}{\bar{x}} \quad \text{where: } \bar{x} = \text{mean travel time}$$

Coefficient of variation: c.v. - the relative variability in the travel times, expressed as a percentage (%). The c.v. values can be calculated from empirical data using Equation Z.

Relative allowable error:  $e$  - this is the maximum allowable error divided by the mean travel time, again expressed as a percentage (%).

This formulation is simply an extension of the standard deviation formulation, and is mathematically equivalent. This formulation may be favored if it is desirable to explicitly display the coefficient of variation in the calculation.

### ***Finite Population Correction Factor***

If the sample size actually comprises a large percentage of the population, the following equation will take that into account and adjust the necessary sample size downwards.

$$n = \frac{n'}{1 + \frac{n'}{N}}$$

where:  $n'$  = uncorrected sample size, and  
 $N$  = population size.

### ***Sample Range Formulation***

$$n = \left[ \frac{Z_{\alpha} \bar{R}}{d\varepsilon} \right]^2$$

where:  $Z_{\alpha}$  = normal, two-tailed statistic for a confidence of  $1 - \alpha$ ,  
 $d$  = ratio of  $\bar{R}$  to  $\sigma$ ,  
 $\varepsilon$  = user-specified allowable error or interval half-length,  
 $n$  = sample size

$$\bar{R} = \max_{i=1}^n v_i - \min_{i=1}^n v_i$$

where:  $v_i$  = running speed associated with travel time run  $i$ .

A recent paper by Quiroga and Bullock [ITE, 1998] compared sample size estimates made with this formulation (they also replaced the Z-statistic with the t-statistic) against the standard deviation formulation. They found the two estimates to be very comparable. However, this formulation is better suited to probe vehicle methods because the calculation of  $\bar{R}$  calls for running speeds, which cannot be determined from point measurement methods such as license plate matching that do not measure delay. Although the use of a statistic like the range of running speeds may be an intuitive one,

this equation still incorporates the standard deviation measure; thus, this equation is not any easier to use than the standard deviation formulation.

### ***Recorded Volumes Formulation***

$$N = \frac{pqK^2}{E^2}$$

where:  $N$  = minimum number of correctly recorded inbound plates,  
 $p$  = proportion of correctly recorded inbound plates that will be recorded as a match at an outbound station,  
 $q$  = proportion of correctly recorded inbound plates that will not be recorded as a match at an outbound station (either not a through vehicle, or missed or incorrectly recorded at the outbound station),  
 $K$  = constant corresponding to the desired confidence level (two-sided test), and  
 $E$  = permitted error in the proportion estimate of observance.

This approach is discussed by Box and Oppenlander [ITE, 1976] and Schaefer [ITE, 1988]. This approach is not recommended for the main reason that it does not deal explicitly with travel time or travel time variation. This equation calculates the number of license plates to record to determine the proportion of vehicles traveling through a study section (past both an upstream and downstream recording station). Additionally, estimating the proportion of correctly recorded vehicle plates at each recording station that will result in a match is not necessarily any easier than estimating the variance of travel time. There is no theoretical relationship between the proportion of vehicles recorded and the variability of the travel times. Furthermore, formulations that consider travel time variability are more intuitive to use.

### **7.3 Data Description**

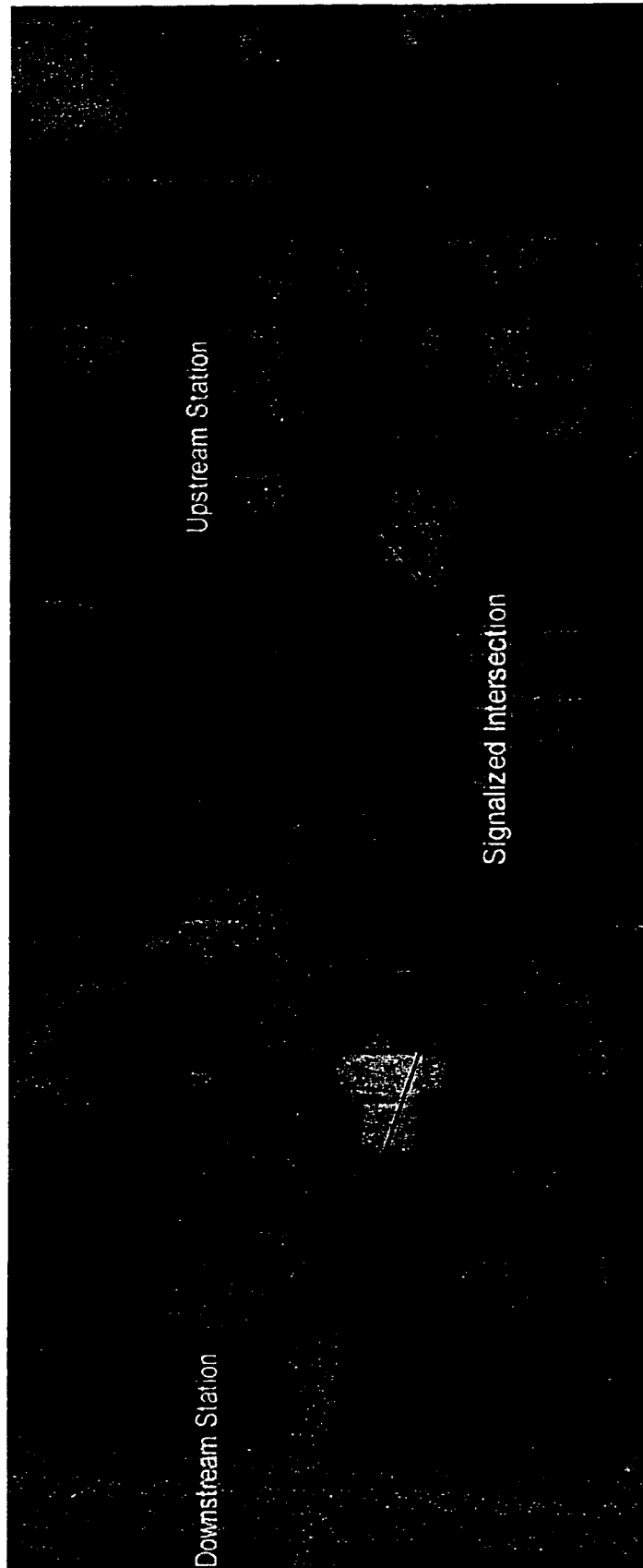
Travel time data were collected manually on both a freeway section and an arterial section. These data are described in the immediately following sections. Throughout the rest of this document, these data will be used in examples or to illustrate various issues, relating to sampling and bias, as appropriate.

### 7.3.1. Arterial Data

Data were collected simultaneously at 15<sup>th</sup> and 7<sup>th</sup> avenues of NE 45<sup>th</sup> Street, a major arterial in the University district. This data collection effort consisted of placing two data collection personnel at both the origin and destination location. At each location, one data collector entered license plates using the portable computer keyboard, while the other used voice recognition. License plates were collected from just the outside (curb) lane for through vehicles at the origin station. All vehicles turning left and right onto NE 45<sup>th</sup> street from 15<sup>th</sup> were also collected. At the destination station, license plates were collected from just the outside lane. Figure 16 shows an aerial photo of the arterial study section. The total study section length from 15<sup>th</sup> Avenue (Upstream) to the 7<sup>th</sup> Avenue (Downstream) is 2,150 feet. There are 7 signalized intersections along this section, but signal delay from the 15<sup>th</sup> Avenue signal was not included in the travel time measurements.

Data were collected on two different days, separated by about six weeks. During both days, data collection was conducted during the hour of 4:00 to 5:00, an hour of the PM peak period for this facility. In addition, reference data were simultaneously obtained by video recording all passing vehicles/license plates (in just the curb lane) at the origin and destination locations. These video tapes were then manually reviewed and the license plate data were entered into a software database. Keyboard and voice entries were compared to the video data for results calculations. After considering the various site characteristics, the video cameras were placed on the near sidewalk, downstream of the time-stamp reference locations. The main problem with this type of placement was FOV interference by pedestrians. At the destination location, pedestrian traffic was lighter, so FOV interference was infrequent; but at the origin, it was much more common. Furthermore, in order to get the curb lane through vehicle license plates adequately sized and placed in the camera FOV, it was not possible to capture the license plate of every right and left turning vehicle. As a result, there were some vehicles that were part of the match pool for the data collected by the human observers that were not contained in the video match pool.

**Figure 16. Aerial Photo of 45th Data Collection Site**



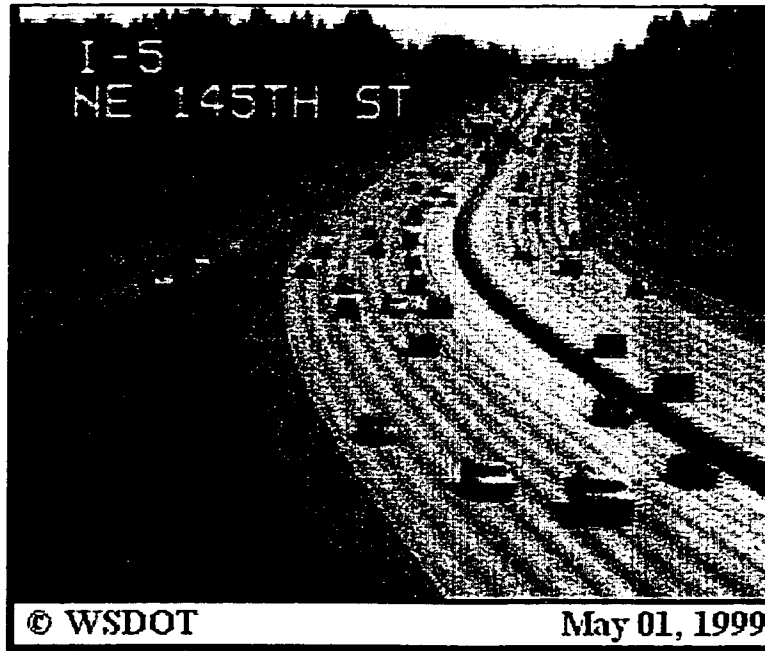
### Sampling Methodology

For these data collection sessions, the data collectors tried to sample as many vehicles as possible, while time-stamping the entries as close as possible to a designated reference point. The data collectors were instructed to ignore only bus and motorcycle vehicle types. Buses make frequent stops along this corridor, and their plate numbering scheme is often not unique. Motorcycles have very small plates and are only mounted on the rear of the cycle. The data collectors were also instructed on the plate entry scheme, which was as follows: For centennial passenger car plates, enter the first three numbers and first letter; for older style passenger car plates, enter the first letter and the last three numbers; for all other plates enter the first four characters.

Each day, two separate data collection sessions were conducted during the hour. This was done for the purpose of testing two different time-stamping techniques for the voice recognition method. The data collectors at the downstream station started their data collection two minutes after the upstream data collectors to allow for the travel time of the first group of vehicles recorded at the upstream end.

#### *7.3.2 Freeway Data*

License plate data were collected from two adjacent interchange locations on Interstate-5 in north Seattle, on Tuesday, June 4, 1996 from 3:45 PM to 6:30 PM. The data were obtained by videotaping the license plates of passing freeway vehicles. High-speed 8-mm (Hi-8) video cameras were set up on the NE 130<sup>th</sup> Street and NE 145<sup>th</sup> Street arterial overpasses, overlooking the northbound direction of travel on I-5. One camera was positioned over the high-occupancy vehicle lane and another camera was positioned over the inside general-purpose lane (adjacent to the HOV lane), for both overpasses. Figure 17 shows pictures of the data collection sites.

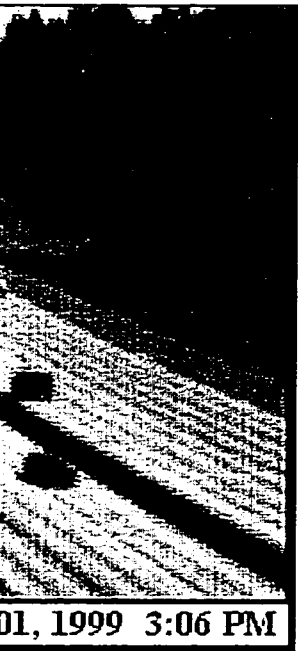


Looking South Towards NE 130<sup>th</sup>



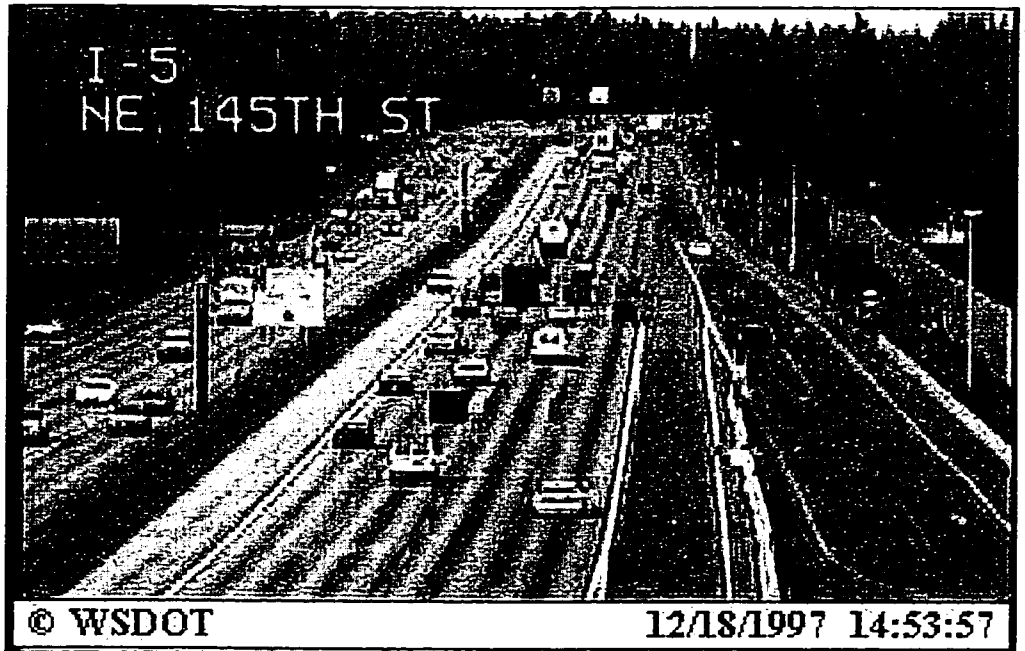
Figure 17. Schematic of I-5 Data





01, 1999 3:06 PM

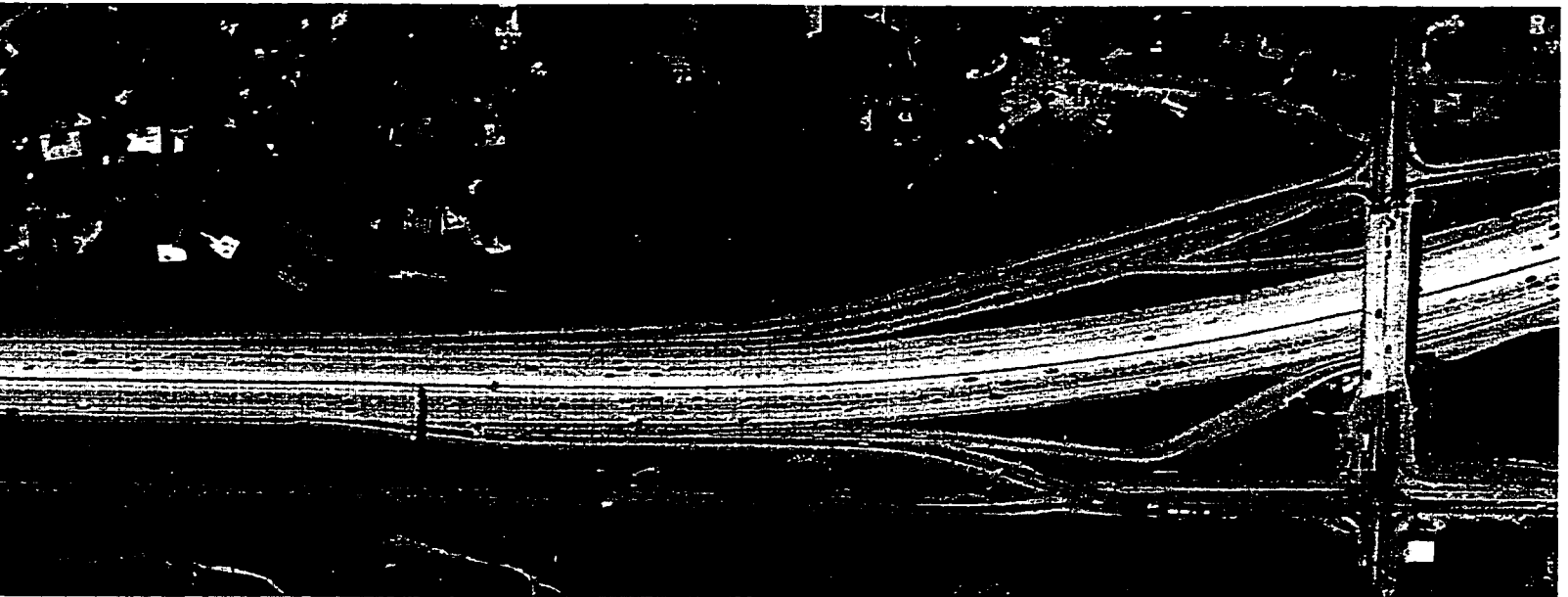
0<sup>th</sup>



© WSDOT

12/18/1997 14:53:57

Looking North



-5 Data Collection Site (North to the right)



The tapes were manually reviewed and the license plate for each passing vehicle for all four cameras was entered into a database, along with the corresponding time-stamps. These data were entered for the time period of about 4:00 PM to 5:30 PM, capturing the transition from free-flow to stop-and-go to back to free-flow during the peak period. Additionally, one of several vehicle types (classifications) was recorded for each observed vehicle.

## **7.4 Travel Time Results**

The basic statistics for both sets of data are presented in the following subsections.

### *7.4.1 Arterial Data*

License plates were matched for all three data sets (i.e., video, keyboard entry, and voice entry). Table 19 shows the results of the matching process for all three data sets, as well as the volume statistics, for the first two data collection sessions. Table 20 displays the travel time statistics for the third data collection session (first session of second data collection day).

Table 19. NE 45<sup>th</sup> Travel Time Statistics, 1<sup>st</sup> and 2<sup>nd</sup> Session

|                                     | Video                  | Keyboard               | Voice Rec              | Video           | Keyboard        | Voice Rec <sup>b</sup> |
|-------------------------------------|------------------------|------------------------|------------------------|-----------------|-----------------|------------------------|
| Total Usable Volume <sup>a</sup>    | 186 / 238              | 186 / 238              | 186 / 238              | 162 / 226       | 162 / 226       | 162 / 226              |
| Recorded Volume                     | 146 / 231 <sup>b</sup> | 162 / 188 <sup>c</sup> | 155 / 217 <sup>c</sup> | 128 / 210       | 144 / 179       | 102 / 156              |
| Sample Rate (%)                     | 78.5 / 97.1            | 87.1 / 79.0            | 83.3 / 91.2            | 79.0 / 92.9     | 88.9 / 79.2     | 63.0 / 69.0            |
| Number of Matches                   | 56                     | 41 <sup>e</sup>        | 44 <sup>f</sup>        | 39 <sup>f</sup> | 36 <sup>e</sup> | 21 <sup>e</sup>        |
| Match Percentage <sup>d</sup>       | 38.4                   | 25.3                   | 28.4                   | 30.5            | 25.0            | 20.6                   |
| Mean Travel Time (sec)              | 179.6                  | 181.7                  | 177.2                  | 214.4           | 285.4           | 287.8                  |
| Std. Deviation of Travel Time (sec) | 43.7                   | 42.8                   | 41.0                   | 65.7            | 66.7            | 55.7                   |
| Maximum Travel Time (sec)           | 264                    | 257                    | 257                    | 429             | 436             | 408                    |
| Minimum Travel Time (sec)           | 107                    | 108                    | 109                    | 145             | 158             | 190                    |
| Median Travel Time (sec)            | 182                    | 182                    | 178.5                  | 267             | 277             | 278                    |
| Mode of Travel Time (sec)           | 179                    | 179                    | 177                    | 309             | 203             | 258                    |

<sup>a</sup> This is the total volume minus buses and vehicles with no front plate or an unreadable plate (by either video or human). The first number is upstream value, second number is downstream value. The total volume was 202/265 for session 1, and 184/257 for session 2.

<sup>b</sup> Vehicles with plates that were not within the camera field-of-view and vehicles with unreadable plates from the video (e.g., glare, blocked by another vehicle or pedestrian) were subtracted from usable volume in footnote a.

<sup>c</sup> This is just the total number of license plates entered into the software program.

<sup>d</sup> This is based on the ratio of the number of matches to the lower of the recorded volumes at either the upstream or downstream station.

<sup>e</sup> Three matches were identified as outliers and removed.

<sup>f</sup> Two matches were identified as outliers and removed.

<sup>g</sup> This data collection was done using a manual post license plate entry timestamp technique.

Table 20. NE 45<sup>th</sup> Travel Time Statistics, 3<sup>rd</sup> Session

|                                     | Video                  | Keyboard               | Voice Rec              |
|-------------------------------------|------------------------|------------------------|------------------------|
| Total Usable Volume <sup>a</sup>    | 122 / 198              | 122 / 198              | 122 / 198              |
| Recorded Volume                     | 118 / 190 <sup>b</sup> | 120 / 135 <sup>c</sup> | 109 / 186 <sup>c</sup> |
| Sample Rate (%)                     | 96.7 / 96.0            | 98.4 / 68.2            | 89.3 / 93.9            |
| Number of Matches                   | 23                     | 18                     | 21                     |
| Match Percentage <sup>d</sup>       | 17.8                   | 15.8                   | 19.3                   |
| Mean Travel Time (sec)              | 432.6                  | 437.9                  | 436.9                  |
| Std. Deviation of Travel Time (sec) | 116.3                  | 126.2                  | 138.5                  |
| Maximum Travel Time (sec)           | 644                    | 641                    | 687                    |
| Minimum Travel Time (sec)           | 268                    | 267                    | 245                    |
| Median Travel Time (sec)            | 411                    | 455                    | 413                    |
| Mode of Travel Time (sec)           | 271                    | 530                    | 643                    |

- <sup>a</sup> This is the total volume minus buses and vehicles with no front plate or an unreadable plate (by either video or human). The first number is upstream value, second number is downstream value. The total volume was 149/220 for session 1.
- <sup>b</sup> Vehicles with plates that were not within the camera field-of-view and vehicles with unreadable plates from the video (e.g., glare, blocked by another vehicle or pedestrian) were subtracted from usable volume in footnote a.
- <sup>c</sup> This is just the total number of license plates entered into the software program.
- <sup>d</sup> This is based on the ratio of the number of matches to the lower of the recorded volumes at either the upstream or downstream station.

### Travel Time Comparisons

An analysis of variance for the three methods (session 1) yields the results displayed in Table 21. Comparing the calculated F-value of 0.118 to the F-distribution reference value of 3.06 indicates that there is no statistically significant difference between the travel time values obtained by each of the three methods, at the 95% level of confidence. This result was expected, just by inspection of the large standard deviations relative to the means.

Table 21. ANOVA Results for Session 1

#### Summary

| <i>Groups</i> | <i>Count</i> | <i>Sum</i> | <i>Average</i> | <i>Variance</i> |
|---------------|--------------|------------|----------------|-----------------|
| Video1        | 56           | 10060      | 179.64         | 1909.18         |
| Keyboard1a    | 41           | 7450       | 181.71         | 1831.21         |

|            |    |      |        |         |
|------------|----|------|--------|---------|
| Voice Rec1 | 44 | 7798 | 177.23 | 1683.02 |
|------------|----|------|--------|---------|

**ANOVA**

| <i>Source of Variation</i> | <i>SS</i>  | <i>df</i> | <i>MS</i> | <i>F</i>     | <i>P-value</i> | <i>F crit</i> |
|----------------------------|------------|-----------|-----------|--------------|----------------|---------------|
| Between Groups             | 428.162    | 2         | 214.081   | <b>0.118</b> | 0.889          | <b>3.062</b>  |
| Within Groups              | 250623.072 | 138       | 1816.109  |              |                |               |
| Total                      | 251051.234 | 140       |           |              |                |               |

The same conclusion can be drawn about the travel time measurements for session 2 as session 1, as indicated by the F-values in Table 22.

**Table 22. ANOVA Results for Session 2**

**Summary**

| <i>Groups</i> | <i>Count</i> | <i>Sum</i> | <i>Average</i> | <i>Variance</i> |
|---------------|--------------|------------|----------------|-----------------|
| Video2        | 39           | 10701      | 274.38         | 4318.30         |
| Keyboard2     | 36           | 10273      | 285.36         | 4446.01         |
| Voice Rec2    | 21           | 6044       | 287.81         | 3106.56         |

**ANOVA**

| <i>Source of Variation</i> | <i>SS</i>  | <i>df</i> | <i>MS</i> | <i>F</i>     | <i>P-value</i> | <i>F crit</i> |
|----------------------------|------------|-----------|-----------|--------------|----------------|---------------|
| Between Groups             | 3346.851   | 2         | 1673.425  | <b>0.408</b> | 0.666          | <b>3.094</b>  |
| Within Groups              | 381836.774 | 93        | 4105.772  |              |                |               |
| Total                      | 385183.625 | 95        |           |              |                |               |

An analysis of variance (95% level of confidence) for session 3 gives a calculated F-value of 0.010. When compared to the F-distribution reference value of 3.153, no statistically significant difference between the travel time values obtained by each of the three methods is again indicated.

**Sample Size Adequacy**

Solving the standard deviation sample size equation for the error term based upon the given sample size for the video matches and a 95% confidence interval yields:

$$\varepsilon = \frac{s * t}{\sqrt{n}} = \frac{43.7 * 2.00}{\sqrt{56}} = 11.7 \text{ seconds.}$$

where: t = 2.00 for 55 d.f.

With the coefficient of variation approach,  $c.v. = 43.7/179.6 = 0.243$ , yielding:

$$e = \frac{s * t}{\sqrt{n}} = \frac{0.243 * 2.00}{\sqrt{56}} = 0.065, \quad \varepsilon = 0.065 \times 179.6 = 11.7 \text{ seconds.}$$

This equals a potential  $\pm 6.5\%$  travel time error. To achieve a 5.0% (9 seconds) error, assuming the same standard deviation and confidence level, 95 samples would be required. These results are comparable with other reported results summarized in table 4-3 of the 'Travel Time Data Collection Handbook' [TTI, 1998]. For this data set, another 39 travel time samples would have needed to be collected to obtain the 5.0% error and 95% confidence level. At the rate that samples were obtained for this video data set, approximately one every 26 seconds, license plates would have needed to be collected for another 16.9 minutes to obtain the additional 39 samples. Of course, if data had been collected for another 17 minutes, the standard deviation could have been significantly different.

The second data collection session was started 12 minutes after the first one ended, and as can be seen from Table 19, the traffic conditions changed considerably over the next 25 minutes, resulting in a mean travel time 96 seconds longer and a standard deviation 22.3 seconds greater. Additionally, the number of matches for this session was considerably lower, with only 38 being obtained for another 25 minute period. Although the volume was a little lower for this period, which is expected given the increased travel time, the sample rate was similar to the previous session, except for the voice recognition method, which used a different time-stamp technique.

$$\varepsilon = \frac{s * t}{\sqrt{n}} = \frac{65.7 * 2.03}{\sqrt{39}} = 21.4 \text{ seconds}$$

where:  $t = 2.03$  for 38 d.f.

$$e = 21.4/274.4 = 7.8\%$$

From this information, it is clear that at least one intermediate data collection point between the two endpoints would have been necessary to achieve an appreciably lower error rate. For just the two data collection locations, even capturing all possible matches would not yield an error value of 5% or less, as evidenced by the video data. Of

course with every data collection point added, another source of potential selection bias and measurement error gets introduced.

The results for the second data collection day are dramatically different than those for the first data collection day. The average travel time was just over 7 minutes, with the standard deviation on the order of 2 minutes. These very congested conditions resulted in a lower traffic volume traversing the section than during the first data collection, which in turn led to a lower number of matches. Nonetheless, the computed error rate for this session was similar to that for the first session.

$$\varepsilon = \frac{s * t}{\sqrt{n}} = \frac{116.3 * 2.074}{\sqrt{23}} = 50.3 \text{ seconds}$$

where:  $t = 2.074$  for 22 d.f.

$$e = 50.3/432.6 = 11.6\%$$

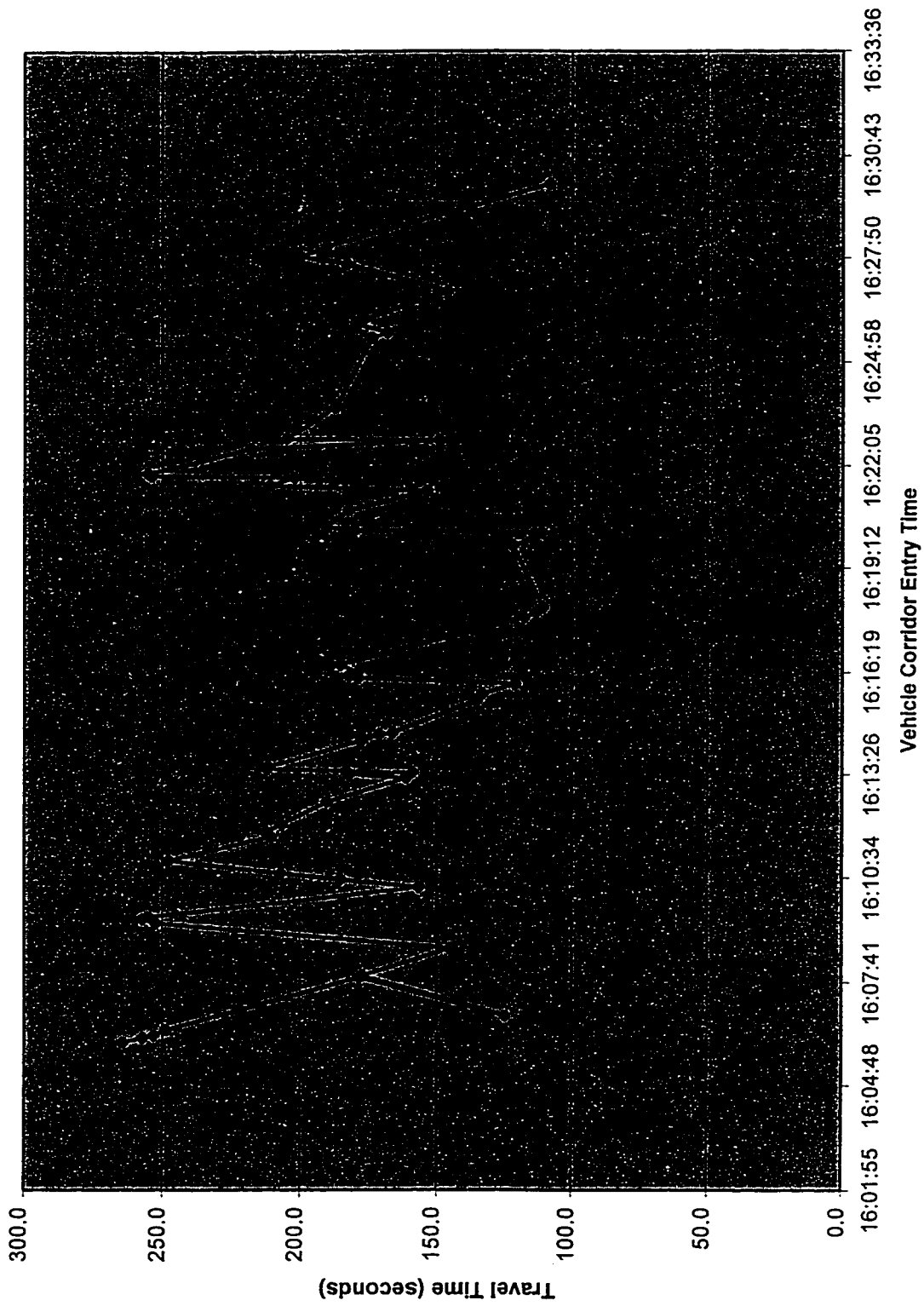
### Travel Time Distributions

Figure 18 shows a plot of vehicle matches (and their corresponding travel time) versus the time of vehicle entry, for all travel time matches during the first data collection session. What this plot illustrates very clearly is the cyclical pattern that the travel times follow, analogous to a sine wave. This results from the interrupted nature of the flow, and correlates with level of progression experienced by the matched vehicles. As the figure indicates, many vehicles experience very poor progression, and many others experience very good progression. From a sampling perspective, this is certainly more challenging than the situation of uninterrupted flow, where the travel time changes steadily and consistently.

For arterial data, it is likely that the travel time distribution will often be bi-modal or even tri-modal. That is, there will be a peak associated with lower travel times for those vehicles that experience good progression, and a peak associated with higher travel times for those vehicles that experience poor progression. With the tri-modal distribution, there would be another mode associated with the average travel time. Figure 19 shows the travel time distribution for the first data collection session on NE 45<sup>th</sup> Street.

This distribution clearly indicates three different modes, one associated with very low travel times, one for very high travel times, and one for the mean travel time. Although by statistical definition there will only be one mode for a distribution, it is clear that there are two other “secondary” modes for this data set. This does highlight one of the main difficulties with probe vehicle sampling. If a limited number of probe vehicle sample runs are made (which is often the case), it could be the case that a number of runs experience poor progression and large signal delay. Or other times, a majority of runs might get “lucky” with the progression and incur little signal delay. This issue is not trivial to deal with, because with a limited number of runs, it is difficult to say whether good progression or poor progression is truly the average condition.

However, interrupted-flow arterial data may also resemble a reasonably normal distribution also, as evidenced by the second data collection session. Figure 20 again shows the cyclical pattern of travel times, but this set of data contain smaller percentages of very low and very high travel times. Figure 21 looks much more similar to a normal distribution, with only one obvious mode near the mean travel time.



**Figure 18. Vehicle Matches versus Data Collection Time (Session 1)**

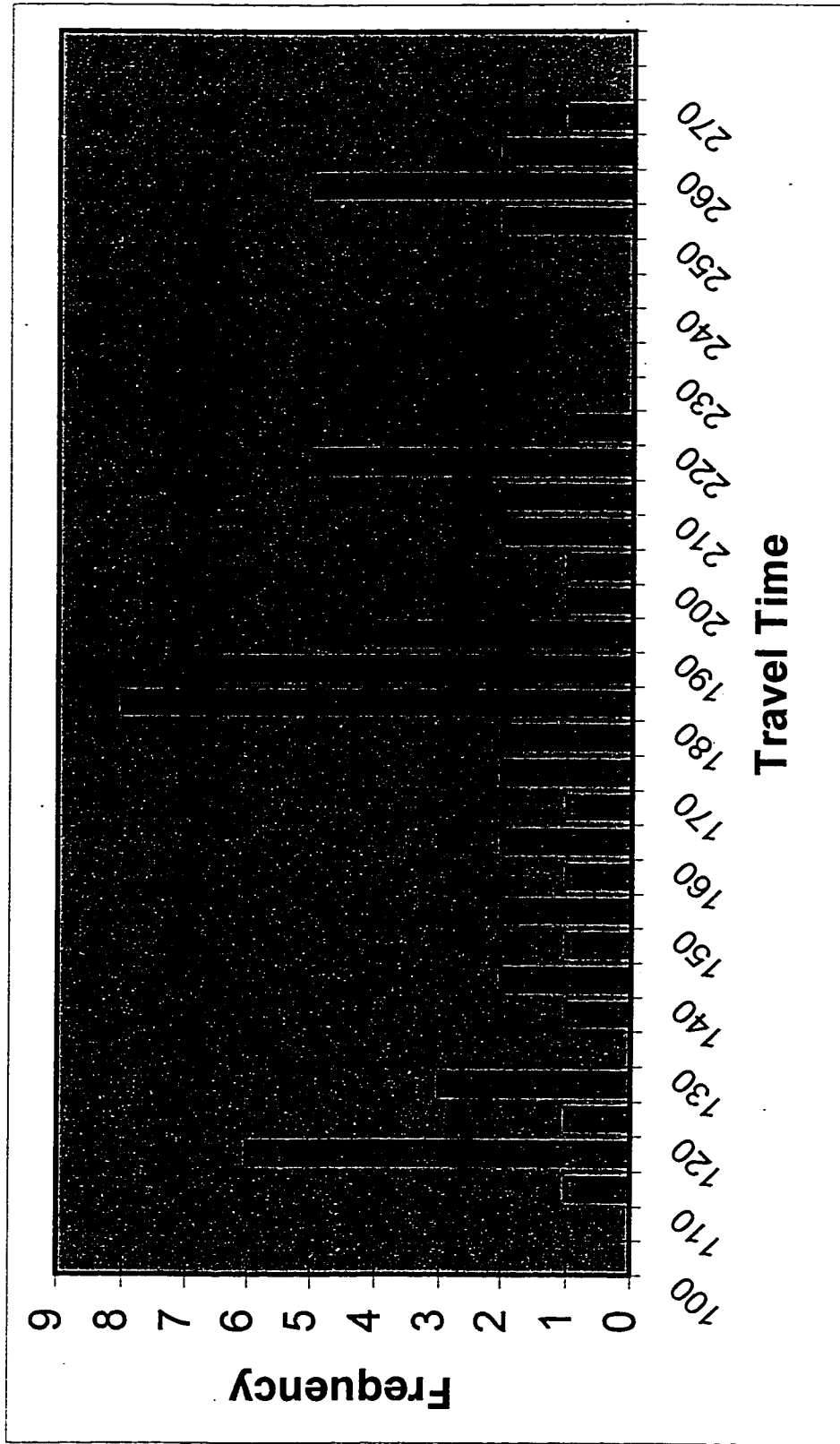


Figure 19. Arterial Travel Time Distribution (Session 1)

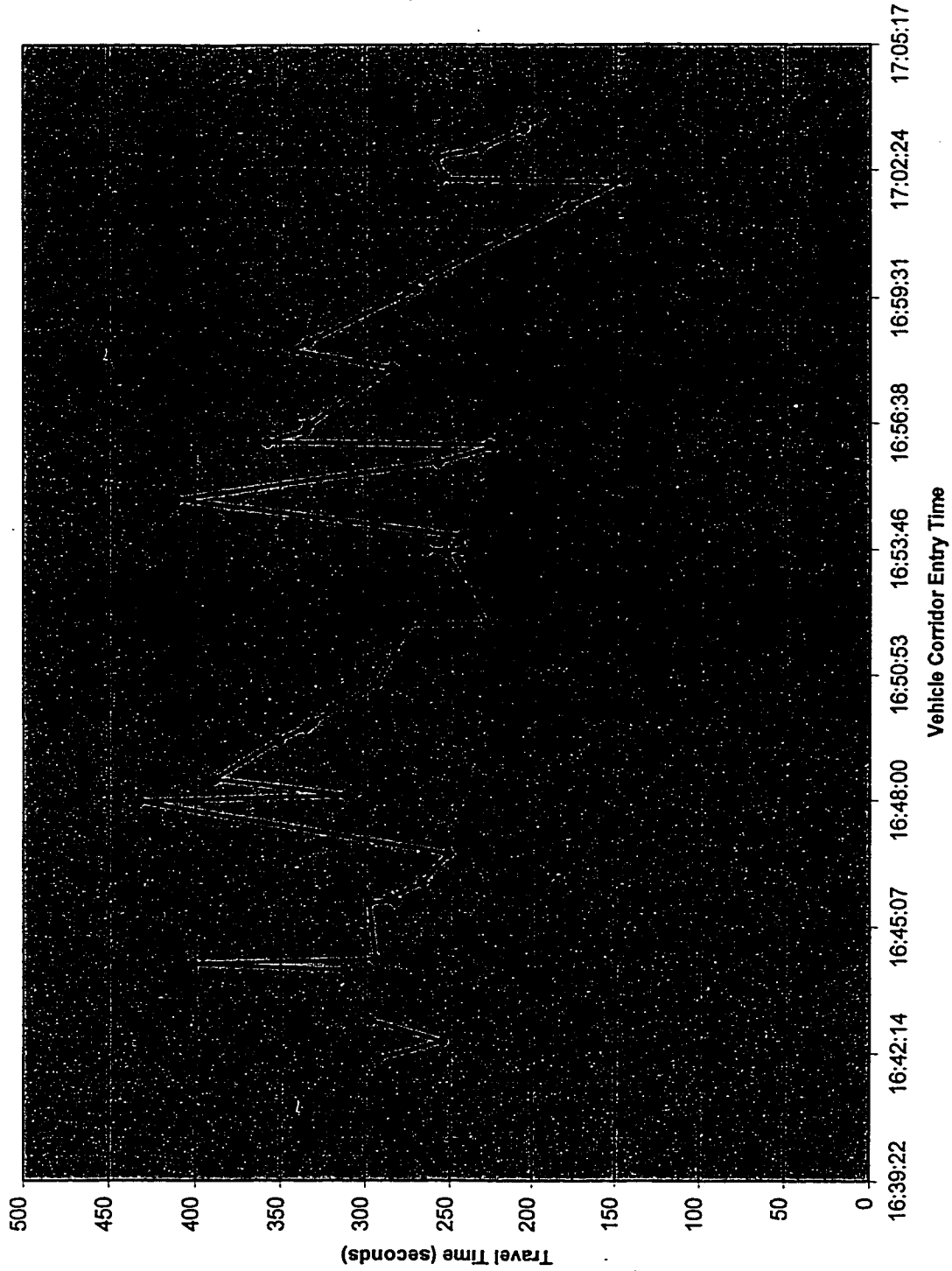


Figure 20. Vehicle Matches versus Data Collection Time (Session 2)

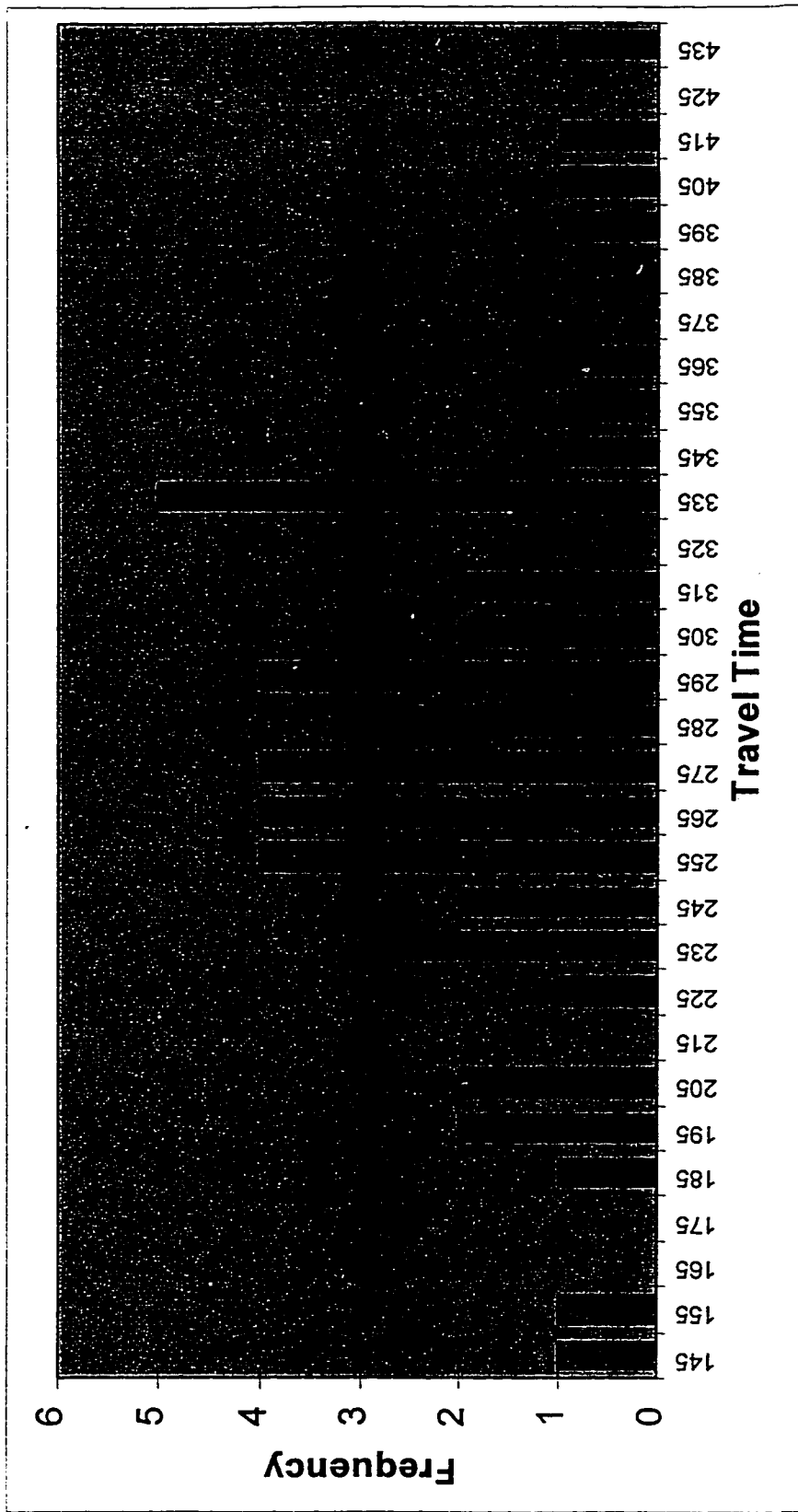


Figure 21. Arterial Travel Time Distribution (Session 2)

### Travel Time Daily Variation

Although much discussion has been made about sampling within any particular time period, it is equally important to understand the necessity of sampling across multiple time periods. While it is important to collect a statistically reliable sample during any particular time period, those results only give you information about that particular time period. Thus, sampling must be performed not only within days, but also across days. It is very possible that variation between days can be greater than the variation within a particular day (i.e., collection time period). Unfortunately, it is difficult to give simple guidelines on how many time periods to sample, and exactly when they should be collected (see Chapter 9 for additional information) due to the numerous factors that affect daily variation. Table 23 does demonstrate the variation possible between different data collection days and different time periods for the same facility.

**Table 23. Travel Time Comparisons for NE 45<sup>th</sup> Street**

|                                      | 7/23 (1)    | 7/23 (2)    | 9/3 (1)     | 1/12 <sup>a</sup> |
|--------------------------------------|-------------|-------------|-------------|-------------------|
| <b>Time of Day</b>                   | 4:00 – 4:25 | 4:35 – 5:00 | 4:46 – 5:09 | 4:00 – 5:00       |
| <b>Time period</b>                   | 25 min      | 25 min      | 23 min      | 60 min            |
| <b>Total Volume</b>                  | 202 / 265   | 184 / 257   | 149 / 220   | NA                |
| <b>Usable Volume</b>                 | 186 / 238   | 162 / 226   | 122 / 198   | NA                |
| <b>Number of Matches</b>             | 56          | 38          | 21          | 11                |
| <b>Mean Travel Time (sec)</b>        | 179.6       | 275.8       | 425.6       | 156.0             |
| <b>Median Travel Time (sec)</b>      | 182         | 267         | 407         | 150               |
| <b>Std. Dev. of Trav. Time (sec)</b> | 43.7        | 66.0        | 118.3       | 62.5              |
| <b>Max. Travel Time (sec)</b>        | 264         | 429         | 644         | 302               |
| <b>Min. Travel Time (sec)</b>        | 107         | 145         | 268         | 69                |

<sup>a</sup> Data were collected through floating car runs on this day. Two vehicles were used. The runs were made in the inside lane. The other data were collected through license plate matching and from the outside lane.

From the data collected at NE 45<sup>th</sup> Street, on two different days, almost six weeks apart, it is obvious that the issue of time-of-year timing for data collection is not trivial. As the results show for 25 minute periods within a similar time frame (4 – 5 PM), the results are dramatically different—one average travel time is 150 seconds higher than the other, and the standard deviation is almost twice as much. Granted, the statistical

confidence level in both cases is not very high, but this difference is significant enough that there are clearly some traffic factors that are contributing to this difference.

#### 7.4.2 Freeway Data

Due to the very large number of license plates in these data sets, it was necessary to develop an application that would automatically match the license plates. This application is described briefly in Appendix C.

Table 24 displays the summary of the freeway license plate matching volumes. There are two general purpose (GP) lane segments because there was a short power outage shortly after data collection started.

**Table 24. Freeway Vehicle Matching Statistics**

|               | <b>Minutes</b> | <b>Upstream<br/>Volume</b> | <b>Downstream<br/>Volume</b> | <b>Total<br/>Matches</b> |
|---------------|----------------|----------------------------|------------------------------|--------------------------|
| Segment 1 GP  | 10             | 391                        | 336                          | 267                      |
| Segment 2 GP  | 84             | 2382                       | 2000                         | 1799                     |
| Segment 1 HOV | 97             | 2251                       | 2242                         | 2083                     |

Figure 22 and Figure 23 show the travel time results for the GP lane and HOV lane respectively, for the selected time period. Table 25 displays the statistical summaries of four different intervals of a consecutive 5-minutes of data, aggregated on a per-minute and per-five minute basis.

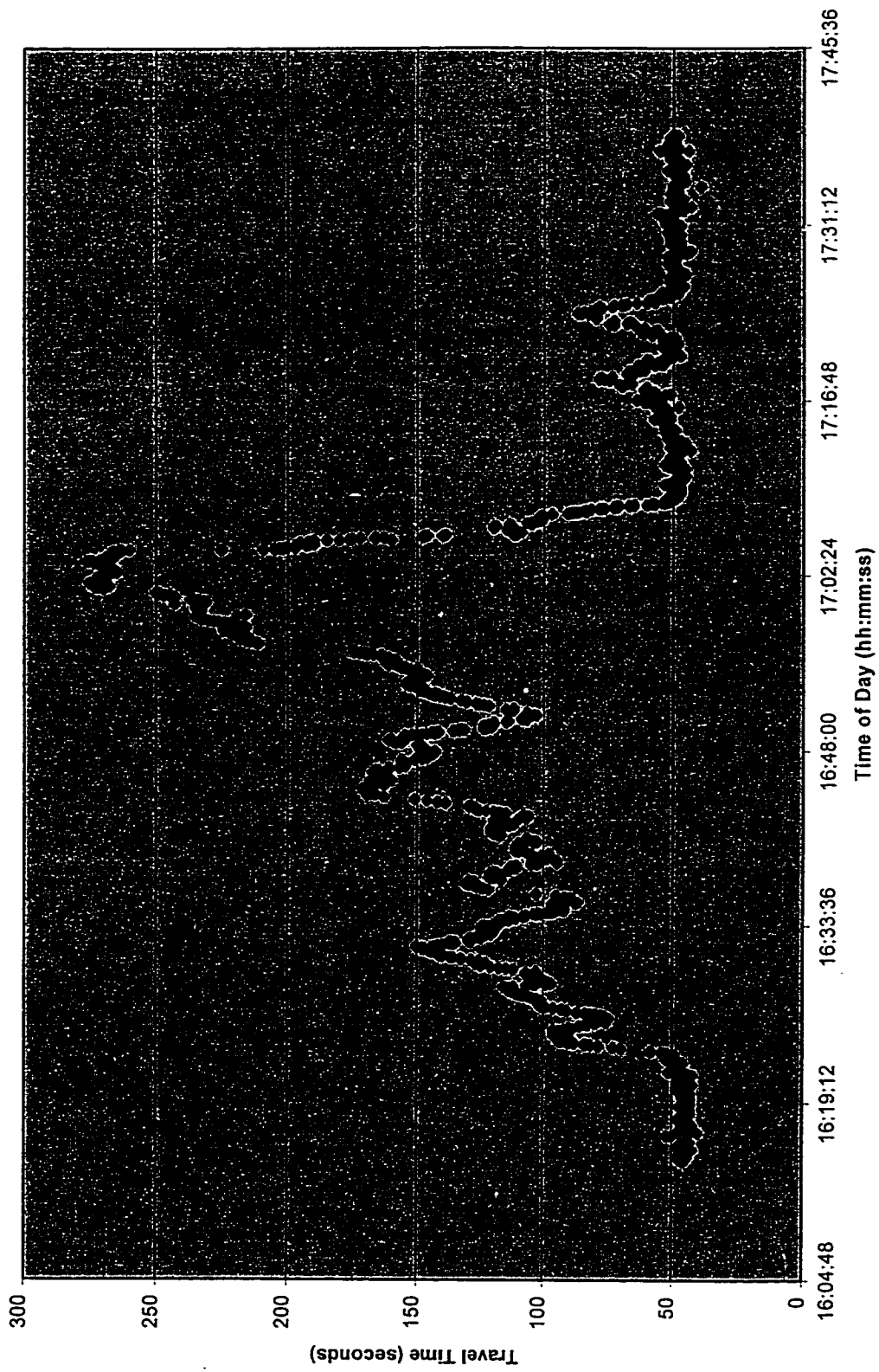
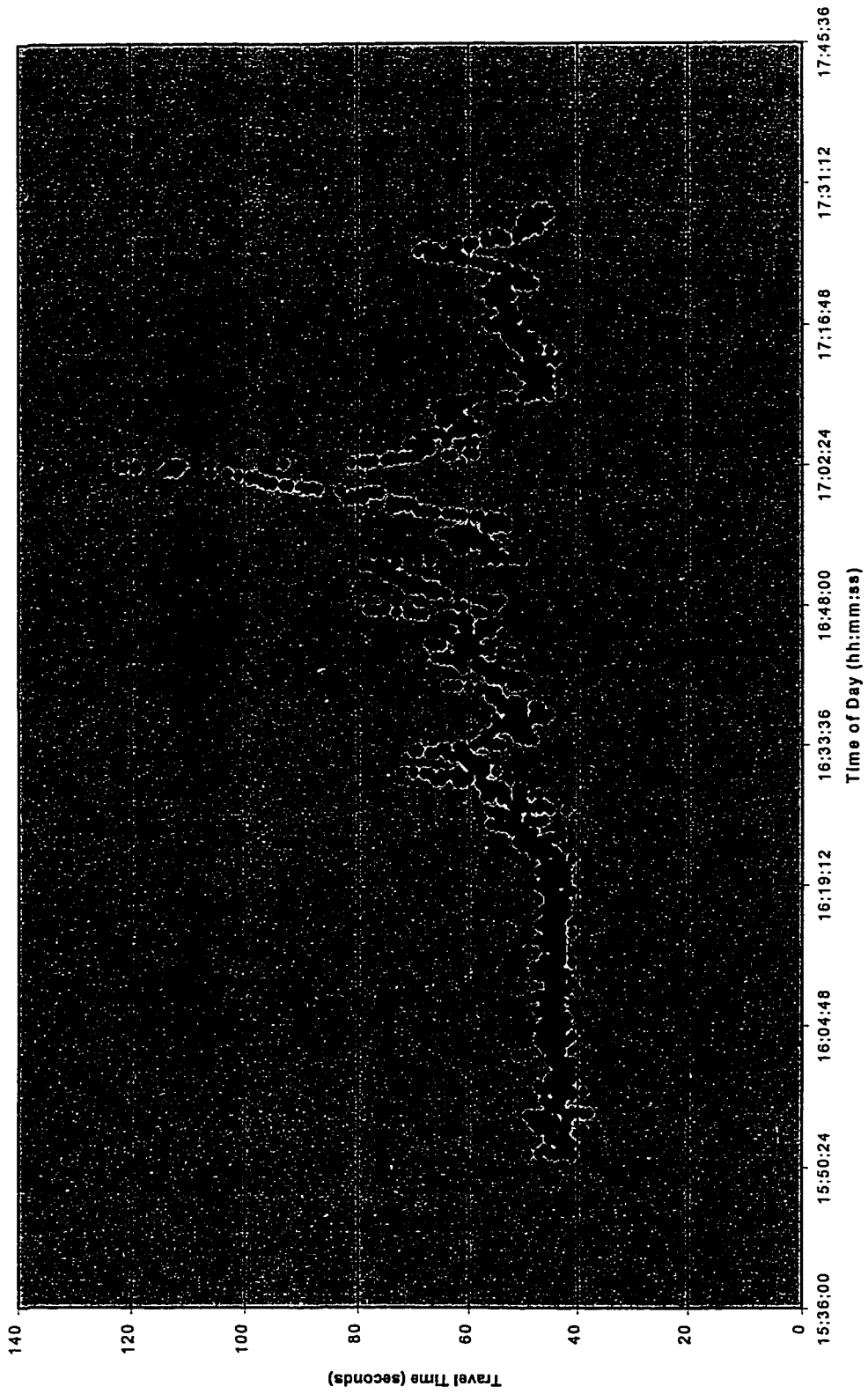


Figure 22. Overall Travel Time Results (GP Lane)



**Figure 23. Overall Travel Time Results (HOV Lane)**

Table 25. Sample Size Comparisons for Freeway Data

| Time Ending | 1-min. avg. (sec) | 1-min std. dev. | Avg. Speed (mph) | # Observations | Required Travel Time Samples |           |          | 5-min. avg. | 5-min std. dev. | Avg. Speed (mph) | # Observations | Required Travel Time Samples |           |          |
|-------------|-------------------|-----------------|------------------|----------------|------------------------------|-----------|----------|-------------|-----------------|------------------|----------------|------------------------------|-----------|----------|
|             |                   |                 |                  |                | 20% error                    | 10% error | 5% error |             |                 |                  |                | 20% error                    | 10% error | 5% error |
| 4:16        | 44.9              | 1.3             | 50.1             | 30             | 0                            | 0         | 1        |             |                 |                  |                |                              |           |          |
| 4:17        | 44.7              | 2.5             | 50.3             | 22             | 0                            | 1         | 5        |             |                 |                  |                |                              |           |          |
| 4:18        | 44.8              | 1.2             | 50.2             | 35             | 0                            | 0         | 1        |             |                 |                  |                |                              |           |          |
| 4:19        | 44.6              | 1.2             | 50.4             | 28             | 0                            | 0         | 1        |             |                 |                  |                |                              |           |          |
| 4:20        | 44.7              | 0.9             | 51.1             | 32             | 0                            | 0         | 1        | 44.6        | 1.4             | 50.4             | 147            | 0                            | 0         | 2        |
| 4:21        | 46.2              | 1.8             | 48.7             | 25             | 0                            | 1         | 2        |             |                 |                  |                |                              |           |          |
| 4:22        | 45.8              | 2.5             | 49.2             | 26             | 0                            | 1         | 5        |             |                 |                  |                |                              |           |          |
| 4:23        | 48.6              | 1.7             | 46.3             | 26             | 0                            | 0         | 2        |             |                 |                  |                |                              |           |          |
| 4:24        | 60.1              | 12.8            | 37.5             | 16             | 4                            | 17        | 70       |             |                 |                  |                |                              |           |          |
| 4:25        | 29.8              | 2.9             | 24.9             | 24             | 0                            | 0         | 2        | 58.1        | 19.0            | 38.7             | 17             | 10                           | 4         | 164      |
| 5:01        | 239.7             | 7.1             | 9.4              | 15             | 0                            | 0         | 1        |             |                 |                  |                |                              |           |          |
| 5:02        | 270.8             | 2.1             | 8.3              | 18             | 0                            | 0         | 0        |             |                 |                  |                |                              |           |          |
| 5:03        | 268.4             | 2.2             | 8.4              | 26             | 0                            | 0         | 0        |             |                 |                  |                |                              |           |          |
| 5:04        | 269.8             | 3.3             | 8.3              | 25             | 0                            | 0         | 0        |             |                 |                  |                |                              |           |          |
| 5:05        | 226.7             | 3.0             | 9.9              | 29             | 2                            | 8         | 33       | 254.6       | 25.5            | 8.8              | 13             | 15                           | 4         | 15       |
| 5:06        | 157.6             | 29.9            | 14.3             | 28             | 3                            | 14        | 55       |             |                 |                  |                |                              |           |          |
| 5:07        | 110.0             | 6.5             | 20.5             | 20             | 0                            | 1         | 5        |             |                 |                  |                |                              |           |          |
| 5:08        | 90.7              | 9.7             | 24.8             | 29             | 1                            | 4         | 18       |             |                 |                  |                |                              |           |          |
| 5:09        | 57.3              | 8.1             | 39.3             | 29             | 2                            | 8         | 30       |             |                 |                  |                |                              |           |          |
| 5:10        | 49.1              | 1.9             | 45.8             | 25             | 0                            | 0         | 2        | 92.6        | 42.7            | 24.3             | 13             | 20                           | 8         | 327      |

The intent of Table 25 is to illustrate the variability of necessary sample sizes for this freeway data for different travel time variances, permitted errors, sample sizes, and data aggregation level. The sample size values assume a 95 percent confidence level, and a z-value of 1.96<sup>3</sup>. Furthermore, these calculations do not take the finite population correction factor into account, although the sample sizes do comprise a very large percentage of the population.

The first five minutes of data represent free-flow conditions. Since these travel times are very consistent for each minute and for all five minutes, a very nominal number of travel time samples is necessary to determine the mean travel time with a low error level, for both 1-minute and 5-minute aggregation levels. The second five minutes represents conditions transitioning from free-flow to heavily congested. Only one minute (4:24) within the five-minute period requires a considerable number of matches. It is apparent that a lot of flow transitioning occurred during this minute. Despite only one of the 1-minute intervals exhibiting a lot of travel time variance, the 5-minute aggregate statistics show a considerable amount of variation, as expected for an average travel time doubling in length from the first minute to the fifth minute. Thus, even though a low error level for the mean travel time estimate could be obtained for four of the five 1-minute intervals with a nominal number of samples, the entire 5-minute aggregation would require a very large number of travel time samples. The third five minutes represents fairly constant stop-and-go conditions. Only the fifth 1-minute interval requires more than a nominal number of samples. Looking at the following 1-minute interval (5:06), it can be seen that the 5:05 interval experiences significant travel time variation due to transitioning from stop-and-go to less congested conditions. However, the 5-minute aggregation still only requires a small number travel time samples for a low error level.

The fourth five minutes represents conditions transitioning from heavy congestion to near free-flow. Three of the five 1-minute intervals exhibit substantial travel time

---

<sup>3</sup> This is a simplification, since an accurate t-value is dependent on the other parameter values.

variation and thus require a significant number of travel time samples. The other two 1-minute intervals require only a nominal number of samples. Within this five-minute period, the travel time drops by a factor of three from the first 1-minute interval to the fifth 1-minute interval. This dramatic change results in an extremely large number of travel time samples being required for a low error level for the 5-minute aggregation measurements. It is easy to see from these four examples that the 5-minute aggregation level is adequate for fairly constant traffic conditions, but completely inadequate for transitional flow conditions. More discussion on the effects of aggregation can be found in section 8.4.

During the transitional flow periods, it can be seen that the sample size requirements (for a high confidence level and low error level) are greater than the actual population of travel time measurements. Even with a high proportion of the population sampled and an adjustment made for the finite population, there will be time periods, even at the one-minute aggregation level, in which the traffic flow is fluctuating too much for a mean travel time measurement to be obtained with a high confidence level and low error.

### Travel Time Distributions

Revisiting the issue of travel time distribution normality, some excerpts of the collected freeway data will be used for examples. Travel time and average speed distributions are shown in Figure 24 through Figure 27. The two top figures correspond with generally free-flow conditions, representing about 10 minutes of data. The two bottom figures correspond with the most congested conditions observed during the entire data collection time, representing about 4 minutes of data.

There are several techniques for testing a distribution for normality, such as a chi-square test or the use the Kolomogorov D statistic, but a very simple technique as used by Berry is the skewness index, defined by the following equation.

$$\text{Skewness Index} = \frac{2(P_{93} - P_{50})}{P_{93} - P_{07}}$$

Where  $P_{xx}$  are percentile values, with the  $P_{50}$  value corresponding the median. If the distribution is perfectly normal, this index will equal 1.0. An index value below 1.0 indicates a distribution skewed toward longer travel times or slower speeds.

Other simple formulations for checking distribution skew have been suggested by May [1990]. The following two formulas can be used to calculate a coefficient of skewness.

$$\text{Coefficient of skewness (1)} = \frac{\text{mean} - \text{mode}}{\text{std.deviation}}$$

or

$$\text{Coefficient of skewness (2)} = 3 \times \frac{\text{mean} - \text{median}}{\text{std.deviation}}$$

If the distribution is perfectly normal, these coefficients will equal 0.0.

Computing these values for the figures displayed below yield the results shown in Table 26.

**Table 26. Distribution Skewness Comparisons**

|                             | Travel Time Dist. |             | Average Speed Dist. |             |
|-----------------------------|-------------------|-------------|---------------------|-------------|
|                             | Free-flow         | Stop-and-go | Free-flow           | Stop-and-go |
| Skewness Index              | 1.20              | 1.03        | 0.85                | 0.99        |
| Coefficient of skewness (1) | -0.81             | -0.53       | 0.83                | 0.54        |
| Coefficient of skewness (2) | 0.77              | -0.51       | -0.66               | 0.54        |

As can be seen from the table, neither distribution is perfectly normal, but the distributions for congested conditions are definitely less skewed than those for free-flow conditions.

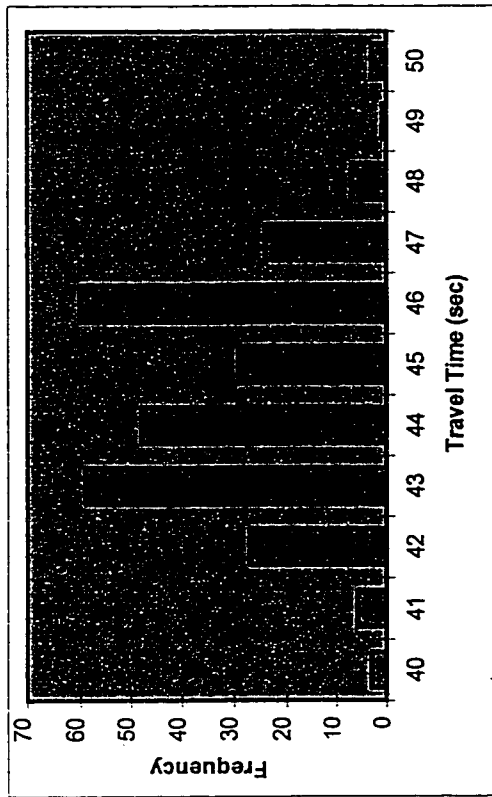


Figure 24. Travel Time Distribution (free-flow conditions)

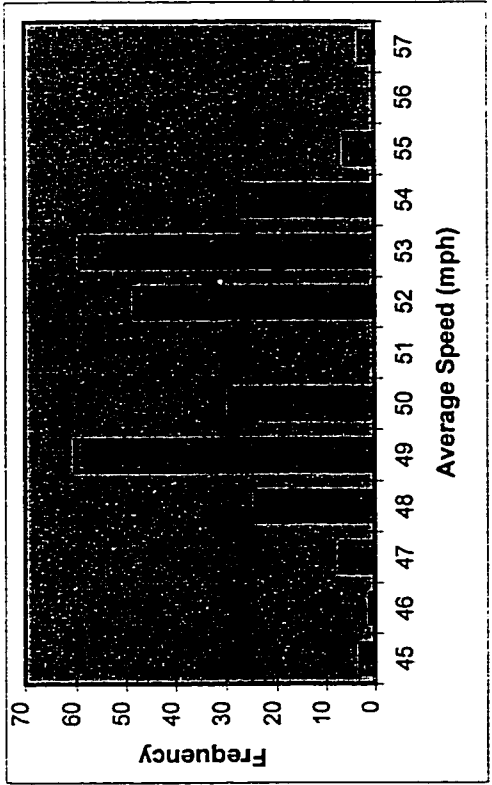


Figure 26. Average Speed Distribution (free-flow conditions)

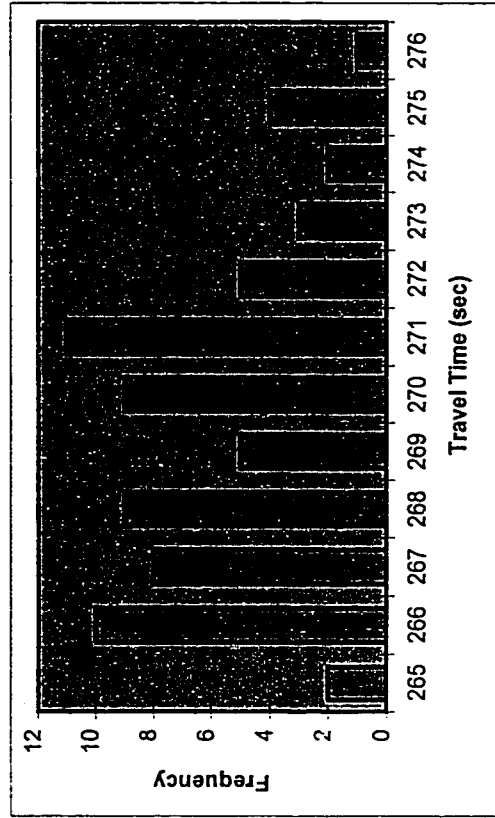


Figure 25. Travel Time Distribution (stop-and-go conditions)

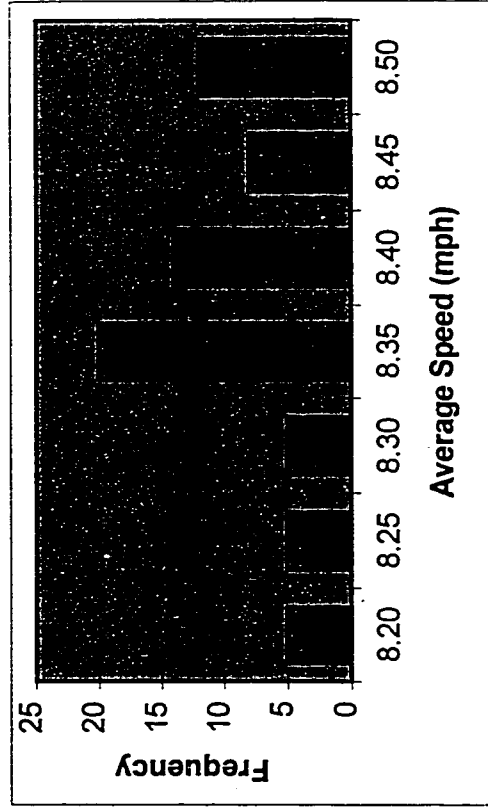


Figure 27. Average Speed Distribution (stop-and-go conditions)

## 7.5 Sampling Capabilities

This section summarizes the sampling capabilities of the keyboard and voice recognition sampling methods. Additionally, this section explores the effect of different plate entry schemes on the travel time results for the collected freeway data. This is accomplished by simulation based upon empirically determined sampling rates.

### *7.5.1 Field Measured Sampling Rates*

#### 7.5.1.1 Arterial

Table 27 presents the various sampling rates obtained for both the keyboard entry and voice recognition methods for the arterial data collection experiments. Furthermore, it shows what those sampling rates translate to for an hourly sample size. As indicated in the table, there were a variety of flow conditions and a few different time stamp techniques used for both methods.

Table 27. Sampling Rate Ranges for Arterials

| Data Collection Date <sup>a</sup> | Flow Rate (vphpl) | Flow Type <sup>b</sup> | Voice Recognition |                  | Keyboard Entry  |                  |
|-----------------------------------|-------------------|------------------------|-------------------|------------------|-----------------|------------------|
|                                   |                   |                        | Sample Rate (%)   | Sample Size/Hour | Sample Rate (%) | Sample Size/Hour |
| 5/8/98 <sup>c</sup>               | 905               | 1                      | NA                | NA               | 51.4            | 465              |
| 5/19/98 <sup>d</sup> (1)          | 880               | 2                      | NA                | NA               | 73.6            | 648              |
| 5/19/98 <sup>d</sup> (2)          | 936               | 2                      | NA                | NA               | 66.7            | 624              |
| 5/19/98 <sup>c</sup> (1)          | 880               | 2                      | NA                | NA               | 50.0            | 440              |
| 5/19/98 <sup>c</sup> (2)          | 936               | 2                      | NA                | NA               | 52.1            | 476              |
| 7/1/98 (1)                        | 740               | 2                      | 83.6              | 619              | 85.8            | 635              |
| 7/1/98 (2)                        | 776               | 2                      | NA                | NA               | 70.8            | 549              |
| 7/9/98 (1)                        | 848               | 2                      | 83.3              | 706              | 70.8            | 600              |
| 7/9/98 (2)                        | 756               | 2                      | 93.5              | 707              | 58.6            | 443              |
| 7/23/98 (1,1)                     | 485               | 3                      | 83.3              | 404              | 87.1            | 422              |
| 7/23/98 (1,2)                     | 636               | 3                      | 91.2              | 580              | 79.0            | 502              |
| 7/23/98 (2,1)                     | 442               | 3                      | 63.0 <sup>e</sup> | 278              | 88.9            | 393              |
| 7/23/98 (2,2)                     | 629               | 3                      | 69.0 <sup>e</sup> | 434              | 79.2            | 498              |
| 8/12/98 (1)                       | 432               | 4                      | 96.3              | 416              | NA              | NA               |
| 8/12/98 (1)                       | 432               | 4                      | 89.8              | 388              | NA              | NA               |
| 8/12/98 (2)                       | 388               | 4                      | 90.7              | 352              | 82.5            | 320              |
| 8/12/98 (3)                       | 376               | 4                      | 96.8              | 364              | 77.7            | 292              |
| 8/12/98 (4)                       | 432               | 4                      | NA                | NA               | 71.3            | 308              |
| 8/12/98 (4)                       | 432               | 4                      | NA                | NA               | 75.9            | 328              |
| 9/3/98 (1,1)                      | 389               | 3                      | 89.3              | 347              | 98.4            | 383              |
| 9/3/98 (1,2)                      | 562               | 3                      | 93.9              | 528              | 68.2            | 383              |
| 9/3/98 (2,2)                      | 612               | 3                      | 63.8 <sup>f</sup> | 390              | 63.0            | 386              |

<sup>a</sup> The first number in parentheses is the data collection session. The second number in parentheses is the data collection station, 1 for upstream, 2 for downstream.

<sup>b</sup> 1—free flow, 2—heavy platooning due to near upstream signal, 3—congestion and data collection at signal, 4—free flow, light traffic, and data collection at signal

<sup>c</sup> The focus of these data collection exercises was maximum time-stamp accuracy; thus, the sample rate was relatively low

<sup>d</sup> This focus of these data collection exercises was maximum sample size, without regard for precise time-stamping

<sup>e</sup> This data collection used the post-license plate entry manual time stamp. Additionally, there was a power glitch during the middle of the session at the downstream station that resulted in 6 potential vehicles not being recorded.

<sup>f</sup> This data collection used the pre-license plate entry manual time stamp.

### 7.5.1.2 Freeway

Using human observers to record license plates of freeway vehicles is certainly more challenging than arterial facilities. For safety reasons, it is not practical, nor prudent, to place an observer adjacent to freeway lanes. Thus, positioning observers on arterial overpasses is one of the few feasible alternatives for collecting license plate data from freeway facilities. Because of the distance above the roadway for typical overpasses, it is usually necessary for observers to use binoculars to read the license plates. And during times of high-speed flow, higher powered binoculars are usually necessary.

It was desired to obtain a range of sampling frequencies for a range of traffic flow conditions from a freeway overpass location. Unfortunately, sample rate values were only obtained for freeway conditions at the two flow condition boundaries—free flow and stop-and-go. Nonetheless, this provided a good estimate of the expected range of sampling rates. Table 28 summarizes the sample rates obtained from the freeway license plate data collection, as well as the potential sample sizes per hour for each specific set of conditions.

**Table 28. Sampling Rate Ranges for Freeways**

| Data Collection Date <sup>a</sup> | Flow Rate (vphpl) | Flow Type <sup>b</sup> | Voice Recognition |                  | Keyboard Entry  |                  |
|-----------------------------------|-------------------|------------------------|-------------------|------------------|-----------------|------------------|
|                                   |                   |                        | Sample Rate (%)   | Sample Size/Hour | Sample Rate (%) | Sample Size/Hour |
| 4/30/98 <sup>c</sup>              | 960               | 1                      | 50.9              | 489              | NA              | NA               |
| 9/29/98 <sup>d</sup> (1)          | 1374              | 2                      | 83.9              | 1153             | NA              | NA               |
| 9/29/98 <sup>d</sup> (2)          | 1308              | 2                      | NA                | NA               | 80.6            | 1054             |
| 9/29/98 <sup>d,f</sup> (3)        | 1446              | 2                      | NA                | NA               | 61.2            | 885              |
| 9/29/98 <sup>d</sup> (4)          | 1466              | 2                      | 77.8              | 1141             | NA              | NA               |
| 9/29/98 <sup>e</sup> (5)          | 1452              | 2                      | 80.3              | 1166             | NA              | NA               |
| 9/29/98 <sup>e</sup> (6)          | 1356              | 2                      | NA                | NA               | 57.6            | 781              |
| 9/29/98 <sup>e,f</sup> (7)        | 1480              | 2                      | NA                | NA               | 78.1            | 1156             |

<sup>a</sup> The first number in parentheses is the data collection session.

<sup>b</sup> 1—free flow, 2—stop-and-go

<sup>c</sup> It was obvious, however, that higher powered binoculars were still necessary.

<sup>d</sup> Since vehicles were moving very slowly (stop-and-go traffic) and the observation overpass was relatively low, the license plates were able to be read without the assistance of binoculars.

- <sup>e</sup> Although the conditions were similar to those in footnote c (except other traffic direction was collected), the data collectors were instructed to use binoculars (low power) anyway.
- <sup>f</sup> This data collection was performed by two people simultaneously—one person reading the license plates aloud, the other typing them into the program.

### *7.5.2 Sampling Simulation*

Since sampling at multiple locations along a freeway for travel time measurement was not performed, a simulation approach was used. From the empirically determined sample rates at individual locations, approximate sampling distributions were developed for use in simulating license plate sampling from the freeway data collected by video.

#### 7.5.2.1 Sampling Function

Exact quantification of the vehicle license plate sampling process by human observers is a complex issue. However, with empirically determined sampling rates for various conditions, rough approximations can be made which are likely to yield reasonable results for simulation purposes. A discussion on the components of sampling by human observers will now be discussed.

The three main components to sampling a vehicle license plate are as follows:

- Visual location of the license plate: This is basically random. It is mainly a function of plate location on the vehicle.
- Plate character recognition/reading: This is a function of distance from the observer to the vehicle, a function of the number of characters read, and a function of vehicle speed.
- Plate character array entry time: A function of the entry method used. This component will probably have the most variation between personnel.

No attempt was made to isolate the contribution of each of these individual components to the overall amount of time required for sampling a vehicle license plate from the traffic stream, but some discussion on how they might vary by method is in order.

The time required for visual location of the license plate is unlikely to vary significantly by data collection method used. However, for a methods which allow the observer to keep their eye on traffic all the time, such as voice recognition, there may be a slight advantage over methods which necessitate an observer to look away from traffic on occasion. For a method like voice recognition, the second two components can actually be accomplished in the same step. That is, in the process of reading the license plate (aloud), the license plate is also being simultaneously entered. For the keyboard method, the same could be accomplished if the data collector is a very accomplished typist. For the average data collector, however, they are likely to read and memorize the plate characters first, and then type them into the program, usually looking at the keyboard while doing so. For the first and second components, it will generally take more time when using binoculars to find and read the license plates.

Analysis of the empirical freeway sampling data for free-flow and stop-and-go conditions indicated that the distribution of headways between successive sampled vehicles generally followed a distribution similar in form to a Pearson Type III distribution. Based upon this observation, and the calculated means and standard deviations, sampling distributions were created for three different traffic flow ranges. The three flow ranges were free-flow, stop-and-go, and transitional flow. Since, as mentioned, data were only obtained for free-flow and stop-and-go conditions, the transitional flow distribution was developed with the assumption that the sampling rate varies linearly between the sampling rates for free-flow traffic and stop-and-go traffic.

Although the general distributional form for sampling is similar to a Pearson Type III distribution, using this specific functional form was not the objective. Instead, it was desired to develop the three distributions such that they followed this general form, but were still comparable to the empirical distributions in mean and standard deviation. Figure 28 displays the three developed sampling distributions. Table 29 displays the results of the empirical distributions and the model distributions, as well as the headway statistics for the entire ground truth volume.

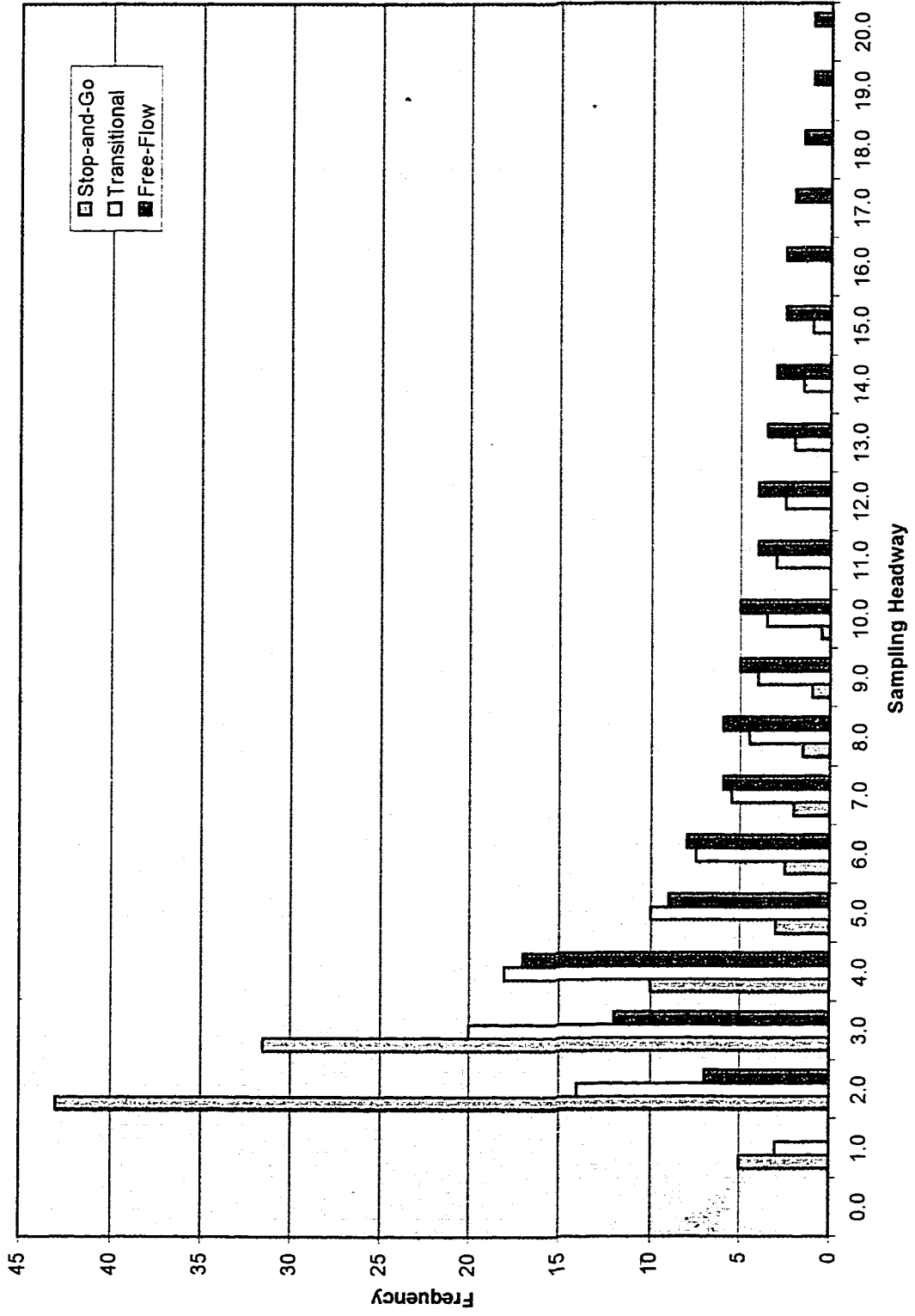


Figure 28. Modeled Manual Sampling Distributions for Freeway Traffic

**Table 29. Manual Sampling Distributions for Freeway Traffic**

|                                 | <b>Avg.<br/>Headway</b> | <b>Std. Dev.<br/>Headway</b> |
|---------------------------------|-------------------------|------------------------------|
| <b>Free-Flow</b>                |                         |                              |
| All Vehicles                    | 3.77                    | 2.91                         |
| Sampled Vehicles                | 7.40                    | 4.89                         |
| Distribution Model              | 7.53                    | 4.61                         |
| <b>Transitional<sup>a</sup></b> |                         |                              |
| All Vehicles                    | 3.06                    | 1.98                         |
| Sampled Vehicles                | 5.17                    | 3.07                         |
| Distribution Model              | 5.29                    | 3.28                         |
| <b>Stop-and-Go</b>              |                         |                              |
| All Vehicles                    | 2.34                    | 1.04                         |
| Sampled Vehicles                | 2.93                    | 1.25                         |
| Distribution Model              | 2.96                    | 1.55                         |

<sup>a</sup> This was estimated based on a linear relationship between the free-flow and stop-and-go flow conditions.

These distributions are based on the voice recognition method. While there is only a small difference between the stop-and-go sample and accuracy statistics of the voice recognition and keyboard method, there is a big difference in the sample rate for free-flow conditions where the observer needs to use binoculars. The voice recognition method sample rate is higher in this case because it is easy to use binoculars and speak at the same time. It is not nearly as easy to use binoculars and type at the same time.

#### 7.5.2.2 Sampling Methodology Testing

A software program was written to facilitate the testing of different sampling schemes (this is the same application that performed the basic matching, a brief review of the simulation aspects are also covered in Appendix C). Essentially, the program assumes a data collector at the origin location overpass and the destination location overpass for the lane of choice (either GP or HOV). Before simulation, the user selects the plate characters that will be entered (e.g., plate characters 1-4). The program then converts all license plate text strings to the appropriate strings representing the plate

character entry scheme. As the program simulates the flowing of the vehicles (based on the video ground truth data), a probabilistic decision is made as to whether each vehicle that passes under the overpass is sampled or not. This decision is based upon the headway between this vehicle and the previous vehicle sampled and the sampling rate distribution developed from the empirical sample rate studies. The sample headway is determined by implementation of the probability curves shown in Figure 29, developed from the distributions shown in Figure 28. This is done for each sampling location. Once a vehicle has been selected for sampling, another decision is made as to whether the license plate of that particular vehicle is entered correctly or not. This is done randomly and based upon empirically determined accuracy rates. A conservative, yet realistic, accuracy rate of 90% was assumed and implemented in the program with a uniform distribution. Once the vehicle sampling is done for the required time period, the program performs license plate matching. The travel time results are then compared to the ground-truth results.

The following thresholds were used to define the different flow conditions for this freeway data set:

- Free-flow—travel time < 50 seconds (45 mph)
- Transitional-flow— $50 \leq$  travel time < 150 sec (15 mph)
- Stop-and-go—travel time  $\geq$  150 seconds

There were two objectives for this simulation. One was to examine the tradeoffs in sample size and spurious matches due to different plate character entry schemes. The second was to see if the obtained samples would provide a reasonable representation of the true population travel time distribution.

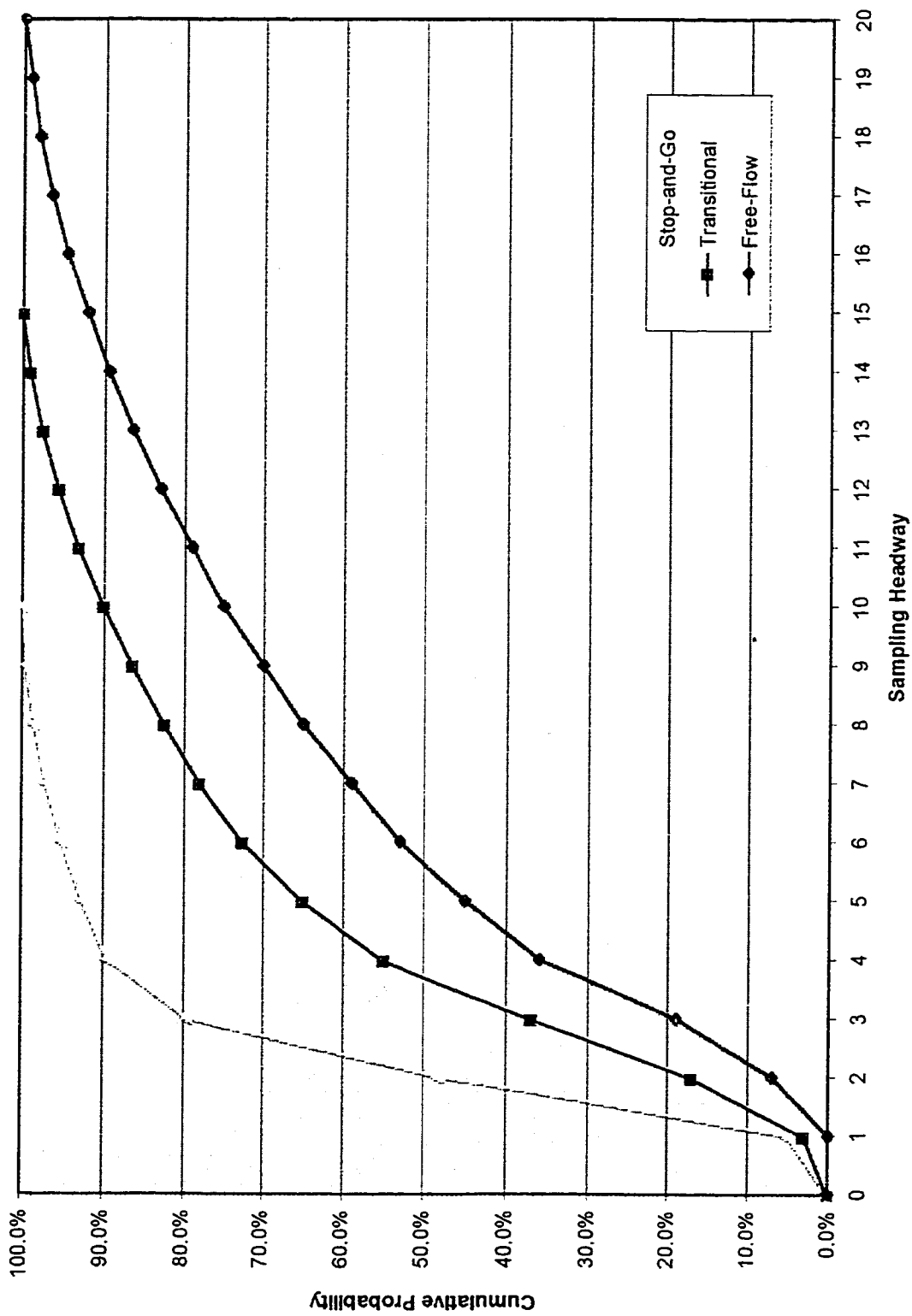


Figure 29. Sampling Probability Functions

### 7.5.2.3 Results and Interpretation

Table 30 shows the results of various sampling schemes using the above simulation methodology. The results in the table are based upon an upstream volume of 2345 vehicles and a downstream volume of 1973 vehicles. The total number of matches for these vehicles is 1775. This is slightly less than the numbers reported in Table 24 because some vehicles that did not have license plates but were confirmed matches were removed from the data set for simulation purposes.

**Table 30. Simulated Travel Time Results, GP Lane (I-5, 130th - 145th)**

| Plate Characters Entered, Entry Accuracy Rate | # of Runs | Total Matches | Spurious Matches | % Spurious Matches | Correct Matches |
|---|-----------|---------------|------------------|--------------------|-----------------|
| Chars 1-4, 100%                               | 10        | 350           | 45               | 12.7               | 305             |
| Chars 3-6, 100%                               | 10        | 321           | 16               | 5.1                | 305             |
| Chars 1-3, 5, 100%                            | 5         | 309           | 10               | 3.2                | 299             |
| Chars 1-3, 6, 100%                            | 5         | 318           | 16               | 5.0                | 302             |
| Chars 1-4, 90%                                | 3         | 286           | 39               | 13.5               | 247             |
| Chars 3-6, 90%                                | 3         | 267           | 12               | 4.4                | 256             |
| Chars 1-3, 100% <sup>b</sup>                  | 1         | 767           | 451              | 58.8               | 316             |
| Chars 4-6, 100% <sup>b</sup>                  | 1         | 494           | 207              | 41.9               | 287             |
| Chars 1-5, 100% <sup>b</sup>                  | 1         | 294           | 6                | 2.0                | 288             |
| Chars 2-6, 100% <sup>b</sup>                  | 1         | 309           | 2                | 0.6                | 307             |

<sup>a</sup> For these runs, there were from 1 to 5 vehicles whose plates were incorrectly recorded at both the upstream and downstream end. These were not counted as correct matches, but if the same mistake was made at both stations when entering these plates, they could appear to be correct matches.

<sup>b</sup> These statistics do not reflect the slightly different sampling rate distributions for a total number of characters other than 4.

For the recording of 4 characters from the license plate, the correct match percentages are about 14 for 90 percent entry accuracy and about 17 for 100 percent entry accuracy. Recording of the first 4 characters generally results in 3 numbers and 1 letter. This is because Washington State license plates for passenger cars predominately use a format of three numbers followed by three letters. This is the format for the centennial plates which have been issued since 1986. Prior to this, passenger car plates used a

format of three letters followed by three numbers. These plates currently comprise a very low percentage in the overall vehicle population.

As expected, the sample schemes that use the first four characters result in a greater percentage of spurious matches than the other schemes. This reflects the predominance of license plates with 3 numbers followed by 3 letters and the smaller number of unique character combinations using numbers instead of letters. While there are out-of-state licenses and other vehicle plate formats in the data set, Washington state passenger cars make up the majority of the traffic stream. The number of spurious matches for the plate entry scheme of characters 3-6 was about 1/3 of those for the plate characters 1-4 entry scheme, on average. However, the total number of correct matches were basically the same for the two entry schemes. Thus, the greater number of matches for the 1-4 character entry scheme is offset by the greater number of spurious matches. In practice, however, one should expect a slightly lower number of correct matches with the 3-6 character entry scheme. This is because this scheme requires a little more time for assimilation and entry by the data collector. For vehicles such as trucks, where there are no spaces between the first and second set of three characters (many trucks have license plates with continuous characters), finding the third character from which to start reading can take a little more time for the data collector. In addition, for the keyboard method, it is likely that it would take a little longer to enter 1 number and 3 letters than 3 numbers and 1 letter. For non-touch typists, it is generally easier to locate the numbers on the keyboard than the letters. For the voice recognition method, it also would probably take a little longer for entering 3 letters and 1 number due to the greater degree of recollection required for the "letters". Since the "letters" are really military alphabet words that must be memorized, it is reasonable to expect this to take a little longer than the numbers. Thus, the first four characters scheme will probably lead to the largest correct match sample size. However, if the goal is to minimize spurious matches and still obtain a reasonable sample size, the last four characters (or 3-6) would be a better strategy.

It was suspected that a scheme that used the first 3 characters and either the fifth or sixth character would result in fewer spurious matches than using the first 4 characters.

For passenger cars, this would generally result in the recording of the first 3 numbers and either the second or third letter instead of the first 3 numbers and the first letter. The reasoning behind this theory is because Washington state issues license plates alphabetical order. In other words, all possible combinations of license plates that have an 'A' as the first letter are issued before any plates that have a 'B' as the first letter (e.g., 000 AAA – 999 AZZ). Currently, Washington state is not even halfway through the alphabet yet. Thus, the first letter of centennial passenger plates is limited to just 'A' through 'K', instead of 'A' through 'Z' like the second and third letters. This makes it more likely to obtain a spurious match using the first letter over the second or third letters. As expected, the spurious matches were indeed lower for both the '1-3, 5' and '1-3, 6' schemes than the '1-4' scheme. Interestingly though, the total number of correct matches were not significantly different. Again, the total correct matches for both of these schemes was similar to that obtained with the '1-4' scheme. If a low number of spurious matches are desired, either the '1-3, 5' or '1-3, 6' method is preferable to the '1-4' method. It is possible, though, that these methods may require a slightly longer time for plate character assimilation by a data collector than the relatively straight forward first four characters technique.

What is of interest for the last two schemes is the spurious match percentage. As one would expect, for recording of 5 characters, the spurious match rate is very low. The number of matches is probably higher than would be realized for recording of 5 characters. This is because the average sample headway would generally increase for more characters entered. This simulation, however, still used the sampling distributions based on the entering of 4 characters.

Figure 30 displays the travel time results from two simulation options, overlaid on the ground truth data. This figure shows that both schemes provide good "coverage" of the travel time throughout the entire time period and for all flow conditions.

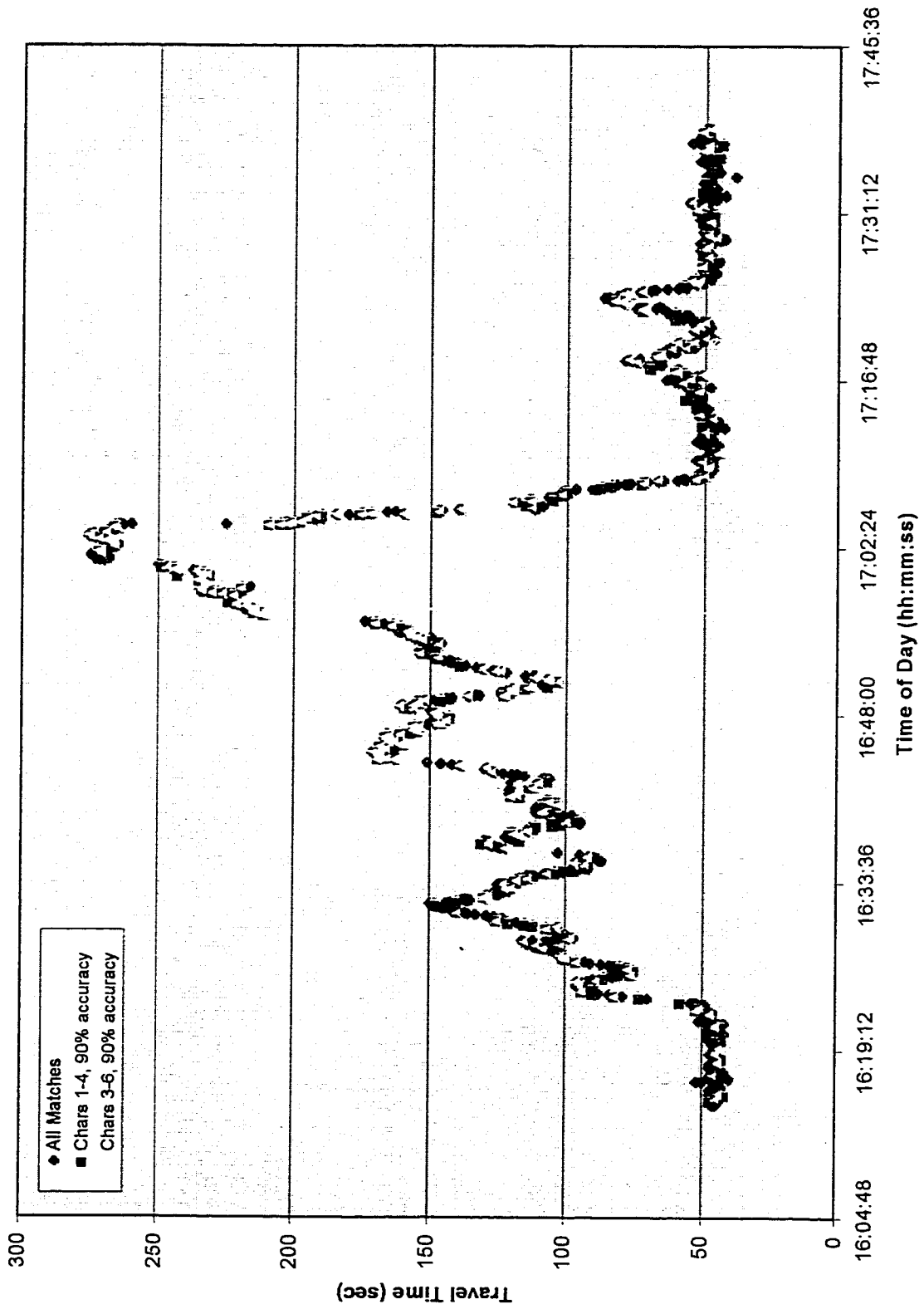


Figure 30. Travel Time Sampling Comparisons

For comparison purposes, Figure 31 shows the travel time tracking obtainable from floating car runs at various headways. Five-minute headways provide fairly course “coverage” of the travel times, but smoothes out most of the high and low points for transitional flow conditions. This headway level would probably be adequate for ATIS purposes, but would not be adequate for operational purposes. The two-minute headway offers very good tracking of the travel time—it smoothes out the minor fluctuations, but follows the general trends very well. The one-minute headways offer excellent tracking, as expected, but with twice the resources (vehicles, personnel) required over the two-minute headway option, does not appear to be warranted for the minor difference in travel time tracking ability.

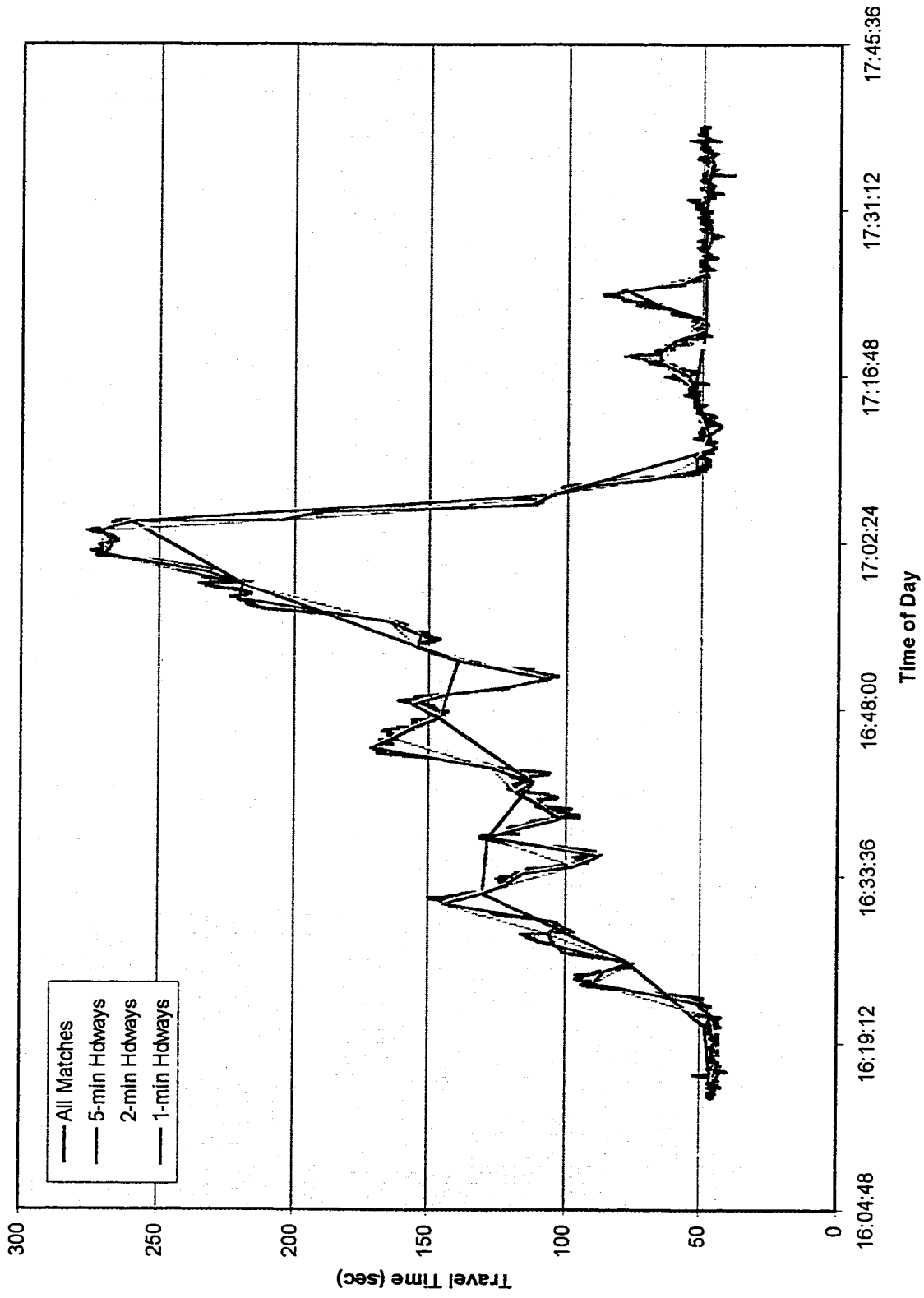


Figure 31. Travel Time Tracking by Floating Cars at Various Headways (GP Lane)

## 7.6 Conclusions and Recommendations

The cyclical nature of travel times on signalized arterials makes obtaining travel time measurements with a high confidence level or low error very challenging. The relatively high error level obtained in these arterial data collection efforts would probably not be adequate for evaluating operational changes, such as revised signal timing. Even with video recording of all passing vehicles, there were not enough through vehicles to provide a sample size that would give a high level of confidence or low error. For this particular section, another intermediate data collection would have been necessary, or recording vehicle plates from the inside lane also, to capture more through vehicles. The challenge is that additional personnel are required one way or the other. If one person samples from both lanes, the sample rate for each lane individually will be much lower. Matching also becomes a difficult proposition. With fewer vehicles sampled from each lane, the overlap of the same vehicles collected at both stations becomes smaller; thus the match rate will be lower. Furthermore, without designating the lane of travel for a recorded vehicle, travel time statistics will not be able to be separated by lane. Ideally, a separate data collector should be used for each lane of traffic. Since collecting two lanes of traffic with a separate observer for each would require four data collection personnel, it would be more cost effective to record for just one lane and use an intermediate recording location, which would require just one additional data collector. Adding an intermediate data collection point does create the possibility that the two sub-segment travel times may not be equivalent to the travel time for all vehicles traveling from beginning to end of the segment. Careful attention must be paid to this potential situation. The cyclical nature of travel times on signalized arterials also makes the mean travel time an inadequate measure by itself. While the median is useful for comparing to the mean for determining symmetry of the distribution, it also is not a terribly meaningful measure for signalized arterial flow. The range of expected travel times and the modes are the most meaningful measures for describing the travel time distribution characteristics on this type of facility. The distinction between mode and modes is an important one. If the distribution is truly

a normal distribution with only one obvious peak centered on the mean, than the reporting of one mode is sufficient. However, in many cases, there will be one or more sub-modes corresponding with travel times for poor and/or good progression. In this case, it is important to report these other modes. The reporting of only a single mode in this case can be very misleading, especially if that mode is not similar to the mean, which is highly likely.

With three different methods (video, keyboard entry, voice recognition) for measuring the arterial travel time, the travel time statistics between all methods were very comparable. In fact, no statistically significant difference in the mean travel times between the three methods was found. This was the case even during the second data collection session when the number of matches for the voice recognition method was considerably lower than the other methods due to using a different timestamp technique. However, with a lower number of matches, the confidence level will usually be lower or the error value higher. Even though the voice recognition method had a lower standard deviation value for the fewer number of matches, the error value was still higher than that for the other methods (8.8% versus 7.7% and 7.9%), using a 95 percent confidence interval. For all the data collection sessions, a low enough error or high enough confidence interval was not obtained to be used for assessing operational changes. For this particular arterial section, there were two major intersecting arterials that significantly reduced the percentage of through traffic between the two data collection locations. Aside from this influence, it is important to consider how representative the obtained travel time statistics are of the traffic conditions that one wants to evaluate. As seen in the results, the travel time statistics varied greatly during the PM peak period for the different days of data collection. Thus, it is equally important to consider the sampling requirements between days as it is within days.

For freeway traffic, attention must be given to the selected aggregation intervals for which sample size requirements will be based. For fairly constant traffic conditions, whether free-flow or stop-and-go, the sample size requirements will be modest, regardless of the aggregation interval. For transitional flow conditions, the sample size

requirements can be excessive. For transitional flow, 1-minute aggregations of travel time statistics are recommended over 5-minute aggregations. It is still possible to obtain low confidence levels or high error levels for individual 1-minute intervals, but the overall statistics will generally be robust enough to overcome these deficiencies on an intermittent basis, and still convey reliable travel time information for the entire period.

The simulated sampling of license plates from freeway traffic based on the voice recognition method showed that good tracking of the true travel time was possible. Although the sample rate is much lower for free-flow conditions than for transitional and stop-and-go conditions, the free-flow conditions will generally have less variance in travel times than the other conditions; thus, fewer samples are required for accurate travel time estimation. As for the simulation of entering different quantities and combination of license plate characters, the entering of 4 characters offers the best compromise. It is clear that the recording of only 3 characters is totally inadequate, no matter the letter/number composition. While filtering out the spurious matches is not a difficult task, it can be tedious. Without the benefit of ground truth data like this study had, the identification of spurious matches could be more difficult. The addition of a fourth plate character recorded will give the researcher considerably more confidence in the match results. If the minimization of spurious matches is a primary objective, the recording of the last four characters instead of the first four will result in significantly fewer spurious matches for the current vehicle license plate composition in Washington state. This will also apply in other states where passenger car vehicle plates are predominantly of the format of three numbers followed by three letters. However, it is likely that this scheme will also result in a little lower sample rate. To further minimize the spurious matches, the second or third letter on the plate should be entered instead of the first letter. This is because Washington state and other states assign license plates in an alphabetic manner; thus, there are fewer possible letters for the first letter than the other letters.

Since the statistical procedures discussed in this chapter rely on the assumption of a normal, or approximately normal, distribution, further study of travel time distributions for all facility types and flow conditions is needed to ensure whether this assumption

holds for all conditions. Additionally, the issue of differences in the travel time distribution from the corresponding speed distribution deserves further attention.

## CHAPTER 8

### TRAVEL TIME MEASUREMENT BIAS ISSUES

---

#### 8.1 Introduction

There are three objectives for this chapter. The first objective is to describe what the significant bias issues are to consider in travel time measurement. The second objective is to demonstrate some of these bias issues through data collection experiments, as well as from examples from the previous literature. The third objective is to discuss how these bias issues can vary by data collection technique.

For any point sampling technique (e.g., license plate matching), there are two possible sources of error that could be introduced into the measurement of travel time: 1) selection bias, which can be manifested in two different ways; a) inaccurate representation of the traffic stream population composition in the sample; b) over-sampling of either slower or faster vehicles in the traffic stream; and 2) inconsistent reference location time-stamping (also referred to as time measurement error). Additionally, aggregation bias can occur if the chosen data aggregation interval is not appropriate for the given traffic conditions. And finally, incorporation of statistical outlier points into the data analysis will lead to biased results.

If the sampling of vehicles from the traffic stream is done in a truly random fashion, and bias is not incorporated into the measurements, the sample distribution of travel times should be representative of the population travel time distribution, assuming enough samples are taken.

#### *8.1.1 Historical Measurement Techniques*

##### 8.1.1.1 License Plate Matching

When the license plate matching technique was first investigated as a method for determining travel time in the late 1940's, portable computers obviously did not exist. The most sophisticated pieces of recording equipment at the time were portable voice

recorders. Even though the voice recorder was more automated than using pencil and paper, there was still a lot of manual processing time required for transcribing the plate and time-stamp entries after the data collection. Thus, it was very advantageous to implement a sampling scheme which required data collectors to only collect a small subsample of the traffic stream license plates which also did not result in a biased estimate of travel time. This would result in greatly reduced post-processing time in the office, and thus reduced study costs. With the advent of portable computers, however, reductions in sample size are not necessary for post-processing concerns as all processing is done automatically via computer software.

Additionally, with the use of voice recorders or pencil and paper, stopwatches were utilized to record the precise time a vehicle passed the observation point. This process would entail pressing the "lap-time" button on the stopwatch to hold the passing time in the display window, while the current time continues to elapse in the "background". Thus, depending on positioning, the data collector could press the "lap-time" button as a vehicle crosses the reference point and then read the license plate into the voice recorder and then say the stopwatch time into the recorder. As it is, this process undoubtedly was lengthy enough that a 100 percent sample size was never possible, except possibly for the most severely congested conditions. The one advantage in utilizing a stopwatch in this way is that it is easy to ensure that the user gets the exact time of the vehicle crossing the reference point.

In an effort to reduce post-processing time, several sampling techniques were considered. One method involved sampling only vehicles with specific license plate endings [Berry, 1952; Walker, 1957]. Another method consisted of sampling only vehicles of a particular color [Berry, 1952; Sosslau and Honts, 1975]. However, these restrictive sampling techniques have the potential to introduce selection bias into the measurements. With the advent of the portable computer, however, the need for artificial sampling schemes which are intended to reduce the sample size to as small as practicable are no longer necessary. Since the data reduction is performed by computer, it makes sense to collect as much as data as possible. The issue now is not "*How does one reduce*

*the amount of data to a manageable level and still achieve statistical validity?”*, but “*How does one collect as much data as possible while minimizing potential sources of bias at the same time?”*.

## **8.2 Selection Bias for Point Sampling**

### **8.2.1 Selection Bias Due to Inaccurate Representation of Traffic Stream Composition**

For manual forms of point sampling (e.g., license plate recording), without a specific sampling strategy in mind, a data collector may be predisposed to recording information for certain types of vehicles disproportionately more than others. Because of the difference in plate styles for different types of vehicles (e.g., passenger cars, commercial vehicles, transit vehicles), a data collector may subconsciously enter license plates of only passenger cars, for example, because they get used to entering that particular format. Different vehicle classifications (e.g., pickup trucks, motorcycles, transit buses, commercial vehicles, etc.) all have different plate formats, in terms of the number of plate characters and arrangement of those characters. Thus, the data collector may ignore certain vehicle types just to avoid any potential confusion about which particular characters to enter. Under a sampling methodology in which the data collector is instructed to not collect certain vehicle types, or another scheme that leaves this to the data collector’s discretion, this result is more than plausible.

Bias due to inaccurate representation of vehicle type percentages is directly related to the sampling strategy, or lack thereof, of the data collector. Since it is virtually impossible for the data collector to know the traffic stream composition of the population before the study (although a study of historical classification data can help), a sampling scheme that does not rely on this knowledge is certainly desirable.

#### **8.2.1.1 Example Demonstration of Selection Bias on an Arterial**

Table 31 and Table 32 present the vehicle classification statistics obtained from the video for the first two data collection sessions on NE 45<sup>th</sup> Street. The two columns

under each heading represent the vehicle type frequency for the first and second data collection sessions, respectively.

The vehicle type matches recorded by each of the data collection methods were compared to the total traffic stream as captured on video for the first data collection session. This is shown in Table 33.

**Table 31. Vehicle Type Percentages for First Arterial Data Collection Session**

| Vehicle Type                    | Total Volume     |                 | Veh Type %       |                 |
|---------------------------------|------------------|-----------------|------------------|-----------------|
|                                 | 15 <sup>th</sup> | 7 <sup>th</sup> | 15 <sup>th</sup> | 7 <sup>th</sup> |
| Motorcycle                      | 1                | 1               | 0.5              | 0.4             |
| Compact / Economy Passenger Car | 29               | 72              | 14.5             | 27.3            |
| Luxury / Sedan Passenger Car    | 85               | 99              | 42.5             | 37.5            |
| High Performance / Sports Car   | 0                | 3               | 0.0              | 1.1             |
| Sport Utility Vehicle           | 24               | 22              | 12.0             | 8.3             |
| Mini-van                        | 11               | 26              | 5.5              | 9.9             |
| Pick-up (small to large)        | 15               | 28              | 7.5              | 10.6            |
| Passenger Car with Trailer      | 0                | 0               | 0.0              | 0.0             |
| Pick-up with Trailer            | 0                | 0               | 0.0              | 0.0             |
| Recreational Vehicle            | 0                | 0               | 0.0              | 0.0             |
| Bus                             | 12               | 8               | 6.0              | 3.0             |
| Single-unit Truck               | 6                | 5               | 3.0              | 1.9             |
| Semi-tractor with Trailer       | 0                | 0               | 0.0              | 0.0             |
| Unknown                         | 17               | 0               | 8.5              | 0.0             |
| <b>Total</b>                    | <b>200</b>       | <b>264</b>      | <b>100.0</b>     | <b>100.0</b>    |

**Table 32. Vehicle Type Percentages for Second Arterial Data Collection Session**

| Vehicle Type                    | Total Volume     |                 | Veh Type %       |                 |
|---------------------------------|------------------|-----------------|------------------|-----------------|
|                                 | 15 <sup>th</sup> | 7 <sup>th</sup> | 15 <sup>th</sup> | 7 <sup>th</sup> |
| Motorcycle                      | 0                | 1               | 0.0              | 0.4             |
| Compact / Economy Passenger Car | 39               | 73              | 21.2             | 28.4            |
| Luxury / Sedan Passenger Car    | 64               | 93              | 34.8             | 36.2            |
| High Performance / Sports Car   | 4                | 6               | 2.2              | 2.3             |
| Sport Utility Vehicle           | 18               | 22              | 9.8              | 8.6             |
| Mini-van                        | 16               | 19              | 8.7              | 7.4             |
| Pick-up (small to large)        | 14               | 21              | 7.6              | 8.2             |
| Passenger Car with Trailer      | 0                | 0               | 0.0              | 0.0             |
| Pick-up with Trailer            | 0                | 0               | 0.0              | 0.0             |
| Recreational Vehicle            | 0                | 0               | 0.0              | 0.0             |
| Bus                             | 14               | 14              | 7.6              | 5.4             |
| Single-unit Truck               | 6                | 6               | 3.2              | 2.3             |
| Semi-tractor with Trailer       | 0                | 0               | 0.0              | 0.0             |
| Unknown                         | 9                | 2               | 4.9              | 0.8             |
| <b>Total</b>                    | <b>184</b>       | <b>257</b>      | <b>100.0</b>     | <b>100.0</b>    |

Table 33. Vehicle Classification Frequencies for Origin/Destination Averaged and for Matched Vehicles (Session 1)

| Vehicle Classification          | Orig.      | Dest.      | Avg        | Avg %         | Video     | Video %       | Kybd      | Kybd %        | Voice Rec      | Voice Rec %   | All Matches | All Matches % |
|---------------------------------|------------|------------|------------|---------------|-----------|---------------|-----------|---------------|----------------|---------------|-------------|---------------|
| Motorcycle                      | 1          | 1          | 1          | 0.4%          | 0         | 0.0%          | 0         | 0.0%          | 0              | 0.0%          | 0           | 0.0%          |
| Compact / Economy Passenger Car | 29         | 72         | 51         | 21.8%         | 11        | 19.6%         | 9         | 24.4%         | 10             | 22.7%         | 12          | 18.2%         |
| Luxury / Sedan Passenger Car    | 85         | 99         | 92         | 39.7%         | 24        | 42.9%         | 14        | 31.7%         | 17             | 38.6%         | 28          | 42.4%         |
| High Performance / Sports Car   | 0          | 3          | 2          | 0.6%          | 0         | 0.0%          | 0         | 0.0%          | 0              | 0.0%          | 0           | 0.0%          |
| Sport Utility Vehicle           | 24         | 22         | 23         | 9.9%          | 9         | 16.1%         | 7         | 17.1%         | 6              | 13.6%         | 9           | 13.6%         |
| Mini-van                        | 11         | 26         | 19         | 8.0%          | 4         | 7.1%          | 4         | 9.8%          | 3              | 6.8%          | 5           | 7.6%          |
| Pick-up (small to large)        | 15         | 28         | 22         | 9.3%          | 7         | 12.5%         | 6         | 14.6%         | 5              | 11.4%         | 9           | 13.6%         |
| Passenger Car with Trailer      | 0          | 0          | 0          | 0.0%          | 0         | 0.0%          | 0         | 0.0%          | 0              | 0.0%          | 0           | 0.0%          |
| Pick-up with Trailer            | 0          | 0          | 0          | 0.0%          | 0         | 0.0%          | 0         | 0.0%          | 0              | 0.0%          | 0           | 0.0%          |
| Recreational Vehicle            | 0          | 0          | 0          | 0.0%          | 0         | 0.0%          | 0         | 0.0%          | 0              | 0.0%          | 0           | 0.0%          |
| Bus                             | 12         | 8          | 10         | 4.3%          | 0         | 0.0%          | 0         | 0.0%          | 1 <sup>a</sup> | 2.3%          | 1           | 1.5%          |
| Single-unit Truck               | 6          | 5          | 6          | 2.4%          | 1         | 1.8%          | 1         | 2.4%          | 2              | 4.5%          | 2           | 3.0%          |
| Semi-tractor with Trailer       | 0          | 0          | 0          | 0.0%          | 0         | 0.0%          | 0         | 0.0%          | 0              | 0.0%          | 0           | 0.0%          |
| Unknown                         | 17         | 0          | 9          | 3.7%          | 0         | 0.0%          | 0         | 0.0%          | 0              | 0.0%          | 0           | 0.0%          |
| <b>Total</b>                    | <b>200</b> | <b>264</b> | <b>232</b> | <b>100.0%</b> | <b>56</b> | <b>100.0%</b> | <b>41</b> | <b>100.0%</b> | <b>44</b>      | <b>100.0%</b> | <b>66</b>   | <b>100.0%</b> |

<sup>a</sup> this vehicle was a school bus

Generally, the vehicle type composition of the matched sample of vehicles were similar between the various data collection methods and the total vehicle population. However, the voice recognition method matched sample was closer overall to the true population composition than the keyboard entry method matched sample. In particular, the keyboard sample is under-represented in sedan passenger cars and over-represented in sport utility vehicles and pick-ups. In looking at the sedan passenger car statistics, the sample rates at the upstream location were 45.1% (73 vehicles) for keyboard entry and 41.9% (65 vehicles) for voice entry. At the downstream location, the sample rates were 39.9% (75 vehicles) for keyboard entry and 41.0% (89 vehicles) for voice entry. After filtering out incorrect entries, there were 60 lpc's correctly recorded with the keyboard entry method and 50 lpc's correctly recorded with the voice recognition method at the upstream location. At the downstream location, there were and 71 lpc's correctly recorded with the keyboard entry method and 80 lpc's correctly recorded with the voice recognition method. Although the entry accuracy rates for lpc vehicle plates were lower for every data collector than their average for all vehicles, they were considerably lower for both data collectors at the upstream location. The reason for this is not known. Upon inspection of the original data files, the difference in number of lpc's matched between the keyboard entry and voice recognition methods is basically random. While entry errors contributed to a reduced number of matches for both methods, there happened to be more instances of a particular vehicle that was seen at both locations not being entered at both locations by the keyboard method than the voice recognition method.

Some of the other percentage differences in Table 33 are more a result of changes in small numbers having a bigger effect percentage-wise, rather than some systematic bias or error. It should also be noted that since the data collectors did not record buses, the percentages would change slightly if buses were removed from the population distribution.

Given the instructions to only ignore buses for plate recording, data collectors had no obvious bias, or propensity, to sample any vehicles of a particular type disproportionately. As shown in chapter 7, there was no statistically significant

difference in the travel times obtained by any of the methods.. Because of the interrupted nature of the flow, and the slower free-flow speed, the vehicle composition is not a critical factor on this type of facility. Vehicles are much more likely to maintain their ordering and not try to improve their travel time by frequent lane changing. Thus, for this type of facility, the travel time would not be significantly affected even with significant deviations in the sample vehicle type proportions. However, extreme deviations in vehicle type proportions (e.g., sampling only trucks) may have a measurable effect. Thus, a sampling plan that specifically targets only a certain type of vehicle should be avoided. A statistical evaluation of the travel times of different vehicle classes for this data set would not be meaningful because the sample sizes of the different vehicle classes are too small.

For arterial facilities that are not interrupted by signals, there will likely be more sensitivity in the travel time measurement to deviations from the population vehicle type proportions. This is because there are more opportunities for vehicles to improve their travel time by changing lanes and passing other vehicles.

#### 8.2.1.2 Example Demonstration of Selection Bias on a Freeway

Table 34 lists the vehicle classifications utilized and several examples of the specific vehicles that were included under each vehicle type. The vehicle classifications were based upon general vehicle performance characteristics (e.g., acceleration, braking, maneuverability).

Table 34. Vehicle Classifications Key

| Vehicle Classification          | Vehicle Class. Code | Example Vehicles <sup>a</sup>  |
|---------------------------------|---------------------|--|
| Motorcycle                      | MC                  | Self explanatory   |
| Compact / Economy Passenger Car | CPC                 | Honda Civic, Honda CRX, Toyota Tercel, Toyota Corolla, Ford Escort, Mazda 323, Hyundai Accent, Nissan Sentra, Nissan Stanza, Geo Metro, Dodge Neon, Mitsubishi Mirage, Mitsubishi Eclipse, Subaru Justy, Volkswagen Golf |
| Luxury / Sedan Passenger Car    | LPC                 | Acura Integra, Acura Legend, Toyota Camry, Ford Taurus, Ford Probe, Toyota Paseo, Honda Accord, Honda Prelude, Nissan Altima, Nissan Maxima, Subaru Legacy, Mazda 626, Dodge Intrepid, Volkswagen Jetta, Buick Regal     |
| High Performance / Sports Car   | HPC                 | Chevrolet Corvette, Chevy Camaro, Ford Mustang, Pontiac TransAm, Mazda Miata, Porsche, Jaguar  |
| Sport Utility Vehicle           | SUV                 | Nissan Pathfinder, Toyota 4Runner, Ford Explorer, Ford Bronco, Chevrolet Blazer, Chevrolet Suburban, Isuzu Trooper, Isuzu Rodeo, Jeep Cherokee   |
| Mini-van                        | MV                  | Dodge Caravan, Ford Aerostar, Windstar, Volkswagen Vanagon, Mazda MPV, Dodge Caravan, Plymouth Voyager   |
| Pick-up (small to large)        | PU                  | Ford Ranger, Dodge Ram   |
| Passenger Car with Trailer      | PCT                 | Self explanatory   |
| Pick-up with Trailer            | PUT                 | Self explanatory   |
| Recreational Vehicle            | RV                  | Self explanatory   |
| Bus                             | BUS                 | Public transit, school buses, charter buses  |
| Single-unit Truck               | SUT                 | Delivery type vehicles   |
| Semi-tractor with Trailer       | STT                 | Self explanatory   |

<sup>a</sup> The example vehicles are not completely representative of the all the vehicle types entered under that classification.

Table 35 displays the vehicle composition percentages in the traffic stream (for the sample of matched vehicles).

**Table 35. Vehicle Classifications for Freeway Data**

| <b>Vehicle Classification</b>   | <b>GP Frequency</b> | <b>GP %</b>  | <b>HOV Frequency</b> | <b>HOV %</b> |
|---------------------------------|---------------------|--------------|----------------------|--------------|
| Motorcycle                      | 2                   | 0.1          | 35                   | 1.7          |
| Compact / Economy Passenger Car | 392                 | 19.0         | 337                  | 16.2         |
| Luxury / Sedan Passenger Car    | 843                 | 40.8         | 853                  | 40.9         |
| High Performance / Sports Car   | 23                  | 1.1          | 12                   | 0.6          |
| Sport Utility Vehicle           | 206                 | 10.0         | 212                  | 10.2         |
| Mini-van                        | 119                 | 5.8          | 225                  | 10.8         |
| Pick-up (small to large)        | 422                 | 20.4         | 296                  | 14.2         |
| Passenger Car with Trailer      | 1                   | 0.05         | 1                    | 0.05         |
| Pick-up with Trailer            | 0                   | 0.0          | 1                    | 0.05         |
| Recreational Vehicle            | 1                   | 0.05         | 2                    | 0.1          |
| Bus                             | 0                   | 0.0          | 84                   | 4.0          |
| Single-unit Truck               | 25                  | 1.2          | 26                   | 1.2          |
| Semi-tractor with Trailer       | 32                  | 1.5          | 0                    | 0.0          |
| <b>Total</b>                    | <b>2066</b>         | <b>100.0</b> | <b>2084</b>          | <b>100.0</b> |

Figure 32 contains plots of the travel times of several different vehicle types in the GP lane. Figure 33 is a similar plot for the HOV lane, but just for three significantly different vehicle types.

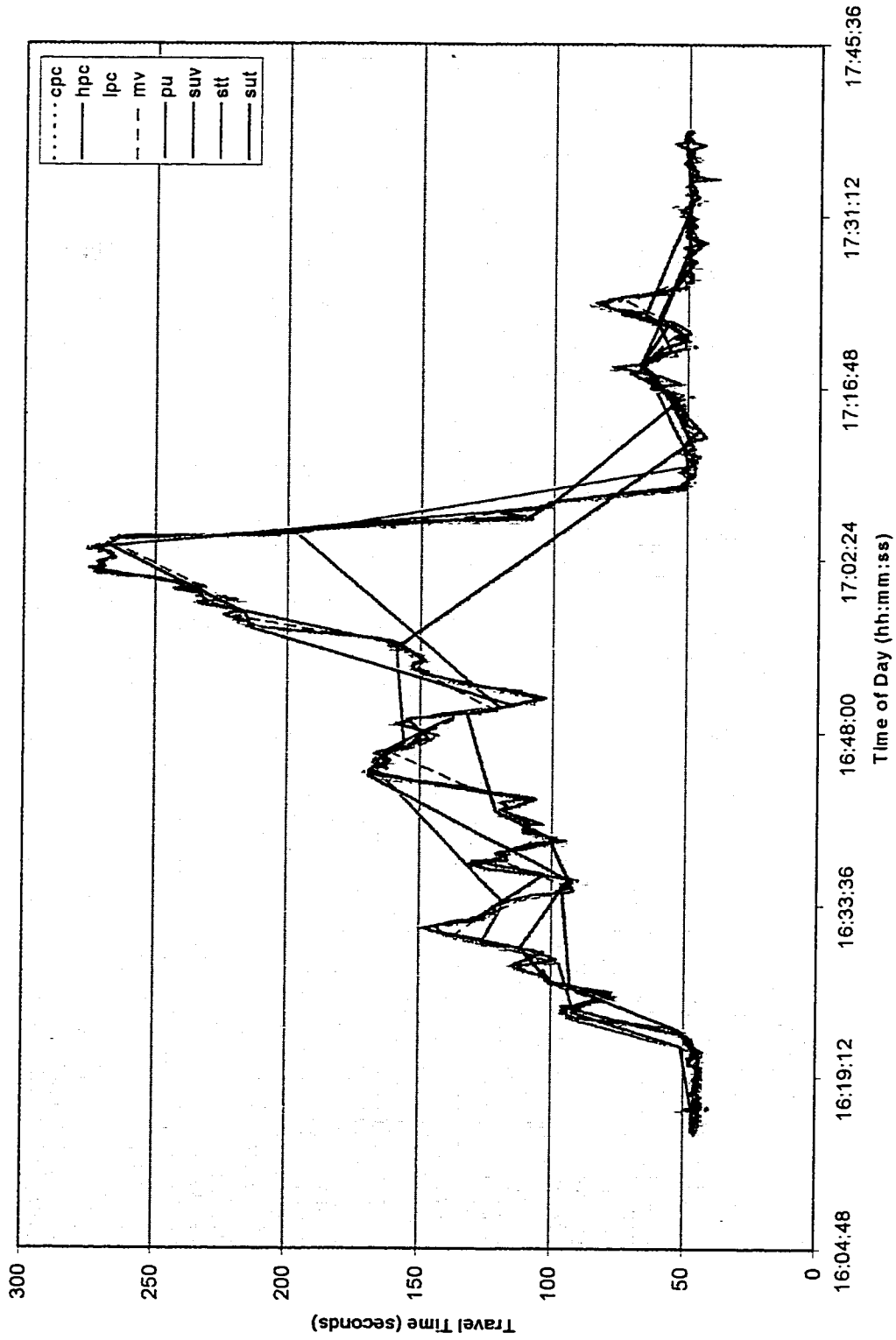


Figure 32. Travel Time Results by Vehicle Classification (GP Lane)

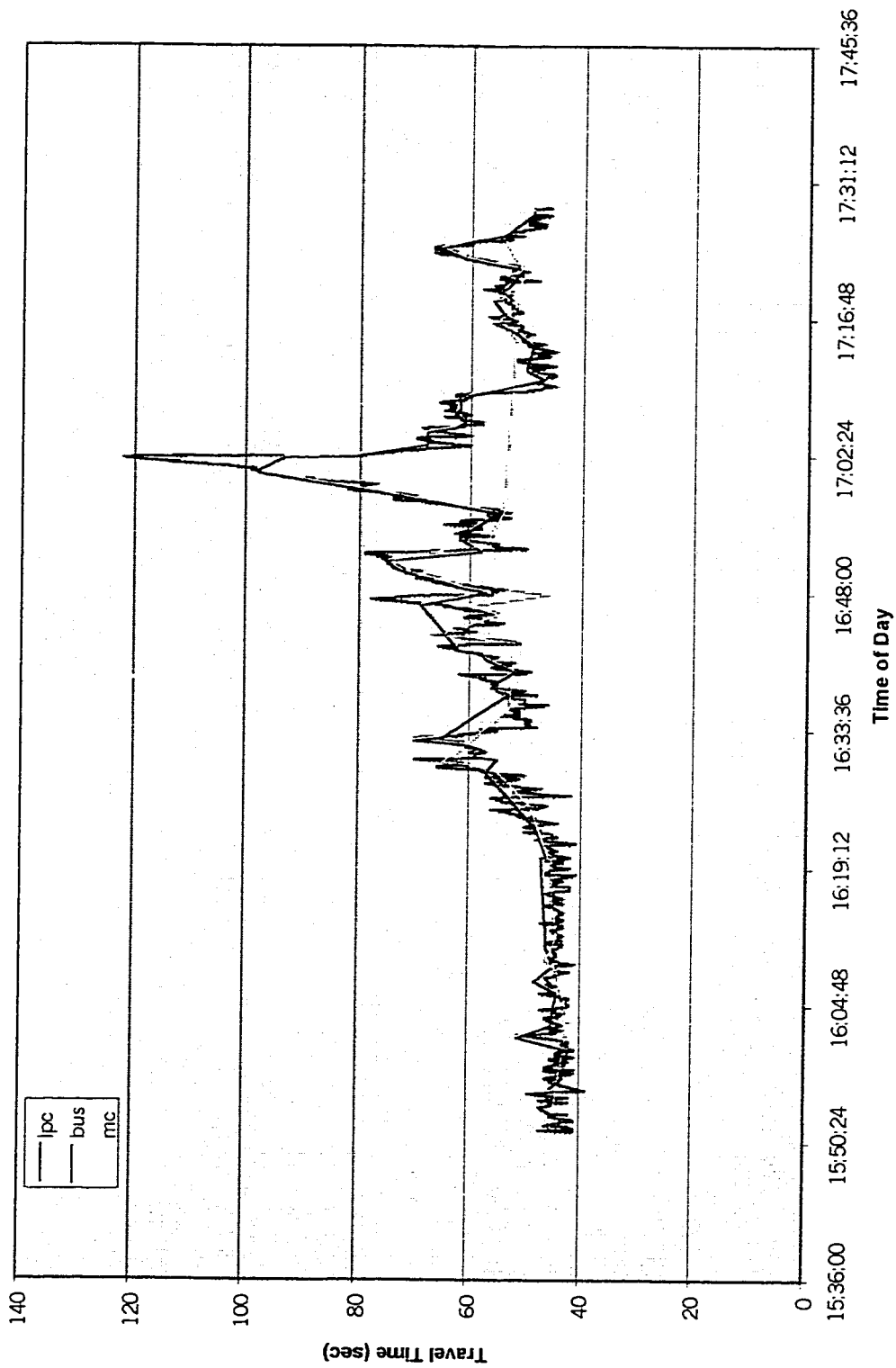


Figure 33. Travel Time Results by Vehicle Classification (HOV Lane)

It was suspected that during relatively non-congested flow conditions, there would be significant differences in average travel times between some different vehicle types in both the GP lane and the HOV lane. As has been previously discussed, as the volume demand continues to increase on a roadway section, there will be less opportunity to drive a desired speed; thus, speeds (and travel time) become more dependent under congested conditions. It was decided to compare some travel time statistics under relatively non-congested conditions, where drivers still have considerable freedom to choose their desired speed. The time period for which these statistics were taken was 3:51 – 4:21 PM. This time period ends several minutes before the onset of congestion begins.

**Table 36. Comparison of Speeds under Non-congested Conditions**

| Veh. Type    | General Purpose Lane <sup>a</sup> |                       |                      | HOV Lane       |                       |                      |
|--------------|-----------------------------------|-----------------------|----------------------|----------------|-----------------------|----------------------|
|              | No. of Samples                    | Avg. Trav. Time (sec) | Std. Dev. Trav. Time | No. of Samples | Avg. Trav. Time (sec) | Std. Dev. Trav. Time |
| cpc          | 75                                | 44.59                 | 1.86                 | 85             | 44.26                 | 1.73                 |
| lpc          | 173                               | 44.79                 | 1.81                 | 191            | 44.15                 | 1.81                 |
| hpc          | 7                                 | 45.57                 | 2.23                 | 2              | 43.5                  | 0.71                 |
| mv           | 18                                | 44.89                 | 2.19                 | 54             | 44.28                 | 1.62                 |
| pu           | 119                               | 44.76                 | 1.54                 | 96             | 44.45                 | 1.90                 |
| suv          | 50                                | 44.14                 | 1.74                 | 48             | 44.08                 | 1.80                 |
| sut          | 5                                 | 45.20                 | 1.30                 | 13             | 44.08                 | 1.75                 |
| stt          | 7                                 | 44.00                 | 1.83                 | 0              | 0                     | 0                    |
| bus          | 0                                 | 0                     | 0                    | 15             | 45.73                 | 2.19                 |
| <b>Total</b> |                                   | <b>44.685</b>         |                      |                | <b>44.276</b>         |                      |

<sup>a</sup> Due to the power outage for the GP lane camera, these data only represent 15 minutes during the ½ hour time period.

As it turns out, for these data, there are no statistically significant differences between any average travel times for different vehicle types on either lane. Although by inspection, one might suspect a significant difference between the average travel time of buses and lpc's, or buses and suv's in the HOV lane, for example, but this is not the case. Even by inspection of Figure 33, it would appear that the bus travel times are consistently higher than those for the lpc's. If by inspection, one chose to perform a two-sample t-test for these vehicle types, a statistically significant difference would in fact be identified.

However, by so doing, one will have contaminated the  $\alpha$  error value due to selection bias. Thus, the statistically correct method is to proceed with an F-Test (ANOVA) followed by pairwise comparisons using the Tukey-Kramer method. Proceeding in this manner showed that there were no significant differences between the travel times of any of the vehicle types for either lane. This result is just for this specific data set, and is not necessarily representative of all conditions. Since the GP lane data are from the inside lane, it is possible that this result is perfectly reasonable since there are usually fewer large trucks in this lane and most vehicles in this lane usually want to go faster than the average. In other words, different results certainly might be obtained for different lanes of the freeway, especially for the ones that do not consist primarily of self-selected vehicles that want to go faster than the “norm”. Further research on this topic is recommended.

Generally, though, it can be seen that the average travel times for all vehicle types are lower in the HOV lane than in the GP lane. This is also reflected in the overall average travel time for each lane.

What is evident is that one must be careful about selecting only certain types of vehicles within the traffic stream from which to measure travel time. Although differences in speed between vehicle types may not be a major factor, the frequency of certain vehicle types is definitely an issue. If one were considering equipping buses with AVI equipment, the frequency of buses within the traffic stream may not be sufficient to track the travel time for a roadway section throughout the entire spectrum.

### *8.2.2 Selection Bias Due to Variance in Vehicle Speeds*

The greatest potential for bias in travel time estimates due to over-representation of slower or faster vehicles exists under uninterrupted, low-flow traffic conditions. Under these conditions, individual vehicle speeds, and thus travel times are most likely to be independent of one another. Under interrupted flow conditions on arterials, or higher-flow conditions on freeways, this potential bias is diminished due to the dependence in travel speeds and times.

This effect may also be prevalent for video-based optical character recognition systems. That is, if the system accuracy is related to the number of video frames it “looks at” while a vehicle is within the FOV, then slower vehicles are likely to be recognized correctly more frequently than faster vehicles because they will be present in more frames than faster vehicles. This could bias the results toward a slower travel time than the true population travel time. A similar result may be possible with video image tracking systems, where vehicle identification accuracy is increased with increased “looks” at a vehicle within the field-of-view. This phenomenon did not appear to be present in the results of the Mobilizer tests discussed in chapter 6, but the tests were limited in flow conditions. This issue may surface during periods of transitional flow, from free-flow to stop-and-go, where there can be considerable variance in speeds from one minute to the next and across lanes.

Also, for manual data collection, an observer may unknowingly enter more license plates of slower vehicles because they have more time to view and assimilate the license plates of those vehicles. Given the likelihood of a conditional dependence between vehicle types and vehicle speeds (e.g., sports cars faster, commercial vehicles slower), it is important to be cognizant of these potential biases when developing a sampling plan.

### *8.2.3 Selection Bias Due to Over-sampling of Queued Vehicles*

This issue has some similarities to the previous issue, but is only an issue for interrupted flow facilities. Frequently, signalized intersections are used as data collection reference points. A wide range of traffic flow conditions can be observed at signalized intersections, depending on general congestion conditions and the state of the signal indication for any particular approach during any particular cycle. Flow conditions from free-flow, to stop-and-go, to just stopped (due to red signal indication) are frequently observed at signalized intersections during the peak periods.

Since it is possible for data collectors to record more vehicles when they are moving slower, such as when vehicles are starting up again after a red signal indication

turns to green, it is plausible that a data collector might record a larger proportion of license plates from vehicles having to stop for a red signal indication, or just starting from the queue at the start of green, than when the signal indication is green, except in the case of steady stop-and-go conditions. Thus, it is conceivable that an average travel time for an arterial could be weighted toward a higher travel time due to the incorporation of a larger sample that includes signal delay.

However, given that different sets of vehicles experience signal delay at different intersections within the study section, it is difficult to quantify this issue without more extensive simulation and/or experimentation along an entire corridor.

Through visual inspection of the NE 45<sup>th</sup> arterial data, it was determined that there was no obvious pattern that queued vehicles, or vehicles just starting from a queue, were sampled at any greater frequency than non-queued vehicles. The main reason for this is because the maximum sample rate (on the order of 2-3 seconds per vehicle) for the keyboard and voice recognition methods was about the same as the minimum headway of the vehicles.

### **8.3 Measurement Error Bias (Reference Point Time-Stamping)**

Travel time studies are based on specific locations of reference. For most studies, the distance between these locations is based on an upstream reference point and a downstream reference point. For travel time studies that rely on data collected by human observers from the field, the ability of the data collectors to enter and time-stamp the vehicle of interest at the exact point of reference is not a trivial issue. For shorter arterial sections, inaccuracies of time-stamping at the reference point can be significant. Of course, this might be mitigated by looking at longer sections, but then the loss of sample size due to ingress/egress points could become an issue, potentially requiring more data collection personnel at intermediate points.

There are a couple of key questions of interest. One, in a field setting, how much measurement error can be expected due to inaccuracy in time-stamping, and will that error bias the travel time measurements in a significant manner? And two, how are the

errors in time-stamp accuracy distributed? Are they distributed normally, or is there a systematic bias that can be expected in certain cases? If there is an unavoidable error present in the accuracy of the time-stamping, this must be accounted for in the formula for determining an adequate sample size. This naturally will require a larger sample size and thus more labor costs.

Also, the issues of sampling for greatest sample size and sampling for time-stamp accuracy are competing issues. That is, following the '*sample as many vehicles as possible*' strategy is likely to result in the greatest error and variance in time-stamp accuracy, whereas '*sampling for greatest time-stamp accuracy*' is likely to lead to the smallest sample size.

In order to sample for greatest time-stamp accuracy, it is necessary for the data collector to be sufficiently upstream of the chosen reference point to give them time to read and enter the license plate and then observe when the vehicle reaches the reference point and then time-stamp the entry. Inherent in this process is the fact that other vehicles will pass the data collector during the observation process of another vehicle which is about to be time-stamped. If the data collector time-stamps a license plate entry immediately, regardless of the vehicle's position relative to the reference location, the data collector will be able to enter other license plates during the aforementioned observation time.

At any given point along an arterial, a data collector can observe vehicle speeds ranging from free-flow to stop-and-go. Obviously, under stop-and-go conditions, it is much easier to not only time-stamp vehicles at a specific point, but also collect a 100 percent sample size, or close to. However, under free-flow conditions, sample size and time-stamping precision is very likely to be reduced. This entire range of flow conditions can be experienced at a signalized intersection within a short time interval.

Inaccuracies in time-stamp reference location are a result of several factors. The most significant is usually the platooning or headway characteristics of vehicles. This is because a vehicle behind the lead vehicle of the platoon cannot be observed until that vehicle gets close to the data collector, due to the small headways and the license plates

of following vehicles being occluded by leading vehicles. Another significant factor is the physical arrangement of the data collector relative to the reference location. If a data collector reads license plates from behind and then tries to enter the plate and time-stamp it before the vehicle reaches the reference location, this could lead to more errors because a faster moving vehicle will get to the reference location more quickly. A better strategy is to be upstream of the reference location and try to read the front license plates. Even if the data collector cannot read the plate before it passes, they can still try to read the rear plate. Generally, the front plate is better anyway since many vehicles have things like trailer hitches or bicycle racks that make it difficult to read rear license plates. However, not all states require a front license plate. Washington State law requires front mounted license plates, but even this is no guarantee that all vehicles will have one, as shown in Table 42 (Appendix D).

### *8.3.1 Example Demonstration of Measurement Error in Reference Point Time-Stamping*

Time-stamp accuracy data were obtained for many of the arterial tests described in chapter 5. Since specific pavement reference markings are not typically a luxury afforded for portable travel time studies, existing pavement markings (e.g., crosswalk stripe, raised pavement markers) visible within the camera FOV were utilized. The video camera time-stamp was synchronized with the portable computer used to simultaneously collect the license plate data. The time-stamp entries from either the keyboard input or voice recognition method were compared to the time-stamp entries of the corresponding vehicles on the videotape. If the videotape time-stamp is taken as the true time-stamp<sup>4</sup>, then there will be a  $\Delta t$  for each license plate observation from the computer entry.. Even though video data can be used with a very precise time base (e.g., tenths or seconds), there are some practical constraints that limit the precision of the time stamp in these data collection exercises. Since the video camera and the laptop computers are all

---

<sup>4</sup> The issue of time-stamping precision also exists for video (Kou and Machemehl, 1997). However, as discussed in chapter 6, the time-stamping of vehicles from the video is more highly controlled, and any errors are generally random in nature, not systematic.

synchronized together, they are usually done from a wristwatch with seconds. Although a stopwatch with more precision could be used, it is extremely difficult for a human to synchronize these devices to anything more precise than a second. An external electronic device that could be connected to the computers and cameras would be required for that kind of precision to be obtained.

Table 37 presents the timestamp accuracy results of the applicable data collection experiments. The values in parentheses are the results based on the absolute value of the timestamp errors. A positive timestamp value means that the data collector time-stamped the license plate entry after the vehicle crossed the reference point.

Table 37. Time Stamp Accuracy Results

| Data Collection Date <sup>a</sup> | Keyboard Entry              |                                  |                 | Voice Recognition           |                                  |                 |
|-----------------------------------|-----------------------------|----------------------------------|-----------------|-----------------------------|----------------------------------|-----------------|
|                                   | Avg. Time Stamp Diff. (sec) | Time Stamp Diff. Std. Dev. (sec) | Min / Max (sec) | Avg. Time Stamp Diff. (sec) | Time Stamp Diff. Std. Dev. (sec) | Min / Max (sec) |
| 5/8/98 <sup>b</sup>               | -0.495 (0.604)              | 0.656 (0.555)                    | -3/1            |                             |                                  |                 |
| 5/19/98 (1) <sup>b</sup>          | 0.309 (0.436)               | 0.674 (0.599)                    | -4/1            |                             |                                  |                 |
| 5/19/98 (2) <sup>b</sup>          | 1.377 (1.377)               | 0.720 (0.720)                    | 0/4             |                             |                                  |                 |
| 7/1/98 (1)                        | -1.750 (1.855)              | 1.135 (0.952)                    | -5/2            | 0.553 (0.913)               | 1.121 (0.851)                    | -3/3            |
| 7/1/98 (2)                        | -0.512 (0.976)              | 1.168 (0.818)                    | -4/4            |                             |                                  |                 |
| 7/9/98 (1)                        | 0.986 (1.189)               | 1.345 (1.168)                    | -3/8            | -0.186 (0.779)              | 1.165 (0.884)                    | -4/2            |
| 7/9/98 (2)                        | 1.422 (1.550)               | 1.048 (0.844)                    | -2/4            | 1.782 (1.806)               | 1.057 (1.016)                    | -1/5            |
| 7/23/98 (1,1)                     | 2.475 (2.710)               | 1.832 (1.460)                    | -6/8            | 2.558 (2.776)               | 1.811 (1.453)                    | -14/6           |
| 7/23/98 (1,2)                     | 3.904 (3.947)               | 1.361 (1.231)                    | -4/8            | 3.300 (3.300)               | 0.988 (0.988)                    | 1/8             |
| 7/23/98 (2,1) <sup>c</sup>        | 0.082 (1.945)               | 2.788 (1.992)                    | -19/9           | 2.931 (3.010)               | 1.329 (1.139)                    | -4/6            |
| 7/23/98 (2,2) <sup>c</sup>        | 2.672 (3.317)               | 3.881 (3.343)                    | -44/7           | 3.879 (3.879)               | 1.482 (1.482)                    | 1/16            |
| 8/12/98 (1)                       |                             |                                  |                 | 1.471 (1.471)               | 0.668 (0.668)                    | 0/3             |
| 8/12/98 (1)                       |                             |                                  |                 | 0.763 (0.784)               | 0.642 (0.616)                    | -1/2            |
| 8/12/98 (2)                       | 2.725 (2.725)               | 0.746 (0.746)                    | 1/4             | 3.920 (3.920)               | 0.761 (0.761)                    | 2/6             |
| 8/12/98 (3)                       | 0.644 (1.055)               | 1.072 (0.664)                    | -3/3            | 0.670 (0.780)               | 0.775 (0.663)                    | -1/2            |
| 8/12/98 (4)                       | 0.346 (0.500)               | 0.764 (0.672)                    | -1/3            |                             |                                  |                 |
| 8/12/98 (4)                       | 1.228 (1.333)               | 0.945 (0.787)                    | -1/3            |                             |                                  |                 |
| 9/3/98 (1,1)                      | 0.587 (1.826)               | 3.229 (2.722)                    | -24/4           | -0.064 (1.954)              | 3.703 (3.140)                    | -22/3           |
| 9/3/98 (1,2)                      | 2.104 (2.119)               | 1.186 (1.159)                    | -1/7            | 3.011 (3.129)               | 1.414 (1.127)                    | -11/6           |
| 9/3/98 (2,2) <sup>d</sup>         | 1.287 (1.287)               | 0.746 (0.746)                    | 0/4             | 1.023 (1.023)               | 0.857 (0.857)                    | 0/4             |

<sup>a</sup> The first number in parentheses is the data collection session. The second number in parentheses is the data collection station, 1 for upstream, 2 for downstream.

<sup>b</sup> These data collection experiments were done for the specific objective of obtaining maximum time stamp accuracy.

<sup>c</sup> The voice recognition entry scheme used the manual post-license plate entry timestamp technique for the purpose of trying to obtain more accurate time stamps.

<sup>d</sup> The voice recognition entry scheme used the manual pre -license plate entry timestamp technique for the purpose of trying to obtain more accurate time stamps.

What can be discerned from the numbers in the above table is that a relatively wide range of average timestamp errors are possible (-1.75 to 3.92), and there is no distinguishable pattern to the errors. However, the majority of the average timestamp errors were positive in value. For every data collection effort, data collection personnel tried to position themselves such that they could time their license plate entry with vehicles crossing the reference point. Determining this timing is difficult—it is a function of vehicle speed and the entry time of the data collector. Positioning observers a conservatively large distance upstream of the reference point is also not a reasonable alternative. The further upstream an observer is from the reference point, the more difficult it is for the observer to see the reference point and exactly when a vehicle crosses it. Furthermore, from a sample rate perspective, it is not desirable to have data collectors watching vehicles travel downstream towards the reference point when they could be reading the license plates of upstream vehicles.

As indicated by the bold numbers, there were several instances in which a vehicle was time-stamped much sooner than when it crossed the reference point. This is a result of stop-and-go traffic, either for through vehicles in congestion, or for right-turn-on-red vehicles. It was confirmed through review of the video, that the large negative timestamps were a result of a data collector time-stamping a plate entry just as a vehicle reached the reference point, but then the vehicle was delayed at that point. For vehicles in congested traffic, it is usually difficult for a data collector to know when a vehicle arriving at the reference point will get delayed at that point, since the data collectors is usually focused on reading and entering license plates more than the traffic flow condition at any particular moment. Right turning vehicles that approach the intersection while the signal turns from green to yellow will often perform a slowdown to a brief stop and then check for opposing traffic and then proceed to make the turn if there are no conflicts. Many times, however, they will arrive a second too late, and opposing traffic flow has already started; thus, they have to wait. But oftentimes they continue to creep forward until they get a gap in the traffic stream and then continue with their turn. While performing the data collection, it can be difficult to know that a right-turning vehicle is

slowing down for a signal change or just slowing down to make the turn, as most drivers do. Thus, a data collector may enter and time-stamp a right turning vehicle as they approach the stop line, but then the vehicle may be delayed at that point for many more seconds. If the data are being collected such that signal delay is being incorporated, this is an important consideration. If signal delay is not being considered, than this issue is irrelevant, as you want to timestamp the vehicle as it arrives at the intersection, rather than as it leaves. While it is extremely difficult to prevent these kind of timestamp errors altogether, they should be infrequent. Additionally, their impact is minimized because not all the vehicles with this kind of timestamp error will be matched at another location.

Interestingly, using the manual post-license plate entry technique for the voice recognition method actually resulted in worse timestamp errors than comparable experiments using just the automatic post-license plate entry timestamp technique. However, as shown in Table 4, the manual timestamp technique resulted in considerably lower sample rates than the automatic timestamp technique. The manual pre-license plate entry timestamp technique for the voice recognition method did result in a relatively low average timestamp error (1.02 seconds) and a narrow error range (0 to 4 seconds). This is only one data set, but intuitively, this technique should offer the most accuracy in time-stamping because the observer is not even concerned with the license plate entry until they press a key (or use a voice keyword) time-stamping a vehicle crossing the reference point. Thus, the order of emphasis is reversed—timestamp first, then plate entry, instead of plate entry and then timestamp. But again, this results in a considerably lower sample rate.

To determine the true impact of these timestamp errors, and whether a particular timestamp entry technique is justified over another, one must look at the combined effects of more than one data collection station. It is possible that the timestamp errors from one station to the next will offset each other. Or they may combine for large total travel time errors. Table 38 presents the travel time results from two of the arterial data collection sessions. The table contains the travel time results based on the original data for each method, and the travel time results based on the timestamp corrected values for the keyboard entry and voice recognition methods.

Table 38. Timestamp Corrected Arterial Travel Times

|                                     | Session 1 (7/23/98) |          |                   |           |                   |       | Session 3 (9/3/98) |                   |           |                   |  |  |
|-------------------------------------|---------------------|----------|-------------------|-----------|-------------------|-------|--------------------|-------------------|-----------|-------------------|--|--|
|                                     | Video               | Keyboard |                   | Voice Rec |                   | Video | Keyboard           |                   | Voice Rec |                   |  |  |
|                                     |                     | Orig.    | Adj. <sup>a</sup> | Orig.     | Adj. <sup>a</sup> |       | Orig.              | Adj. <sup>a</sup> | Orig.     | Adj. <sup>a</sup> |  |  |
| Mean Travel Time (sec)              | 179.6               | 181.7    | 180.0             | 177.2     | 176.2             | 432.6 | 437.9              | 435.2             | 436.9     | 433.9             |  |  |
| Std. Deviation of Travel Time (sec) | 43.7                | 42.8     | 42.9              | 41.0      | 41.4              | 116.3 | 126.2              | 124.1             | 138.5     | 139.8             |  |  |
| Maximum Travel Time (sec)           | 264                 | 257      | 257               | 257       | 256               | 644   | 641                | 640               | 687       | 685               |  |  |
| Minimum Travel Time (sec)           | 107                 | 108      | 107               | 109       | 107               | 268   | 267                | 266               | 245       | 239               |  |  |
| Median Travel Time (sec)            | 182                 | 182      | 181               | 178.5     | 179               | 411   | 455                | 458.5             | 413       | 409               |  |  |
| Mode of Travel Time (sec)           | 179                 | 179      | 210               | 177       | 175               | 271   | 530                | 269               | 643       | 269               |  |  |

<sup>a</sup> These values have been adjusted for the timestamp errors. The timestamp error is based on the difference between the manual method timestamp and the video timestamp.

In both sessions, for both manual data collection techniques, the timestamp errors combined to make the travel time estimate higher than the true travel time.

It is also interesting to note how sensitive the mode value is. For distributions that exhibit peaks other than just at the mean, reporting a single mode value can be very misleading.

A data collection experiment was conducted for the purpose of comparing the data sets obtained when trying to obtain very accurate time stamps and when trying to collect the largest sample size possible, without regard for accurate time stamping. This comparison experiment was conducted twice, with the two data collectors switching entry methods for the second session. These were done at the same location, with the same traffic conditions—free-flow, heavy platooning. It should be noted that all data entry was performed using the keyboard. Table 39 summarizes the data collection statistics. Figure 34 and Figure 35 display the frequencies of the timestamp errors for the two data collection strategies for each data collection session.

In the first data collection session, the average timestamp was clearly more accurate when trying to enter vehicles for timestamp accuracy versus sample size. The average timestamp error was considerably lower and the variance of timestamp errors was less. However, the average timestamp error for the timestamp accuracy technique in the second session was higher than any of the other data collection sets. For both data collection sessions, there was a clear sample size advantage for the entry strategy that was attempting to get the maximum sample size by disregarding timestamp accuracy. The entry strategy did not make a significant difference in the entry accuracy rate.

Table 39. Comparison of Maximum Sample Size versus Time Stamp Accuracy

| Data Collection Date <sup>a</sup> | Time Period (min) | Flow Rate (vphpl) | Sample Rate (%) | Accuracy Rate (%) | Avg. Time Stamp Diff. (sec) <sup>b</sup> | Time Stamp Diff. Std. Dev. (sec) <sup>b</sup> | Min / Max (sec) |
|-----------------------------------|-------------------|-------------------|-----------------|-------------------|--|---|-----------------|
| 5/19/98 (0,1)                     | 15                | 880               | 72.6            | 96.9              | 1.259 (0.958)                            | 1.264 (0.150)                                 | 1/5             |
| 5/19/98 (1,2)                     | 15                | 880               | 50.0            | 96.3              | 0.309 (0.436)                            | 0.674 (0.599)                                 | -4/1            |
| 5/19/98 (2,1)                     | 15                | 936               | 66.7            | 98.0              | 0.282 (0.731)                            | 0.962 (0.685)                                 | 2/3             |
| 5/19/98 (2,2)                     | 15                | 936               | 52.1            | 98.3              | 1.377 (1.377)                            | 0.720 (0.720)                                 | 0/4             |

<sup>a</sup> The first number in parentheses is the data collection session. The second number indicates whether the data collection was done for maximum sample size (1), or for time-stamp accuracy (2).

<sup>b</sup> The number in parentheses is based on the absolute value of the timestamp error.

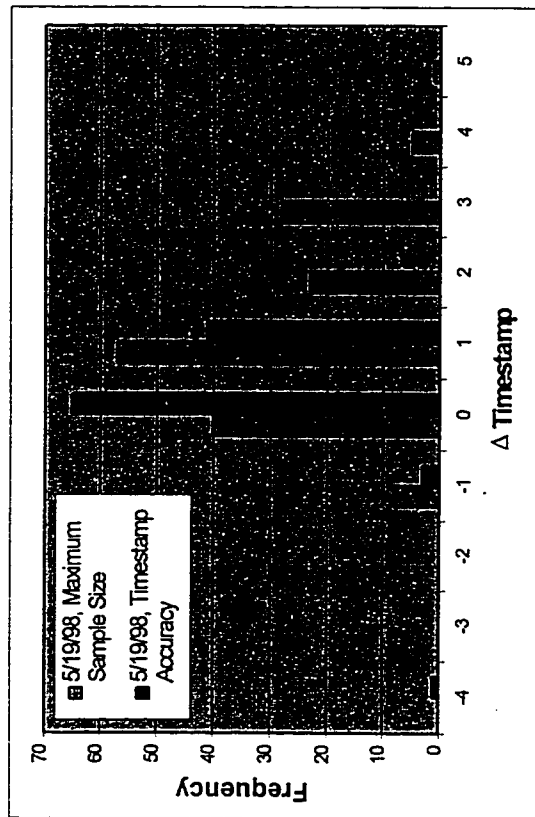


Figure 34. Timestamp Accuracy Comparison (Session 1)

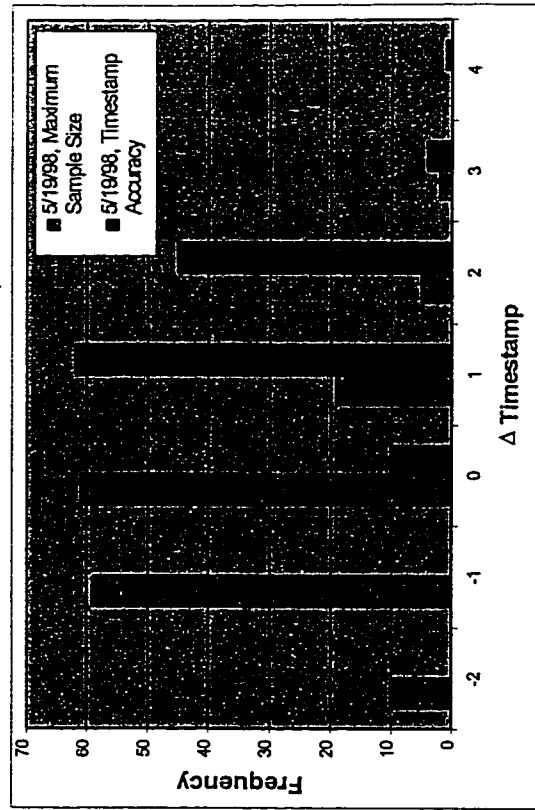
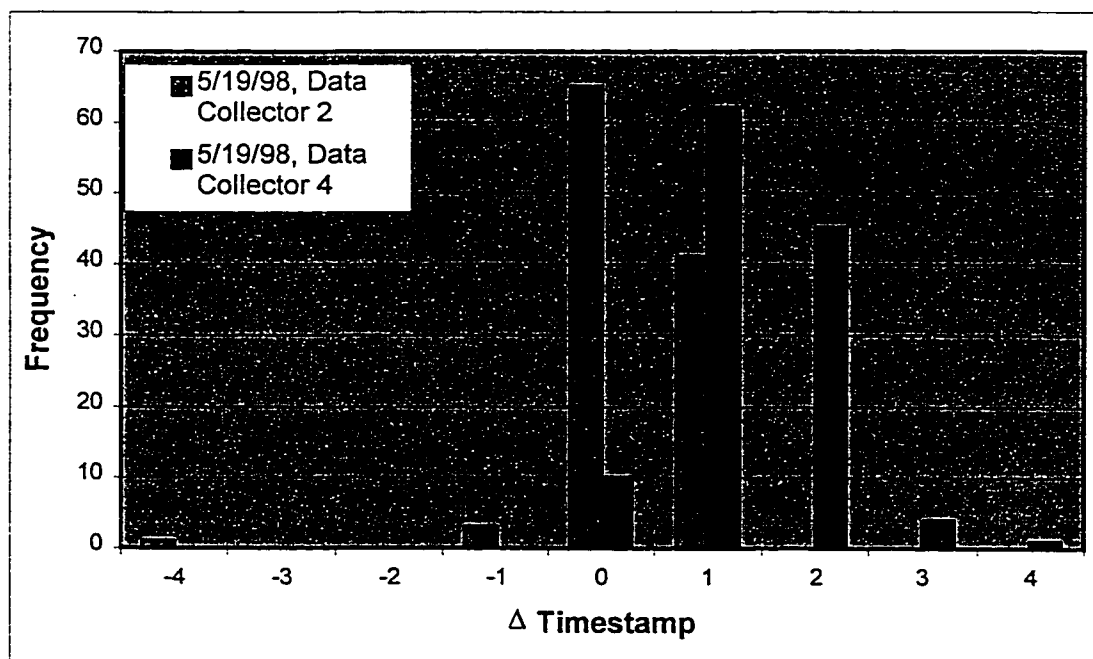


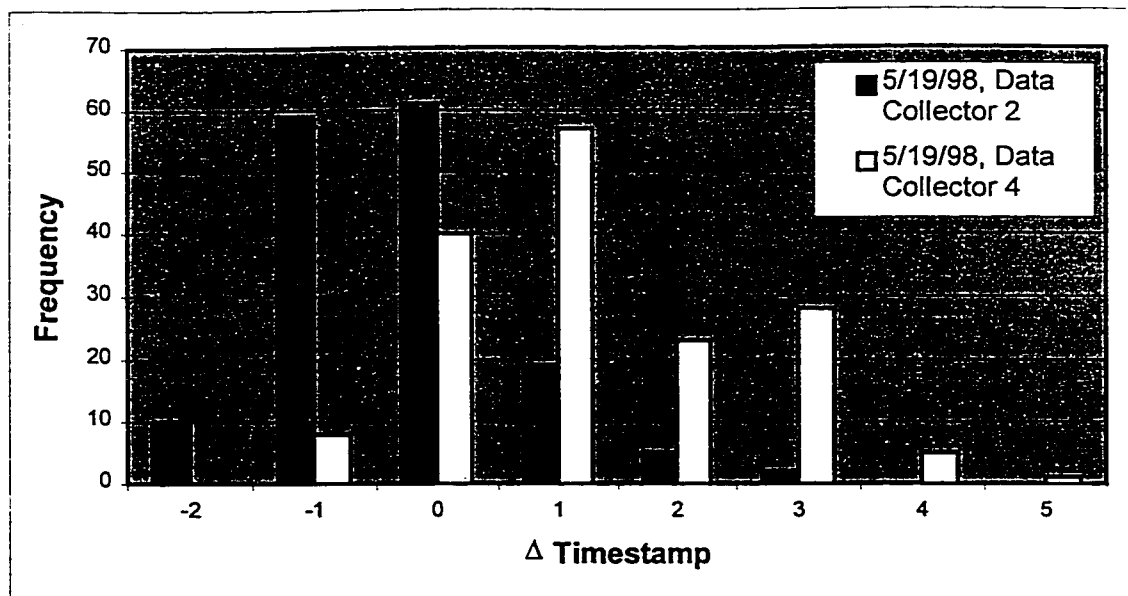
Figure 35. Timestamp Accuracy Comparison (Session 2)

Measurement errors in an experimental setting typically follow a normal distribution. One might expect a similar distribution for deviations of time-stamp entry about the true time-stamp entry point. However, even if it is a normal distribution, it may not be centered about the true time-stamp. It may be skewed either toward late time-stamps or early time-stamps. Typically, though, the distribution will be skewed one way or the other. Again, operator skill plays an important role in the quality and accuracy of the entries.

The following two figures compare the timestamp error distributions between data collectors for the same type of timestamp entry scheme for free-flow, heavily platooned, arterial traffic.



**Figure 36. Timestamp Error Distributions (Timestamp Accuracy Technique)**



**Figure 37. Timestamp Error Distributions (Maximum Sample Size Technique)**

As the results in Table 38 indicated, the cumulative effect of the timestamp errors was a travel time measurement greater than the true travel time. The next two figures illustrate this result. In both figures it can be seen that the downstream timestamp errors are distributed towards larger values more than the upstream timestamp errors. This results in the larger than true travel times.

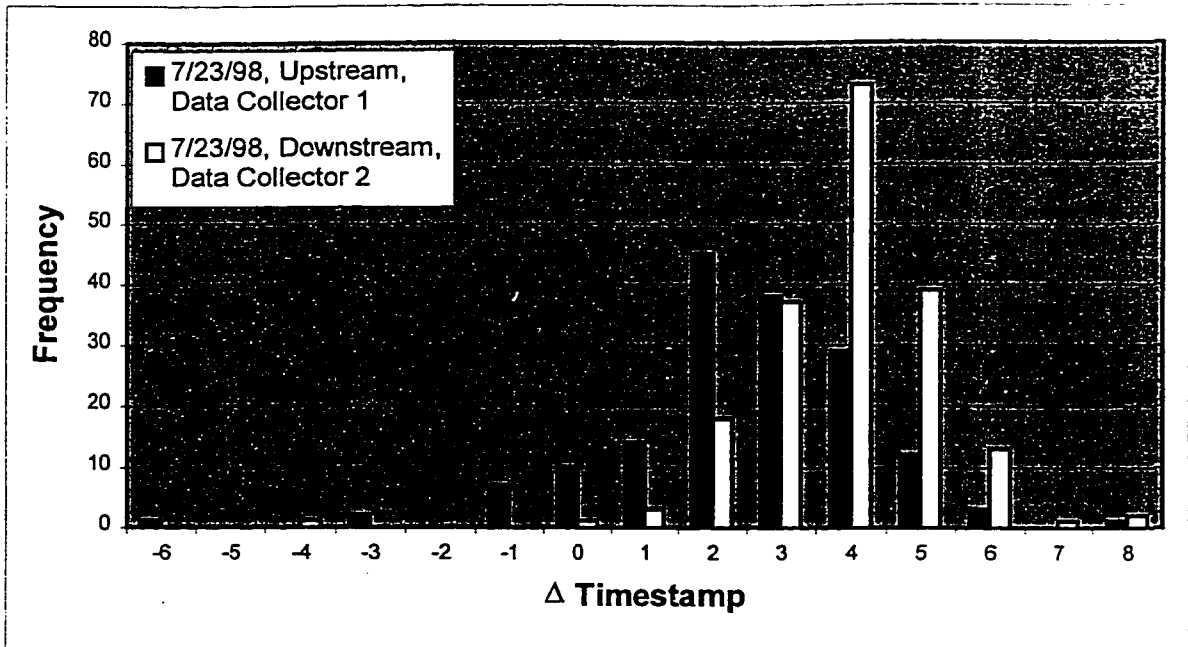


Figure 38. Timestamp Error Distributions (Session 1)

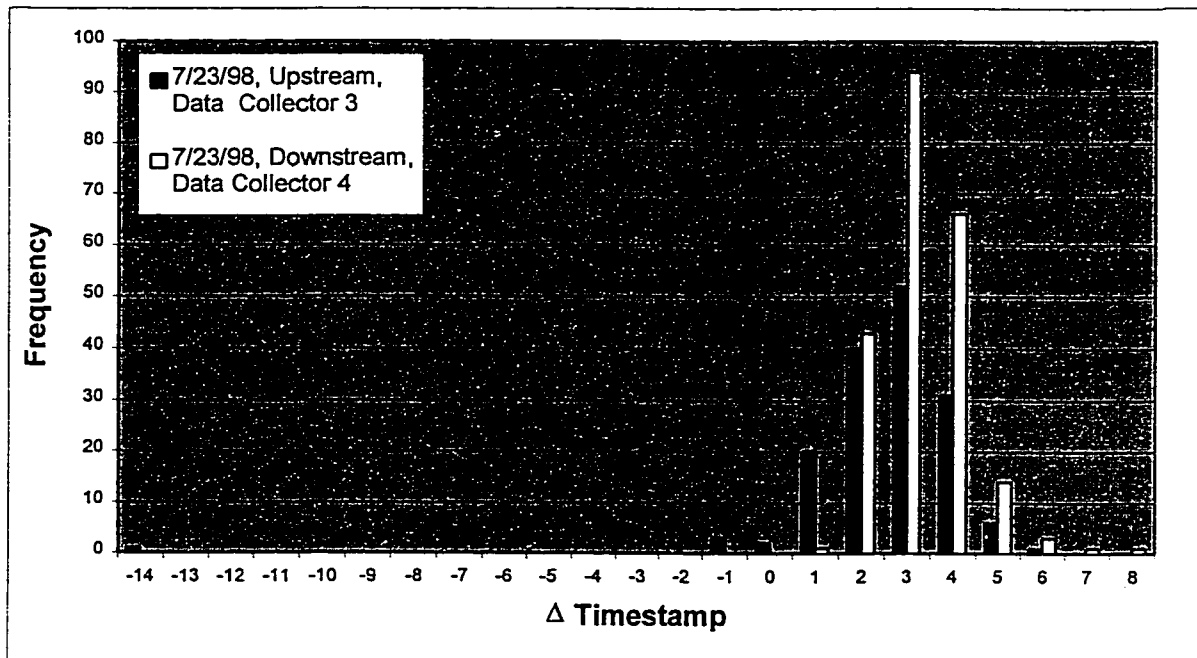


Figure 39. Timestamp Error Distributions (Session 2)

## 8.4 Aggregation Bias

This type of bias is associated with the time interval that is chosen for data summarizing and/or reporting. Ideally, one would like to summarize the travel time statistics over as large a time interval as possible that does not result in gross inaccuracies at smaller time intervals. Aggregation has the effect of “smoothing” the data. To some extent, smoothing is desirable, especially for Advanced Traffic Information Systems (ATIS) purposes. It does not make any sense to report travel times to commuters on a 30-second or even 1-minute basis. That is just too much information for the average commuter to “process”. The commuter is more interested in the general condition than the potentially highly variable condition at any given moment. For the purposes of ATIS, a smaller number of reporting intervals will result in less demand on the DOT computing and electronic resources and less (frequent) information for commuters to “process”.

Figure 40 is a simple illustration of various levels of travel time measurement aggregation for the freeway data.

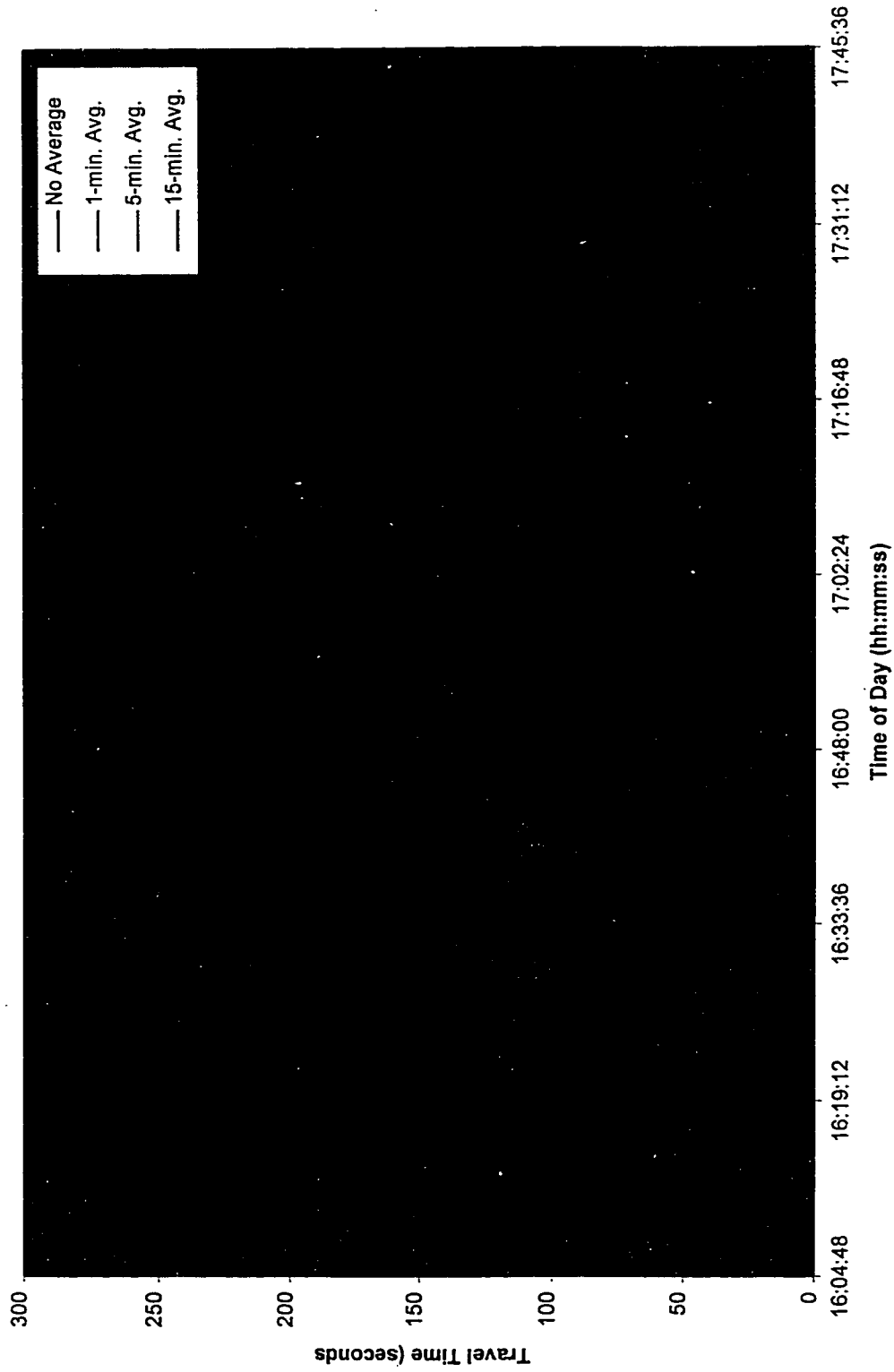


Figure 40. GP Lane Travel Time with various aggregation levels

As can be seen, the 1-minute aggregation level provides a very good “fit” with the raw travel time data. Going to an aggregation level of 5-minutes still seems to provide a good representation of the overall travel time, while still capturing the general peaks and valleys throughout the entire study period. Going to an aggregation level of 15-minutes does not follow the trend of the travel time data, other than just indicating the travel time increases to a peak, and then decreases again. Additionally, the traffic conditions can change completely in the time it takes to aggregate and report data for 15-minute intervals. For example, if a commuter were to get the travel time update at about 5:15 PM for this section/route, they would see that traffic is still fairly congested, when in fact it has returned to nearly free-flow already. The use of a 5-minute reporting period was found appropriate for a study measuring travel time via AVI [Turner and Holdener, 1995].

## **8.5 Error Due to Outliers**

Another potential source of significant error in travel time estimates is a data analysis that inaccurately accounts for outlying data points, either through erroneous inclusion of true outliers, or erroneous exclusion of valid data points. Again, ad-hoc statistical procedures are often used to identify and eliminate outliers, usually based on the judgement of the data collector. However, since this process is usually not very straightforward, there is generally much room for error. The collected arterial and freeway travel time data were examined for outliers, and appropriate analysis techniques are discussed.

### **8.5.1 Arterials**

Since arterials usually have many signals and many driveways, the probability of obtaining outlying data points in a travel time data collection are usually high. As a result, it is desired to have a technique whereby outliers can be readily identified and dealt with appropriately. However, given the nature of point sampling, where information is only gathered at discrete points and not continuously only the route, a

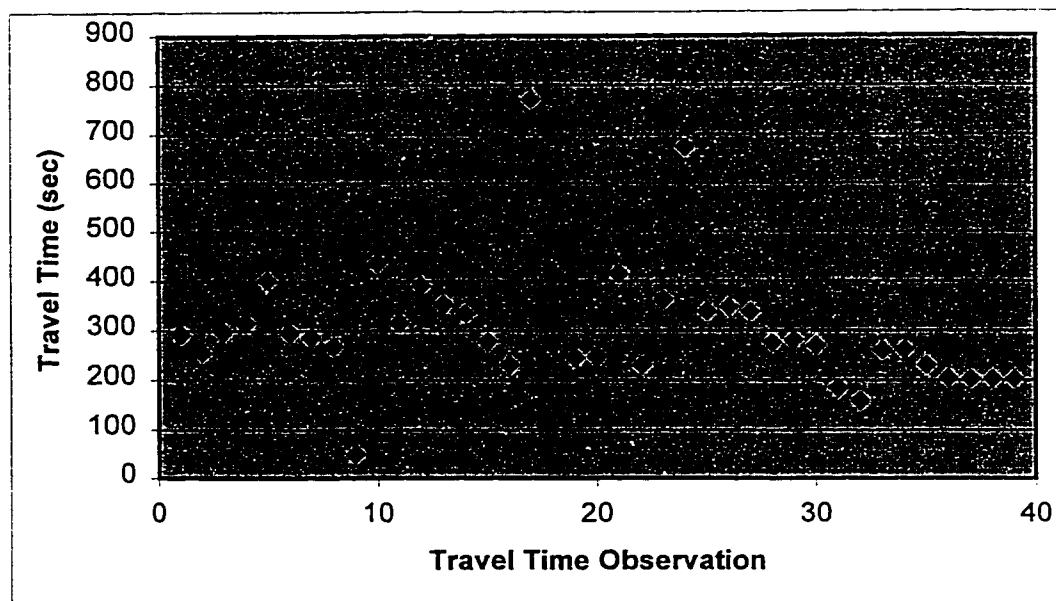
universal technique (i.e., one technique fits all approach) for this purpose does not exist. On busy arterials with many signals and driveways, the treatment of outliers is not necessarily straightforward, especially if data collection points are spaced far apart. Situations in which vehicles miss out on progression opportunities, or vehicles temporarily leave the arterial (e.g., espresso stand) can certainly cloud the true travel time for that arterial section. Many retail and commercial sites that rely heavily on pass-by traffic (e.g., gas stations, fast-food restaurants, etc.) are potentially large contributors to the presence of outliers in an arterial travel time data set. Transit buses also pose a potential outlier problem with the frequent stops made along a roadway facility.

#### 8.5.1.1 Example Demonstration of Outliers for Arterial Data

As it turns out, this particular arterial study section provided very few outlying data points. Although the frontage property to this roadway is heavily developed with commercial and retail property, there are few land uses of the quick visit type (e.g., fast food restaurants, gas station, convenience market, etc.). Several methods have been suggested for identifying outliers, but one of the best and quickest methods is just by visual observation. Especially for data sets with only a few outliers, this can be the most efficient method.

As an example, the keyboard entry data set for the second data collection session had three travel time measurements that were identified as outliers. Plotting the individual travel time data points makes identification of these outliers fairly easy. Figure 41 is a plot of individual travel time observations versus travel time. The observations are in sequential order. Observation 9 (50 seconds), observation 17 (772 seconds), and observation 24 (674 seconds) were identified as outliers. Observation 9 is most likely an outlier because of the extremely low travel time and observations 17 and 24 are most likely outliers because of extremely high travel times. All three of these values fall well outside of the range of all the other measurements (158 – 436 seconds). Cross-referencing with the video data identified observation 9 as a spurious match (i.e., different vehicle with same first four license plate characters). The other two were the

same vehicles at both recording stations, but must have pulled off the arterial for a short period of time.



**Figure 41. Travel Time Plot for Identifying Outliers**

A couple of other data sets contained some negative travel time values. These are obviously easy to identify as outliers. For these data sets, the outliers were easy to identify just through visual inspection. However, sometimes a data set may be obtained that contains a high proportion of outliers. In this case, visual inspection alone will probably not be sufficient. A popular type of plot that is much more comprehensive and robust than the one above is the boxplot. A boxplot identifies the data center, spread, extent and nature of any departure from symmetry, and identification of outliers using spread measures. Based on the boxplot method, values below 88.75 or above 486.75 would be considered mild outliers. The boxplot method uses a measure called the fourth spread, which is the median of the largest half of the observations minus the median of the largest half of the observations, to identify outliers. This fourth spread measure is resistant to outliers.

Another suggested method of dealing with outliers that does not even involve identifying outliers is to use the median value of travel time. This method, however, is

undesirable because there is no variance measure associated with the median, which is needed to evaluate the sample size adequacy.

Another method would be to establish lower and upper bounds for travel time based upon driving experience of the route or some other local knowledge. This method has the drawback that setting these upper and lower thresholds is subjective. Additionally, if they are not adjusted for each data collection, they are likely to be invalid for many data sets due to the variance experienced in traffic conditions over different days.

If outlying points are a very small percentage of the total data points, graphical methods will probably be sufficient for correctly identifying the outliers. For data sets with significant percentages of potential outliers, more rigorous statistical methods should be applied as well. Methods that utilize spread or quartile values are likely to be robust and resistant to outlier influences.

### **8.5.2 Freeways**

Due to the limited access nature of freeways, and absence of traffic signals, the treatment of outlying statistical points in travel time data is usually less complicated. The primary cause of outliers in this instance is the use of partial license plate numbers for entering into the data collection software. This occasionally results in the matching of different vehicles with the same subset of partial license plate characters, referred to as spurious matches.

#### **8.5.2.1 Example Demonstration of Outliers for Freeway Data**

Table 30 from chapter 7 has been reprinted below, for the matching done assuming 100% entry accuracy. This table illustrates the issue of outliers due to spurious matches from partial plate character entries.

**Table 40. Freeway Simulation Spurious Matches**

| <b>Plate Characters Entered, Entry Accuracy Rate</b> | <b>Total Matches</b> | <b>Spurious Matches</b> | <b>% Spurious Matches</b> | <b>Correct Matches</b> |
|--|----------------------|-------------------------|---------------------------|------------------------|
| All  | 1799                 | 20                      | 1.1                       | 1779                   |
| Chars 1-4, 100%                                      | 350                  | 45                      | 12.7                      | 305                    |
| Chars 3-6, 100%                                      | 321                  | 16                      | 5.1                       | 305                    |
| Chars 1-3, 5, 100%                                   | 309                  | 10                      | 3.2                       | 299                    |
| Chars 1-3, 6, 100%                                   | 318                  | 16                      | 5.0                       | 302                    |
| Chars 1-3, 100% <sup>b</sup>                         | 767                  | 451                     | 58.8                      | 316                    |
| Chars 4-6, 100% <sup>b</sup>                         | 494                  | 207                     | 41.9                      | 287                    |
| Chars 1-5, 100% <sup>b</sup>                         | 294                  | 6                       | 2.0                       | 288                    |
| Chars 2-6, 100% <sup>b</sup>                         | 309                  | 2                       | 0.6                       | 307                    |

For these data sets, the identification of outliers was straightforward because they were either negative in value or much larger than the “adjacent” travel times.

## 8.6 Conclusions and Recommendations

When collecting data in the field for travel time measurement, there are several potential sources of bias that can become incorporated into the travel time estimate. Selection bias and timestamp errors are the primary sources. For arterial travel time measurement, selection bias, either through differences in sampled vehicle type proportions, over-sampling of faster or slower vehicles, or over-sampling of queued vehicles was not observed in any systematic manner, and definitely not to an extent where it significantly influenced any travel time results. Sampling as many vehicles as possible, and not targeting any specific vehicle types will typically result in the statistics being insensitive to minor deviations in the sample distribution from the population distribution.

Timestamp error, however, was a significant issue. Through all of the data collection experiments, it was found that it is extremely difficult to control for timestamp errors, regardless of the timestamp technique. The range of average timestamp errors was from -1.75 to 3.92 seconds, a range of almost six seconds. For the combined effect of the timestamp errors for NE 45<sup>th</sup> Street travel time data collections, the range of travel time

errors ranged from 1.0 to 3.0 seconds. Essentially, this practically unavoidable error should be accounted for in your confidence limits and/or sample size requirements. The timestamp error distributions varied substantially, but were distributed more frequently towards positive timestamp errors, meaning the vehicle was time-stamped after it crossed the reference point. It is interesting that no particular timestamp entry technique resulted in consistently lower timestamp errors. Thus, the timestamp techniques aimed at collecting more accurate timestamps do not appear warranted, particularly when the sample size effects are taken into account. A strategy that aims to collect as large a sample size as possible, while still trying to obtain reasonably accurate timestamps will generally result in a data set with more statistical validity because of the much greater sample size.

Outliers were not a significant issue for the arterial travel time data sets collected. Further research on arterial travel time data with many outliers due to vehicles making short stops at roadside businesses is recommended.

For freeway data, travel times were measured for the HOV lane and the inside GP lane. While significant differences in travel time between different vehicle types for free-flow conditions were not found for either lane, data from other freeway lanes should be examined to determine if this situation holds for the other lanes as well. It is well known that average speeds vary across different lanes, but much less is known about how speeds vary within those specific lanes. Under the forced flow nature of congested traffic flow, it is unlikely that speeds will vary significantly by vehicle type. But it is possible that this may not be the case for free-flow conditions and even transitional flow conditions for other freeway lanes. It is likely that the specific use patterns of the HOV lane and inside GP lane result in a greater level of homogenization of either vehicle type or driver type, or both, that contributes to small variances in the average speeds.

For ATIS reporting purposes, the use of 5-minute aggregation levels for reporting travel time is appropriate. The 5-minute aggregation level is the largest interval that should be considered for other travel time uses. The 1-minute aggregation level is

probably the most appropriate for operational control purposes. The use of a 2 or 3-minute aggregation level is a good compromise for other travel time uses.

Outliers in freeway travel time data are primarily a result of spurious matches due to incomplete license plate entries. These are usually easily identified visually through graphical display methods because they are often either negative or many times larger than "adjacent" travel time measurements.

## CHAPTER 9

# GUIDELINES FOR TRAVEL TIME DATA COLLECTION

---

### 9.1 Introduction

Considerable experience with travel time data collection was gained through the research conducted for this document. Additionally, there are numerous other experiences with travel time data collection contained in the body of literature. However, one thing that is generally lacking in all of this literature is a set of guidelines that a practitioner looking to do a travel time study can consult. Thus, the purpose of this chapter is to develop a set of practical guidelines to consider when undertaking a travel time study, based upon the knowledge and experiences of this author. The material in this chapter will also be supplemented with relevant references to other available literature. In particular, a recent publication, *'The Travel Time Data Collection Handbook'* [TTI, 1998] provides a large quantity of useful information regarding travel time data collection. Another relevant publication is NCHRP Report 398, *'Quantifying Congestion'*. A good deal of information from this publication was incorporated into the former publication. The latter publication emphasizes issues and procedures to consider for congestion management. The emphasis of this chapter is strictly upon travel time measurement, which is just one component of congestion management. In many cases, the information provided from this author is more specific than that contained in the other publications. For other areas that this author does not have much direct experience, other relevant publications will be cited, if they exist. It should be pointed out that this set of guidelines is not intended to address every possible question that could arise in any type of travel time data collection effort. Furthermore, this set of guidelines is meant to be primarily qualitative in nature, providing a practitioner a relatively comprehensive set of questions to ask when undertaking a travel time study. The primary quantitative issues are raised in this set of guidelines, but the focus is on giving the inexperienced practitioner a thorough list of issues to be considered.

The travel time data collection effort is divided into four main areas: Planning the travel time study, developing the data collection plan, conducting the study, and data analysis and reporting. The main questions for each area will be listed and then each question will be discussed in more detail in the following sections.

## 9.2 Planning the Travel Time Study

Before developing a data collection effort, several questions must be addressed to determine exactly what the study is going to accomplish.

- What are the characteristics of the study facility?
- What is the purpose of the travel time study?
- What is the appropriate (or necessary) level of accuracy for the chosen parameters?
- What travel time parameters (e.g., mean, mode, standard deviation, range, etc) or information should be reported?

### 9.2.1 *What are the characteristics of the study facility?*

As the study planner, it is essential that you already are or become familiar with the study site. The following items at a minimum should be known.

- Number of cross streets or exits and entrances,
- Number of lanes,
- Number of signals, and
- Suitable vehicle observation locations (for license plate matching).

How these characteristics should be taken into account are discussed later in section 9.3.3.

### 9.2.2 *What is the purpose of the travel time study?*

The first question that needs to be asked is “What is the purpose of the travel time study?” There are many purposes for travel time measurements, but the major ones are: 1) for use as an MOE in comparing before-and-after facility modifications; 2) reporting to the commuting public to help them make travel decisions; 3) general facility operational

evaluation for planning purposes; and 4) for use in real-time traffic control algorithms (e.g., ramp metering rates). The specific data needs and tolerance levels vary for each of these purposes. Once this fundamental question is answered, all remaining questions can be more readily addressed. The first thing answering this question will make clear is what type of travel time measurement “installation” is appropriate—permanent or portable.

A permanent (long term continual measurement) data collection installation is more appropriate for the needs of traffic management centers. A portable (short term or immediate evaluation) data collection installation is more appropriate for the following:

- Scheduled occasional measurements at selected times of the year for transportation infrastructure monitoring for regional planning organizations, and
- Studies done for evaluation of a specific facility over a limited period of time.

The needs of these two types of travel time data collection are very different. Generally, permanent data collection will only be performed by public agencies. Normally only automated methods of data collection will be appropriate for permanent data collection. A permanent data collection installation provides the opportunity to implement the most resource efficient data collection technique that also fulfills all requirements. For portable studies, the practitioner is often faced with developing and executing a data collection plan without a thorough knowledge of the resources that will be required to obtain the statistical results required for the project.

Unfortunately, cost is usually the most significant variable in performing a travel time study. Cost is manifested in all aspects of a data collection effort—personnel time (data collection, data reduction), equipment purchase or rental, etc. However, it is also vital that the practitioner have an understanding of the tradeoffs associated with the reliability of the results for changes in the data collection plan.

### 9.2.3 What is an appropriate level of accuracy for the study of interest (i.e., How accurate does it really need to be)?

The ITE *Manual of Transportation Engineering Studies* [1994] recommends the following absolute errors for different types of studies (also outlined in NCHRP 398, Lomax, et al. 1997):

- Transportation planning and highway needs studies:  $\pm 3$  to  $\pm 5$  mph
- Traffic operations, trend analysis, and economic evaluations:  $\pm 2$  to  $\pm 4$  mph
- Before-and-after studies:  $\pm 1$  to  $\pm 3$  mph

Indiscriminant application of these values is not recommended. Instead, it is recommended that a percentage error be applied to the measure of interest. Use of the above constant values can result in misleading accuracy or confidence level results. As an example, data from one arterial travel time measurement discussed in chapter 7 will be examined. Translating the average travel time into an average travel speed yields 8.6 mph (for all combined 66 matches). Using the ITE guidelines for traffic operations of  $\pm 2$  to  $\pm 4$  mph as an example, it is obvious just from inspection that this range would be totally inadequate. If we assume 8.6 mph to be the true average travel speed, the percentage errors range from 18.9% (for +2 mph) to 86.9% (for - 4 mph), as calculated from the following formulas.

$$\frac{2150 * 3600}{(8.6) * 5280} = 170.5 \text{ seconds} \quad (\text{travel time for section at 8.6 mph})$$

$$\frac{2150 * 3600}{(10.6) * 5280} = 138.3 \text{ seconds} \quad (\text{travel time for section at 10.6 mph})$$

$$\frac{170.5 - 138.3}{170.5} = 18.9\%$$

$$\frac{2150 * 3600}{(4.6) * 5280} = 318.7 \text{ seconds} \quad (\text{travel time for section at 4.6 mph})$$

$$\frac{318.7 - 170.5}{170.5} = 86.9\%$$

Thus, percentage error values would be a better approach. The following are recommended percentage error levels for various types of travel time measurement.

- **Highest Accuracy Level (maximum error  $\leq$  5%)**
  - Travel time data implementation real-time traffic operations control strategies
  - Validation testing of indirectly derived measurements
  - MOE for before-and-after study
- **Medium Accuracy Level (5% < maximum error  $\leq$  15%)**
  - Dissemination of data to real-time traveler information systems
  - MOE for transportation system improvement
  - Calibration of simulation model
- **Lowest Accuracy Level (15% < maximum error  $\leq$  20%)**
  - Planning Study

These percentage values are consistent with what many other researchers have utilized in travel time studies. But again, it should be emphasized that these values should be carefully considered when planning a study.

#### *9.2.4 What travel time parameters (e.g., mean, mode, standard deviation, range, etc) or information should be reported?*

For standard statistical reporting, the focus is typically the measures of central tendency—mean, median, mode, variance, and standard deviation. These values can be computed for any sample of data. For travel time measurements taken on transportation facilities, the question arises as to which measures are the most appropriate to use in reporting travel time measurement results. There are advantages and disadvantages for any reported measurement.

For ATIS applications, commuters are probably more likely to understand the meaning of a reported average travel time than they are a reported variance measure.

However, many may still not fully understand the meaning of an average. They may interpret an average as being their own predicted time of travel, rather than realizing that some times will be less than this, and some times more than this. For signalized arterials, particularly coordinated systems, an average value has even less significance due to the cyclical nature of flow through those facilities. Typically, on this kind of facility, a driver will experience either good progression or poor progression. Thus, their travel time will most likely be above the calculated average or below the calculated average. There may not be many vehicles that actually travel the corridor in the calculated average amount of time. The median travel time does not offer any advantages over the mean in this situation. Two measures that may be more appropriate for signalized arterial facilities are the mode and range. The mode is a measure of the travel time occurring most frequently during the analysis period. For signalized arterials, it may be appropriate to report the two or three most frequently occurring travel times, due to the likelihood of a multi-modal travel time distribution. As indicated in chapter 7, there may be a mode associated with low travel time (good progression) and a mode associated with high travel time (poor progression), and maybe even a mode associated with the average travel time. The range reports the lowest occurring travel time and the highest occurring travel time during the analysis period. Thus, this measure gives a picture of what range of travel times are possible. In general, the range value is probably more understandable to a wider audience than a standard deviation or variance measure.

For uninterrupted flow facilities, the mean and median are still very useful measures. The key is to choose the appropriate aggregation interval for which to report the mean or median value, due to the usual rapidly changing flow conditions on these facilities during the peak periods. As was discussed in chapters 7 and 8, the aggregation interval has a large influence on how representative the reported travel time measures are of the traffic conditions occurring at any specific time. For peak periods, it is recommended to not use aggregation intervals any greater than five minutes in length. For off-peak periods, 10 or 15 minutes may be adequate aggregation intervals. The use of the median value instead of the mean can sometimes be more appropriate if the data

contain a large number of outliers. The presence of many outliers can skew the mean value one way or the other. The median value will always report the center point of the distribution. The median value does have the drawback that a corresponding standard deviation cannot be computed. This issue will be re-addressed in section 9.5.1.

### 9.3 Developing the Data Collection Plan

- What techniques/methods will allow me to obtain the previously determined data needs?
- What are the primary considerations for the data collection techniques?
- How do I ensure a statistically valid study with the chosen technique/method?

#### 9.3.1 *What techniques/methods will allow me to obtain the previously determined data needs?*

The type of study will often dictate the necessary method, or at least narrow the options down considerably. Most available technologies are discussed in Chapter 2. Some additional technologies are discussed in the TTI report (The Travel Time Data Collection Handbook). The most common methods are listed again in Table 41.

**Table 41. Common Travel Time Measurement Techniques**

| <b>Automatic</b>                          | <b>Manual</b>  |
|---|--|
| Probe Vehicle – AVI / AVL                 | Probe Vehicle – Floating Car   |
| Video License Plate Character Recognition | License Plate Matching – pencil/paper, voice recorder, computer, video |
| Video Image Vehicle Tracking              |  |

For more detailed descriptions about these methods, the reader is referred to chapter 2 of this document and the *Travel Time Data Collection Handbook*. It is worth repeating something pointed out in the latter reference: “Travel times collected using different data collection techniques may yield slightly different results. Data managers and users should recognize the potential biases inherent in the technique(s) they are using.” The reader is referred to chapter 8 again for discussion about the various potential sources of bias and the techniques where they may be most prevalent.

Additional guidelines for planning a data collection effort can be found in '*License Plate Matching Surveys: Practical Issues and Statistical Considerations*' [Schaefer, 1988]. This paper presents some specific considerations for license plate matching studies based upon pencil-and-paper license plate recording.

### *9.3.2 What are the primary considerations for data collection techniques?*

#### 9.3.2.1 Data Recording Location Constraints

Arterial facilities offer the most flexibility in terms of suitable data collection methods. Route planning is simpler for floating car studies, roadside locations for license plate observation are more plentiful, and signal poles offer mounting appurtenances for automated equipment methods.

Freeways pose more difficulty because of the general lack of safe and convenient roadside observation locations. Thus, overpasses are usually the only reasonable observation location option.

#### 9.3.2.2 Equipment Resources

The most common tools for manual forms of travel time data collection are video cameras, portable computers, and voice recorders. Since laptop computers have become quite affordable, this may be a resource that is easily available. However, one must be careful to check the hardware requirements necessary to run any specific data collection software. This may reduce the number of computers available for use. While video cameras are also quite common these days, attention must also be paid as to whether the available cameras have the necessary operating specifications to be used for a travel time study. For example, if the cameras are going to be used for recording license plates, they must have the necessary zoom and shutter speed specifications.

Equipment cost and/or availability may be a major constraint. Generally, the less money spent on data collection equipment, the smaller the collected data sample will be. Additionally, certain forms of inexpensive data collection equipment (e.g., pencil-and-paper) could prove to be just as expensive as other means when data reduction is

considered. That is, when you consider the time required to transcribe handwritten data into an electronic format, the overall cost of the data collection effort could be expensive.

It should be noted that power supplies tend to be the Achilles' heel of data collection efforts utilizing electronic equipment. It is strongly recommended that you obtain large power supplies (e.g., car battery). Also be sure to check software operation for each computer before going into the field.

### 9.3.2.3 Personnel Resources

The number of people you have available to perform data collection has a very direct bearing on the type of technology that can be used and what kind of results can be obtained. Schaefer lists several good personnel considerations for license plate matching studies. Keep in mind that when it comes to data collection personnel, you often get what you pay for. In other words, if you hire inexpensive, relatively unskilled labor, the quality of your results are likely to reflect this.

- **Training**—This issue has a large amount of variability. More advanced technologies will generally require more training time. For manual license plate sampling techniques, data collection personnel will need to be trained on the sampling scheme to be employed, and they will need training on the variety of different license plate formats that will be encountered. Additionally, personnel should be instructed on how to respond to curious passer's-by.
- **Data Reduction**—Non-computer based data collection methods will require substantially more data reduction time and expense. While video can provide the richest data collection set, if it is processed manually it will require the most time and expense. However, depending on the needs of the study, shortcuts can be employed to reduce the effort. For example, you can implement a sampling scheme which extracts every 3<sup>rd</sup> or 4<sup>th</sup> vehicle from the video. You can then run the match program on those entries to see if you get statistically desired results. If not, go back and add vehicles to the sample.

Rerun the match program. Check for an adequate number of matches again. Repeat this process as needed. Or you can enter all license plates of the upstream station, and then run the tape for the downstream station and enter only those plates that match with the upstream station. Of course, this may not work very well if you have intermediate data collection stations. Also, some level of quality control is probably still necessary as some vehicles may have a more difficult-to-read license plate at different stations.

#### 9.3.2.4 Costs

Costs are a function of the required equipment (hardware/software, supplies), the personnel (number, training, field time), and the data reduction/analysis time. The Travel *'Time Data Collection Handbook'* provides some cost guidelines on many of the mature techniques. As mentioned in chapter 6, cost guidelines are not yet available for the video image tracking equipment.

Costs for the voice recognition method are similar to those for the standard portable computer keyboard entry method. The only additional expense is the purchase of a high-quality microphone/headset. One area that can result in a substantial savings for the voice recognition method, however, is when binoculars are required to read license plates. One person can still collect a large sample size with the voice recognition method while using binoculars. In many instances, two data collection personnel are necessary when using the keyboard input method and binoculars are needed.

When comparing video-based data collection to portable computer-based data collection, keep in mind that the long run cost effectiveness of the equipment will probably not be as high for video. Although, video cameras are comparable in cost to laptop computers, laptops are usable in many day-to-day activities, while the cameras are more limited in their applications.

#### *9.3.3 How do I ensure a statistically valid study?*

This can prove to be the most challenging aspect of the entire study. The biggest issue that practitioners are faced with is that of the optimal allocation of expensive

resources for data collection. There are many tradeoffs associated with the allocation of resources and the quality of the data collected. One does not want to spend more on the data collection than is necessary, but one also wants to be assured that what data is collected will result in a statistically valid study.

In the simplest terms, to obtain statistical validity, it is necessary to collect enough travel time measurements such that the true travel time distribution for vehicles on that facility for the chosen time period can be determined with a reasonable degree of confidence. 'Reasonable degree of confidence' is of course subjective, but commonly accepted values are usually 90 and 95 percent. Confidence levels below 85 percent should be avoided.

The most dominating variable is the variance, or alternatively, standard deviation. Obviously, traffic streams with little variability in travel time over a given time period for a specific facility will need fewer samples of travel time for one to be reasonably confident of an average travel time, or other statistical measures. Uninterrupted-flow facilities (e.g., freeways) are likely to exhibit the least variability in travel time for small time intervals. Obviously, for large time intervals (e.g., one hour), the variability can be quite large. But for time intervals of five minutes and smaller, the variability is likely to be relatively small. Interrupted-flow facilities such as signalized arterials are likely to exhibit the largest variability in travel time during a specific time period, particularly the peak period. Since travel time formulas rely on knowing at least the variance of the population distribution, it is difficult to determine the necessary sample size before conducting the study. A limited floating-car study may provide a rough estimate of this value to use in the sample size formula, but it could also be very inaccurate. In addition, performing this kind of pre-study data collection uses potentially valuable resources—either time and/or money. Although travel time is correlated with several other traffic flow variables, such as speed, lane occupancy, and volume, speed is the only variable for which the relationship is direct. Of course, speed in itself is not a universally collected measurement. Since the infrastructure of many facilities relies on single loop detector stations, speed cannot be measured directly from these locations. If dual-loop

installations were extensive throughout a freeway corridor, for example, an estimate in speed variability could probably be used as a proxy for travel time variability, and thus a sample size estimate could be made with more precision. Neither volume nor lane occupancy, by itself, is a very reliable surrogate measure of travel time variability.

Aside from the travel time, the other factors that influence necessary sample size are the desired level of confidence and the allowable error. As the desired level of confidence increases, or the allowable error decreases, the required sample size will increase. Of course, the desired confidence level and allowable error should be specified before the data collection effort commences.

Although sample size is the main factor that drives the confidence level of the travel time estimate, other sources of potential bias, as discussed in chapter 8, must also be considered. These include selection bias and measurement error. Controlling for the main potential sources of bias during the study will aid in the accuracy of the measured travel times.

Additionally, it is important to consider the fact that things often do not go exactly as planned. Therefore, it is advisable to not construct the data collection plan such that there is no margin for error. For example, if a computer battery dies with ten minutes left in the data collection period, have a backup plan such that an inadequate data set for statistical validity, or a completely wasted effort, does not result. Other things like data collectors getting off to slow starts, or interruptions from passer's-by should be considered.

Besides the variation inherent during a particular time period on a particular day, there is also variation on a daily and seasonal basis, generally resulting from the constantly fluctuating demand observed on all roadway facilities throughout the year. To justify the expense of an arterial signal timing project, an agency may decide to undertake a before-and-after study. Therefore, it is very important that one can compare the difference in before-and-after travel time without the random variation in travel demand during these time periods masking any potential effects from the signal timing project. Specific guidelines on the number of days to collect data are difficult to give, because

they are facility and location dependent. Nonetheless, it is necessary to factor this consideration into this type of study. Chapter 2 of the '*Travel Time Data Collection Handbook*' offers some general suggestions in this regard. At the very least, one should consider performing before-and-after data collection during the same time of year to factor out seasonal traffic demand effects. Additionally, potential traffic growth between the before-and-after data collection must be accounted for.

The question of an appropriate sampling strategy has to be considered in the context of the purpose of the measurement(s). Generally, facility evaluations for long-range planning purposes are not going to require nearly as high a statistical accuracy as for measuring the impacts of facility modifications in before-and-after studies. The primary factors that influence the obtainable sample size during a data collection effort are described below.

#### 9.3.3.1 Factors affecting potential sample size

##### ***Facility Characteristics***

- *Access frequency* (i.e., on/off-ramps, intersections): This affects the amount of traffic entering/exiting the traffic stream between the endpoints of the study section. This will have a direct impact on the size of the match pool and the resulting match frequency and/or the number of necessary personnel.
- *Signal control level*: Aside from being present at intersections which impact access frequency, they introduce delay to the traffic stream, which adds to the variability of the travel time.
- *Number of travel lanes*: A decision must be made as to how to deal with facilities with multiple lanes in the same travel direction. Since different lanes of the same facility usually have different flow/speed characteristics, and thus, different travel times, it must be decided whether there will be a separate data collector for each lane or whether a more complicated sampling plan will be incorporated for data collectors collecting data from more than one lane of traffic. Under light traffic conditions, it may be possible to make due with one

data collector for multiple lanes, but for peak period conditions, a separate data collector for each lane of traffic is the prudent approach. Lane changing is another factor that convolutes the distinction of separate travel times for separate lanes.

### ***Traffic Stream Characteristics***

- *Flow Range*: Volumes at higher speeds will inherently be more difficult to obtain a large sample size from. Stop-and-go traffic will be the easiest to obtain large sample sizes from.
- *Level of platooning*: Heavy platooning will result in a lower sampling rate than that for uniform flow. This is because it is only possible to sample a fraction of the vehicles within the platoon, and then there will be gaps between platoons in which few vehicles arrive. If this same demand is spread out uniformly in its arrival rate, the data collector will collect a larger sample because the arrival rate will be more similar to the sampling rate capability of the data collector.

### ***Sampling Characteristics***

- *Manual LPM from Video Data*: With this method, it is possible to sample 100% of the traffic stream at each location. However, it is unlikely that a 100% sample proportion will be achieved because of missing license plates (see **appendix D**), pedestrians blocking the camera field-of-view, and/or license plates not caught within the camera's field-of-view (e.g., vehicle tailgating).
- *LPM with Portable Computers*: As previously discussed, the sample proportion is dependent on the number of plate characters entered, which is inversely related to the number of spurious matches. Again, the choice of either the front or rear license plate will affect the potential sample proportion. Under stop-and-go traffic flow conditions, or very uncongested low flow

conditions, it is possible to obtain an extremely high proportion of the vehicles in the sample. Under other flow conditions, the sample size proportion will usually be lower, with higher speed, higher flow conditions usually resulting in the lowest sample size proportion. Heavy platooning will also lead to reduced sample rates.

- *Video Imaging Tracking*: As reported in chapter 6, travel time match rates can vary between 10 and 20 percent for relatively uncongested flow conditions.

One must avoid the temptation to just use sample size values reported in other references without consideration of the specific factors that are present at the study site in question. For example, the ITE *Manual of Transportation Engineering Studies* makes the statement, “Analysts can obtain reliable, unbiased estimates of volumes and travel-times with a minimum of six test runs in each direction under comparable conditions (Mortimer, 1957)” in referring to the moving vehicle method. In another example, The Chicago Area Transportation Study (CATS) reported that 26 license plate matches were necessary for statistical confidence during a given period [Bailey and Rawling, 1991]. It is certainly easy to key in on these numbers as being applicable to your study just because you are using the same data collection method, but the characteristics of each site, as discussed above, will dictate the necessary sample size more than the method itself.

#### **9.4 Conducting the Study**

Once the data collection plan has been formulated, the next step is the execution of that plan. This section will offer some practical guidelines and insights on ways to make the execution of the data collection effort in the field go smoothly. The major components of the data collection exercise consist of the following:

- Equipment/personnel setup in the field, and
- Minimizing sources of error and/or bias during the data collection.

### 9.41. Field Setup

- *Manual LPM from Video Data:* Camera positioning is obviously the critical issue for this type of data collection. If unable to mount the camera directly over the traffic stream, it will be necessary to place the camera at ground level, and probably on a sidewalk or shoulder. The primary concern here is that the camera be positioned so as to minimize field-of-view (FOV) interference from pedestrians and/or bicyclists. Having as narrow an angle as possible between the camera FOV and the vehicle license plates is ideal. Aside from adjusting the zoom and shutter speed to the appropriate levels, the choice of a lens filter may be important (e.g., polarizing lens for reducing sun glare off of plates).
- *LPM with Portable Computers:* Here the most important consideration is the placement of personnel. Observers need to be positioned where they have a clear, unobstructed view of license plates. Additionally, the data collectors will need to be comfortable for an extended period; thus, a folding chair is recommended. Although this is usually difficult for arterial studies, it is recommended that as discrete a location as possible be used, so as to avoid driver distractions. As previously mentioned, a sign that can be displayed indicating that a traffic study is in progress and that the data collectors should not be disturbed is a good idea for locations with pedestrian traffic. No matter how uninteresting you may find the data collection effort, many passers-by cannot seem to resist the urge to quiz you about all the details of what you are doing and why.
- *Video Image Tracking:* As described in detail in appendix B, the data collection process is critical to the success of this kind of system. The camera placement, camera attributes, and field-of-view distance calibration are all significant factors.
- *General Considerations:* Besides exercising standard safety guidelines for field work, the use of two-way radios (a.k.a., walkie-talkies) can be invaluable during data collection exercises. For distances under 2 miles, they can be very

reliable. They take less time to establish communications between parties than cell-phones, and are generally more cost effective, as you only need to pay for the initial cost of the radios. The radios can be very useful for coordinating data collection start and end times, coordinating time synchronization, and reporting equipment problems, among other things.

#### 9.4.2 How do I minimize sources of error and/or bias?

- *Manual LPM from Video Data:* Careful time synchronization and use of a fine interval clock counter is recommended. Diligence and patience should be utilized when performing data reduction from the video tape. Also, data reduction sessions should not last more than 2-3 hours at a time, as eye fatigue and monotony will contribute to increased errors after that length of time.
- *LPM with Portable Computers:* The biggest source of error for this method is inaccurate timestamps. It is virtually impossible to obtain negligible timestamp errors in the field, but careful observer positioning will help to minimize this source of error. Additionally, in situations where periodic stop-and-go traffic conditions occur in the vicinity of the reference point, the data collector must give extra attention to when vehicles actually cross the reference point. To avoid selection bias problems, data collectors should be familiarized with all of the different license plate formats that they would be expected to enter characters from. With the variety of license plates present in the traffic stream, collectors can be temporarily confused about which characters to enter if they are not familiar with that particular format. Also, data collectors should be instructed to sample as many vehicles as possible, and not give preferential entry consideration to any particular types of vehicles.
- *Video Image Tracking:* Due to the many factors to be considered for data collection, readers are referred to appendix B for further information.

## 9.5 Data Analysis and Reporting

The final step is to reduce the data and calculate the desired statistical parameters. When performing the data analysis, the two main questions that are of concern are the following:

- How do I identify and deal with outliers?, and
- How do I report the results?

### 9.5.1 *How do I identify and deal with outliers?*

Aside from the various potential sources of error and bias that have been discussed throughout this document, very significant errors can be incorporated into your results if outliers, are not identified in one's data set and removed. Outliers are usually a result of spurious matches (a partial license plate entry being matched that isn't the same vehicle), or vehicles dwelling off the road in between recording stations. The primary contributors to the presence of outliers are described below.

#### Influences on number of outliers

##### ***Number of plate characters collected***

Schaefer [1988] and Rickman [1990] give some suggestions on the recommended number of plate characters to collect. Although many researchers have suggested three or four characters, it is important to give this issue some consideration because of the associated tradeoffs (i.e., speed of entry, accuracy of entry, number of spurious matches). Essentially, this is just a probability problem. For a six character plate comprised of 3 letters and 3 numbers, there are 17,576,000 unique plates. By collecting only the 3 numbers, for example, there are only 1000 (000-999) unique plates. For the three letters there are  $26^3$  (17,576) combinations. This issue must take into consideration the total number of vehicle plates you expect to record. For a shorter duration study on an arterial, 3 numbers may be appropriate. For a longer duration study on a high volume freeway, 4 or 5 characters may need to be the minimum. While the probability of a spurious

(duplicate) match decreases with the increased number of characters recorded, the possibility of making a data entry mistakes also increases, as well as the amount of time necessary to enter a plate, thus reducing the sample rate. Thus, all these issues must be considered to arrive at an appropriate number that will limit the spurious matches, yet allow for a reasonable sample rate, with a reasonable level of plate entry accuracy. In practice, four characters has been found to be a good compromise for most facilities.

### ***Number of direct access driveways***

Since pass-by and diverted-linked trips are very prevalent in areas with commercial and retail land uses, the number of vehicles exiting a roadway for a short period of time and then reentering the same roadway increases with the amount of these land usage types present along the study facility. Although this can occur while studying either freeway facilities or arterials, it will usually be a much bigger issue for arterials. As shown in the earlier arterial travel time results of chapter 7, it is not unusual to see a difference of 2 to 4 minutes between the minimum and maximum travel times for a heavily signalized arterial corridor of less than  $\frac{1}{2}$  mile in length. It is also not unusual to expect that some vehicles may pull off the road to a site like drive-in banking or a fast food restaurant drive-thru and spend 2 or 3 minutes dwelling at that site before reentering the roadway. For any of these vehicles that get matched, they may have a travel time in the upper end of the spectrum, but without knowledge that they pulled off the road, one may just assume that these vehicles experienced poor progression, when in fact they may have experienced good progression.

At any rate, one can expect outliers to be present in almost any travel time set. These outliers will consist of matches you are not sure are correct, or even if it is a correct match whether that vehicle might have stopped along the route. So the natural question becomes '*Is there any way to reliably identify these outliers?*'. Although the temptation is to want to rely on a single mathematical/statistical formula or procedure to identify outliers, such a formula or method does not exist that readily accounts for all possible sources of data. Some researchers have suggested ad hoc rules such as 'eliminate any

points more than three standard deviations from the mean'. However, in this area, engineering judgement and local knowledge are just as important as any mathematical/statistical rules. In this regard, the very simple process of just “looking” at the data cannot be overemphasized.

If one has decided that only the median is the value of interest, than identification of outliers is not critical. But again, one should “look” at the data to be sure that any potential outlying points are not heavily distributed to one side or the other of the median.

Some previously suggested approaches for dealing with outliers include:

- Upper and lower speed thresholds—subjective and should be determined for each data set separately,
- Use of median value instead of mean—no associated variance measure; thus, sample size formula cannot be utilized, and
- Spread or quartile measures (e.g., a box plot)—statistically robust.

References that provide more detailed discussion about the treatment of outliers include Fowkes [1983] and Hauer [1979]. Additionally, most standard statistics books discuss outliers to some extent.

### 9.5.2 *How do I report the results?*

The ‘*Travel Time Data Collection Handbook*’ and the ITE ‘*Manual of Transportation Engineering Studies*’ offer considerable examples on various data reporting and presentation techniques.

## BIBLIOGRAPHY

---

1. Dew, Andrea M., and Bonsall, Peter W. *Evaluation of Automatic Transcription of Audio Tape for Registration Plate Surveys*. Selected Readings in Transport Survey Methodology, Edited Proceedings of the 3<sup>rd</sup> International Conference on Survey Methods in Transportation, Washington, D.C., January 5-7, 1990, pp. 147-161.
2. IBM<sup>®</sup>, *VoiceType Application Factory for Windows™*, Technical Reference, Version 1.1, 1995.
3. IBM<sup>®</sup>, *VoiceType Developers Toolkit*, Programmer's Guide, Version 3.1, 1997.
4. Rickman, T., Hallenbeck, M., Schroeder, M. *Improved Method for Collecting Travel Time Information*. Transportation Research Record 1271, TRB, National Research Council, Washington, D.C., 1990, pp. 69-88.
5. Schaefer, Mark C. *License Plate Matching Surveys: Practical Issues and Statistical Considerations*. ITE Journal. July 1988. Pp. 37-42
6. Fowkes, A.S. *The Use of Number Plate Matching for Vehicle Travel Time Studies*. Planning and Transport Research and Computation. Summer annual meeting, July 1983.
7. Davidson, Peter W. *Extending the Technique of License Plate Number Matching: Central Edinburgh Experience*. ITE Journal. September 1986. Pp. 41-45.
8. Sosslau, Arthur B. and Honts, Gary E. *Traffic Characteristics Measurement Using License Matching Techniques*. Traffic Engineering. October 1975. Pp. 16-20.
9. Hauer, E. *Correction of License Plate Surveys for Spurious Matches*. Transportation Research A, Vol. 13A.
10. Slavick, M.M. *Errors in Origin-Destination Surveys Done by Number-Plate Technique*. Transportation Research Record 1050, TRB, National Research Council, Washington, D.C., 1985, pp. 46-52.
11. Quiroga, Cesar A. and Bullock, Darcy. *Determination of Sample Sizes for Travel Time Studies*. Institute of Transportation Engineers. ITE Journal on the Web. August 1998. pp. 92-98.
12. Miller, Kenneth S., Harvey, Thomas N., Shuldiner, Paul W., and Ho, Cecilia W. *Using Video Technology To Conduct 1991 Boston Region External Cordon Survey*. Transportation Research Record 1412, TRB, National Research Council, Washington, D.C., 1993, pp. 46-56.

13. Bonneson, James A., Fitts, Joel W. *Traffic Data Collection Using Video-Based Systems*. Transportation Research Record 1477, TRB, National Research Council, Washington, D.C., 1995, pp. 31-40.
14. Kou, Cheng-Chen, Machemehl, Randy B. *Probabilistic Speed-Estimation Measurement Errors Through Video Images*. ASCE Journal of Transportation Engineering, March/April 1997.
15. Breiman, Leo, Lawrence, R., Goodwin, D., Bailey, B. *The Statistical Properties of Freeway Traffic*. Transportation Research, Vol. 11. 1977.
16. Berry, D.S. *Evaluation of Techniques for Determining Over-All Travel Time*. In *Proceedings*. HRB, National Research Council, Washington, D.C., Vol. 31 (1952), pp. 429-439.
17. Walker, William P. *Speed and Travel Time Measurement in Urban Areas*. Bulletin 156. HRB, National Research Council. Washington, D.C. 1957.
18. Muntean Jr., David. *Origin-Destination Travel Surveys in the 1990s Using Microcomputers*. ITE Journal, January 1995.
19. Turner, Shawn M. *Advanced Techniques for Travel Time Data Collection*. Pacific Rim TransTech Conference Proceedings. Seattle, WA. 1995.
20. Liu, Tai K., Gaines, Marsha. *Travel Time Data Collection Field Tests—Lessons Learned*. FHWA, USDOT. 1996.
21. Liu, Tai K. *Field Test of Travel Time Survey Methodologies and Development of a Standardized Data Processing and Reporting System*. Presentation at the National Traffic Data Acquisition conference, September 1994.
22. Bailey, M.D., and Rawling, F.G. *A Computerized Travel Time Study for Northeastern Illinois: Methodology and Commentary*. Working Paper 91-01. Chicago Area Transportation Study, Illinois. June 1991.
23. Hamm, Robert A. *An Evaluation of Travel Time Estimation Methodologies*. Graduate Student Papers on Advanced Surface Transportation Systems, Transportation Engineering Program, Civil Engineering Department, Texas A&M University, 1993.
24. Ford Jr., Gary L. *Travel Time Estimation on Freeways Using Loop Detectors and AVI Technologies*. Compendium: Graduate Student Papers on Advanced Surface Transportation Systems, 1998. Southwest Region, University Transportation Center, Texas Transportation Institute.

25. Washington State Department of Transportation. *Washington State Freeway HOV System Policy: Executive Summary*. 1996.
26. Lomax, T., Turner, S., Shunk, G., Levinson, H., Pratt, R., Bay, P., and Douglas, G. *Quantifying Congestion: Final Report*. NCHRP Report 398, TRB, National Research Council, Washington, D.C., January 1997
27. Turner, S., Lomax, T., Levinson, H. *Measuring and Estimating Congestion Using Travel Time-Based Procedures*. Transportation Research Record 1564. 1996.
28. Turner, Shawn M., Eisele, William L., Benz, Robert J., Holdener, Douglas J. *Travel Time Data Collection Handbook*. Final Report, FHWA-PL-98-035. Federal Highway Administration. 1998.
29. Turner, Shawn M. and Holdener, Douglas J. *Probe Vehicle Sample Sizes for Real-Time Information: The Houston Experience*. IEEE. 1995.
30. Hudson, D. *Computerized Travel Time and Delay Studies*. Proceedings of the 41<sup>st</sup> Annual Meeting of the Institute of Transportation Engineers. July 1988.
31. Epps, A., Chang, J., May, A. *Developing Methodologies for Quantifying Congestion Delay: Final Report*. Report UCB-ITS-RR-94-2. Institute of Transportation Studies, California Department of Transportation, Sacramento, June 1994.
32. IVHS America. *Strategic Plan for Intelligent Vehicle-Highway Systems in the United States*. Report No. IVHS-AMER-92-3. May 20, 1992.
33. Transportation Planning Methods Applications. Abstracts from 5<sup>th</sup> Conference, Seattle, WA, 1995.
34. Nam, Do H. and Drew, Donald R. "Traffic Dynamics: Method for Estimating Freeway Travel Times in Real Time from Flow Measurements". ASCE Journal of Transportation Engineering. May/June 1996.
35. Daley, D.J. Some Statistical Properties of Link Travel Times. National Institute of Statistical Sciences. Technical Report #44. May, 1996.
36. Sen, A., Sööt, S., Ligas, J., and Tian, X. *Arterial Link Travel Time Estimation: Probes, Detectors and Assignment-type Models*. Preprint (#970943) prepared for Transportation Research Board's 76<sup>th</sup> Annual Meeting. 1997.
37. Sisiopiku, Virginia P., Roupail, Nagui M., Santiago, Alberto. *Analysis of Correlation Between Arterial Travel Time and Detector Data from Simulation and Field Studies*. Transportation Research Record 1457. 1994.

38. Sisiopiku, Virginia P., Roupail, Nagui M. *Toward the Use of Detector Output for Arterial Link Travel Time Estimation: A Literature Review*. Transportation Research Record 1457. 1994.
39. Shuldiner, P.W., S.A. D'Agostino, and Woodson, J.B. *Determining Detailed Origin-Destination and Travel Time Patterns Using Video and Machine Vision License Plate Matching*. Transportation Research Record 1551. Transportation Research Board, National Research Council, Washington, D.C. 1996. pp. 8-17.
40. Petty, Karl F., Bickel, Peter, Jiang, Jiming, Ostland, Michael, Rice, John, Ritov, Ya'acov, Scchoenberg, Frederic. *Accurate Estimation of Travel Times from Single-loop Detectors*. Preprint (#971043) prepared for Transportation Research Board's 76<sup>th</sup> Annual Meeting. 1997.
41. Dailey, Dan, HaselKorn, Mark P., Nihan, Nancy L. *Travel Time Estimation Using Cross-correlation Techniques*. Transportation Northwest (TransNow), Final Report #TNW91-02. October 1991.
42. National Traffic Data Acquisition Conference (NATDAC '96). *Proceedings, Volume I and II*. Albuquerque, New Mexico. 1996.
43. Woodson, J.B., P.W. Shuldiner, and S.A. D'Agostino. *Automated Video-based Survey of Travel Times in HOV vs. General Purpose Lanes*. Report No. WA-RD 399.1. Transformation Systems, Inc., Washington State Department of Transportation, Olympia, Washington, October 1995.
44. MNDOT. *Field Test of Monitoring of Urban Vehicle Operations Using Non-Intrusive Technologies*. FHWA Report No. FHWA-PL-97-018. 1997.
45. Chatziioanou, A, Hockaday, S, Ponce, L, Kaighn, S, Staley, C. *Video Image Processing Systems Applications in Transportation, Phase II*. Final Report, California Polytechnic State Univ., San Luis Obispo, CA. 1994.
46. Hoose, N. *IMPACTS: an Image Analysis Tool for Motorway Surveillance*. Traffic Engineering and Control, Vol 33, No 3, March 1992. pp 140-147.
47. Klein, L, Kelley, M [Hughes Aircraft Company]. *Detection Technology for IVHS: Final Report*. FHWA Report No. FHWA-RD-95-100. 1996.
48. Coifman, Benjamin. *Vehicle Reidentification and Travel Time Measurement in Real-Time on Freeways Using the Existing Loop Detector Infrastructure*. Preprint # 981498. Transportation Research Board, 77<sup>th</sup> Annual Meeting, 1998.
49. Dermer, Kay D., Nasburg, Robert E., Lall, B. Kent. *Application of Videotracking Technology to the Measurement of Traffic Statistics in Portland, Oregon*.

- Transportation Research Board, 74<sup>th</sup> Annual Meeting, Washington, D.C., January 22-28, 1995.
50. Hockaday, Stephen. *Evaluation of Image Processing Technology for Applications in Highway Operations – Final Report*. California Department of Transportation Technical Report 91-2. California Polytechnic State University, 1991.
  51. Hughes Aircraft Company and JHK & Associates. *Development of IVHS Traffic Parameter Specifications*, Task A Draft Report for Detection Technology for IVHS, FHWA Contract Number DTFH61-91-C-00076, Fullerton, California, 1994.
  52. Hughes Aircraft Company. *Select and Obtain Vehicle Detectors*. Task D Report for Detection Technology for IVHS. FHWA Contract Number DTFH61-91-C-00076. Fullerton, California. 1993.
  53. Hughes Aircraft Company. *Vehicle Detector Field Test Specifications and Field Test Plan*. Task F Report for Detection Technology for IVHS. FHWA Contract Number DTFH61-91-C-00076. Fullerton, California. 1993, revised May 1994.
  54. Nihan, Nancy L., Leth, Mark, Wong, Abel. *Video Image Processing for Freeway Monitoring and Control: Evaluation of the Mobilizer*. Technical Report, Washington State Transportation Center and Transportation Northwest. 1995.
  55. Nihan, Nancy L., Washburn, Scott S. *Estimating Link Travel Time with a Video Image Tracking System*. Technical Report #TNW 98-03, Transportation Northwest, 1998.
  56. US Patent, # 5,696,503, dated Dec. 9, 1997. *Wide Area Traffic Surveillance Using a Multisensor Tracking System*. Condition Monitoring Systems, Inc.
  57. Press, William H., Teukolsky, Saul A., Vetterling, William T., Flannery, Brian P. *Numerical Recipes in C: The Art of Scientific Computing*. Second Edition. Cambridge University Press. 1992.
  58. May, A.D. *Traffic Flow Fundamentals*. Prentice-Hall, Englewood Cliffs, New Jersey (1990).
  59. Devore, Jay L. *Probability and Statistics for Engineering and the Sciences*. Third Edition. 1991
  60. Box, George E.P., Hunter, William G., Hunter, J. Stuart. *Statistics for Experimenters*. 1978.
  61. Robertson, H.D. (ed.). *Manual of Transportation Engineering Studies*. Institute of Transportation Engineers, Washington, D.C., 1994.

62. Style and Policy Manual For Theses and Dissertations. University of Washington Graduate School. Revised edition, 1999.

## APPENDIX A

### VOICE RECOGNITION PROGRAM FUNCTIONALITY

---

The basic program functionality is described below in a simple outline format.

User specifies new file for recording and pertinent data collection information

User sets number of plate characters to be entered

User specifies time stamp entry scheme

User sets data collection times

User enters data collection details (e.g., time, place, name)

User presses test mode or start recording button

    User presses test mode button, data collection begins immediately

        clock time is irrelevant

        no data are saved to disk

    User presses start recording button

        specified recording start time is later than current clock time

            program goes into stand-by mode

        specified recording start time is equal to earlier than current clock time

            program goes into record mode immediately

Recording has started

    user presses voice recognition activation button

        voice recognition engine becomes active

    user begins speaking license plate characters

        user enters less than specified number of plate characters

            program ignores entry

        user enters specified number of plate characters

            license plate character array voice entries are decoded by voice recognition engine

            voice recognition engine returns text string of entire voice entry

            main application parses string into separate individual characters

            main program converts military word strings to individual letters

            main program converts number word strings to individual numbers

            individual numbers and letters are re-concatenated to form single string of

                all plate characters

            this string is entered into Plate History box and current time is associated

                with entry (i.e., timestamp)

            program provides audible beep to user to signify successful plate entry

            license plate text and timestamp are assigned to arrays for later writing to output file

    process is repeated until user-specified recording stop time is reached

        all data entered since last save interval is written to output file

## **APPENDIX B EQUIPMENT AND DATA COLLECTION CONSIDERATIONS FOR TRAVEL TIME MEASUREMENT WITH THE MOBILIZER VIDEO IMAGE TRACKING SYSTEM**

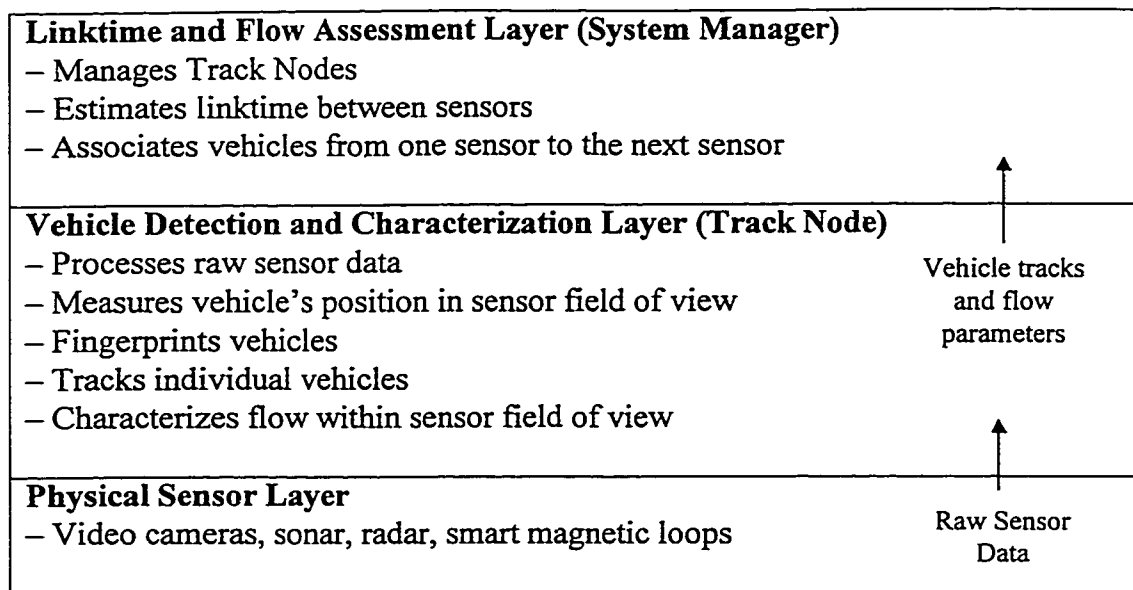
---

The Mobilizer [U.S. Patent No. 5,696,503 (1997)] was developed by Condition Monitoring Systems (CMS). The Mobilizer is a tracking video image system. The advantage of this system is that it not only can provide travel time measurements, but also measurements of all other standard traffic flow parameters. Additionally, it is a passive system; that is, the measurements can be taken without requiring any vehicles to be equipped with special equipment. Travel time data is provided directly by matching vehicles between successive field-of-views (FOV's).

Unlike previous generation video image systems that were capable of only processing specific points within the field-of-view (FOV), the Mobilizer can process individual vehicles as they traverse a FOV. The previous generation video image systems were only capable of providing rough classifications of vehicle length as the only unique vehicle identifying parameter. Thus, these earlier generation video image systems were incapable of determining whether a particular vehicle passed through more than one FOV. The Mobilizer system can extract details from the physical look of a vehicle and use these to give the vehicle a unique identification. If a vehicle in another FOV matches the identification (ID) given to a vehicle in an upstream FOV, the system determines that it has found a match and calculates the travel time between the successive FOV's.

### **Mobilizer Architecture**

In Figure 42, the layered processing architecture of the link-time system is shown. At the lowest (physical sensor) level, there are multiple roadside sensors, each providing an origin or destination sensor input. These sensors can be smart loops, radar sensors, video cameras, or other types of sensors which provide fingerprints of the vehicles. In this study, we used video cameras because at the present time, video offers more unique signature information about vehicles than other non-vehicle invasive technologies.



**Figure 42. Multi-layered architecture for linktime estimation system**

At the second level, the Track Node module processes the raw sensor inputs. It detects vehicles and performs a local tracking function which extracts the vehicle fingerprints and the local traffic statistics from the scene. Up to four Track Nodes may be installed in a single computer, depending on how many camera video inputs are located at that computer.

The third layer, the System Manager, receives the output fingerprints and time/velocity information (tracks) for vehicles from the Track Nodes and performs the functions of linktime estimation between the origin and destination inputs. As many as four Track Node inputs may be processed by one System Manager computer. The System Manager also performs other functions such as managing and monitoring the Track Node functions, providing snapshots from the individual Track Nodes, logging all Track Node and linktime statistics to files, providing numerous real time displays for the operator, and sending data outputs to other computers for ATMS network and Internet functions.

## Hardware Considerations

The Mobilizer system is software driven, using standard personal computer (PC) hardware to function. The software currently runs on a DOS platform, but future plans include upgrading to a Windows NT platform.

There are two hardware related factors that will influence the performance capabilities of the video imaging analysis system. The primary factor is the Central Processing Unit (CPU). CPU speed directly influences the sampling rate of vehicles from the FOV. Since the Mobilizer is a tracking system, it gathers its information by getting multiple "looks" at each vehicle as it travels within a FOV. The more "looks" the system can get at a vehicle within the FOV, the more information it obtains for developing the vehicle signature (or ID). Obviously, to obtain more "looks" (i.e., greater sampling rate), it requires more computing speed. Additionally, the faster the speed of the traffic within the FOV, the faster the sampling rate required.

The sampling speed of the Mobilizer Track Node is usually set to between 2 and 4 samples per second. This allows the system to obtain information about the vehicle's fingerprints every 250-500 milliseconds as it moves through the FOV. The most reliable fingerprint matching information is obtained when the vehicle is close to the camera, so the sample rate must be fast enough to detect a vehicle with a speed of 70 mph in an area near the camera. The specialized image processing hardware installed in a PC for this project was capable of processing two camera inputs with a 233 MHz Pentium processor at 4 samples per second, or one camera input with a 133 MHz processor at about 3.5 samples per second.

A critical piece of hardware is the video digitizing card. This is necessary to process the live or recorded video into the appropriate format that the system software can process. With a very fast cpu, it is possible to use one video card to process two streams of independent video simultaneously. While the system can function with some off-the-shelf video digitizing cards, the video card specified and built specifically for the Mobilizer system will achieve better results than off-the-shelf video digitizing cards.

Video cameras are of course a necessary piece of equipment in the data collection. The quality of the camera image is a significant variable in the video imaging process. The cameras must be capable of providing high-resolution, good signal to noise ratio, high dynamic range (able to pick up both very bright and very dim objects), and high shutter speed to prevent the CCD (Charge-coupled device) focal plane from smearing the data from rapidly moving vehicles.

### **Data Collection Considerations**

The quality of the obtained video data is a critical factor to the success of the travel time measurement capability of the Mobilizer. The three main considerations for obtaining quality video data are: 1) the data collection equipment; 2) the data collection equipment setup, and; 3) the data collection site characteristics.

#### *Data Collection Equipment*

The biggest fundamental change in the data collection procedure versus the earlier research [Nihan, Leth, and Wong, 1995] on this system was the video cameras used. The video camera image quality is a key variable to the success of the Mobilizer system. In the earlier, WSDOT surveillance video cameras were utilized. It was eventually determined that these cameras were inadequate, not just in terms of location, but also because of video image quality. As the WSDOT cameras are currently used for surveillance purposes only, maintaining the cameras to the highest video image performance levels is not an issue. As long as WSDOT personnel can still see the freeway and vehicles through the camera, they have no reason to inspect the camera or adjust its performance. This is clearly inadequate for video imaging. Additionally, while many of the WSDOT cameras probably have the capability of using high shutter speeds, there is no guarantee that the camera is set to use a high shutter speed, as the camera operators never look at still images captured from the video camera. There is also no guarantee that the camera is properly focused because significant differences in the focus level are difficult to perceive to the passive eye. However, small differences in the focus

level can make large differences in the image clarity obtained from capturing a still frame of video.

Additionally, the automatic gain control feature (AGC) is often enabled on the WSDOT surveillance cameras. Previous studies [Hockaday 1991, Hughes and JHK 1993] have mentioned that image quality from cameras that have automatic iris and gain control (AGC) tends to deteriorate for periods of time in reaction to intense light changes. This type of camera darkens the light level in the video image in response to the intense light.

Thus, for this project, the decision was made to purchase and use separate high-performance video cameras for data collection. This change was made for two reasons: 1) Ability to place cameras in a more optimal position, and 2) Ability to ensure video image quality would always be at its best. More details will be provided on these issues in the sections that follow. The cameras purchased for this research were Canon L2 high-quality 8mm video cameras. These cameras provide excellent image resolution and are capable of capturing high quality images under high traffic speeds and varying light conditions.

Due to the typically short time that power is provided by the factory camera batteries, heavy-duty marine 12-volt batteries were usually used for data collection. The cameras were set to a manual mode so that focus and gain control would not be change during the data collection from the initial set up values. Also, different lenses (e.g., polarized, UV) could be utilized depending on the ambient light conditions. The microphone was also installed on the camera to record important information spoken by the operator, such as time stamps.

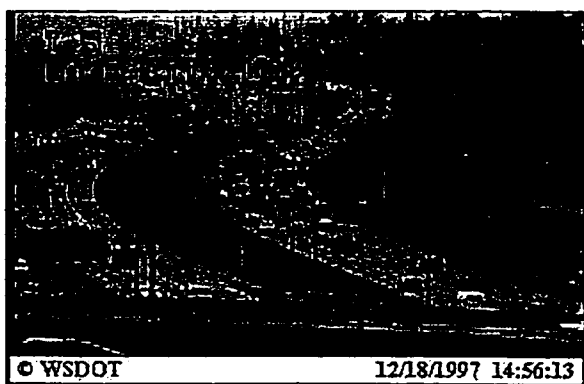
### *Data Collection Equipment Setup*

#### Camera Placement/Perspective

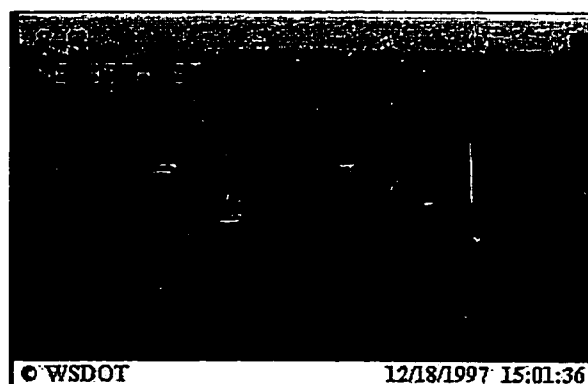
Another fundamental change aside from the camera equipment itself, was the positioning of the cameras for data collection. Departments of Transportation have traditionally placed surveillance cameras in locations that provide good wide-area

coverage of all freeway lanes in both directions. However, these locations are generally less than ideal for vehicle tracking purposes.

For successful tracking of vehicles from one camera location to another, it is essential that the fields-of-view (FOV's) of each camera "look" as similar as possible. Since the tracking algorithms rely on the vehicles "looking" as similar as possible between successive FOV's, the best way to ensure this is to set the cameras up such that vehicles maintain as constant geometry as possible. For the data collection done for this project, the video cameras were positioned on overpasses directly over the center of the lanes used for analysis. As an illustration, you can see in Figure 43 and Figure 44 how two adjacent surveillance cameras positioned near Interstate 5 provide completely different vehicle perspectives.



**Figure 43. Surveillance camera image looking south.**



**Figure 44. Surveillance camera image looking south.**

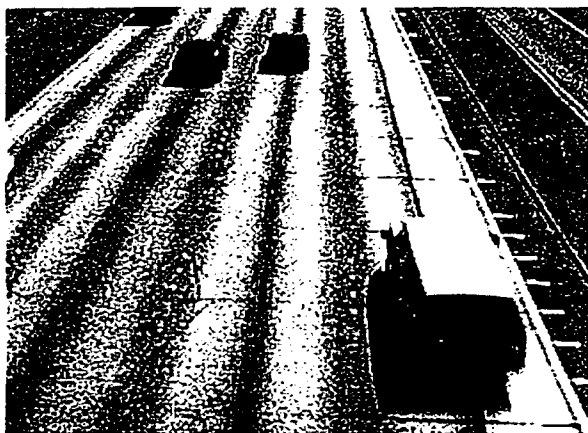
Additionally, lane occlusion due to tall vehicles has often been cited as a problem for video image systems that use images from cameras positioned to the side of the travel lanes, rather than over them analyzed [Hockaday 1991; Hughes and JHK 1993]. With the camera placement used for this project, this problem was also eliminated. Use of cameras placed directly over the roadway also gives the ability to place cameras at a lower height than those mounted on high poles offset from the roadway.

### Shutter Speed

It is necessary to use a shutter speed which is capable of providing unblurred images of the vehicles during a "frozen" camera image. The appropriate shutter speed

will depend on the speed of vehicles. The following images demonstrate the difference in image clarity.

Figure 45 shows an image from a portion of video shot using a shutter speed of 1/60 second. Figure 46 shows an image from another portion of the same video that used a shutter speed of 1/1000 second. Both videos were of freeway traffic traveling between 55 and 65 mph.



**Figure 45. Video image snapshot using 1/60 shutter speed**



**Figure 46. Video image snapshot using 1/1000 shutter speed**

Generally, shutter speeds in excess of 1/500 second should provide adequate results for freeway free-flow speeds. Video images taken of slower traffic on arterials should still obtain good results using shutter speeds of 1/100 - 1/250 second.

For the Mobilizer system, there is really no disadvantage to using a high shutter speed in all cases; thus, it makes sense to use a shutter speed of at least 1/1000 second in all cases. For human video viewing, the main disadvantage to a high shutter speed is that the video can sometimes appear jerky in continuous motion. Since the Mobilizer system is really only looking at very frequent snapshots of the video image, it is better to have the still images look as sharp (unblurred) as possible. Since the Mobilizer system is only sampling a fraction of all possible video frames anyway, the digitized video will still look jerky even using the slowest shutter speed.

### Camera Spacing

In this project, the distance between successive camera locations was generally in the ¼ mile to ½ mile range. This was done for two reasons: 1) the shorter distances mean there is less probability of a given vehicle changing lanes within that distance, and 2) there is less chance of a vehicle entering the traffic stream between the two camera stations that looks similar to one identified in the upstream FOV.

One of the requirements of the Mobilizer system to identify a match is that the vehicle is in the same lane in the downstream FOV as it was in the upstream FOV. While this may seem like a restrictive requirement, allowing matches from different travel lanes opens up a greater propensity for mismatches because of the high percentages of similar looking vehicles in the traffic stream.

While the number of different motor vehicle models manufactured today is greater than it has ever been, it seems as though the average vehicle composition on any given roadway segment at any given time is becoming more homogeneous than ever. This can probably be explained by the extreme popularity of certain types of vehicles. For example, mini-vans and sport utility vehicles are beginning to dominate the average traffic stream. While the Mobilizer had good success at discerning between multiple models of these vehicle types over the short distances used in this study, it is obvious that using large spacings between cameras can yield inaccurate travel times. Large as used in this context is a relative term that depends on the specific area. Areas that have many departure and entry points between the camera locations for a specific roadway facility will have a greater chance of matching two vehicles that look the same but are really not the same.

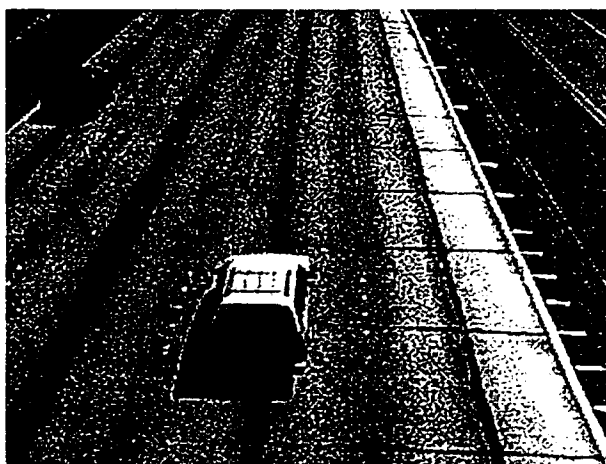
### *Data Collection Site Characteristics*

#### Distance Calibration

In addition to the field of view, the other critical component to successful operation of the tracking system is the accurate setup of the calibration lines within the FOV. Since an important parameter in identifying a particular vehicle is its length, it is

necessary to have points of reference within the FOV for which specific distances are known. For the arterial data, this was accomplished by measuring distances between raised pavement markers. For the freeway data, this was accomplished by painting reference lines on the shoulder at specific distances. Since the Mobilizer system gathers most of its information at the near end of the FOV, it is best to have calibration lines more closely spaced in the first 50 feet of the FOV. In the previous study, the calibration lines were placed at only 50 foot intervals. The Mobilizer development team later determined that calibration line spacings that large were totally inadequate.

Figure 47 shows the calibration lines on the right side of the image, in the shoulder area. The first seven stripes are spaced five feet apart, the next two are spaced ten feet apart, the next four are spaced 25 feet apart, and the remaining stripes are spaced 50 feet apart. The distance from the first calibration mark to the edge of the overpass in this FOV is 55 feet.



**Figure 47. Example calibration line setup**

The process of placing the calibration lines requires considerable effort. It is generally necessary to first set up a camera and determine the field-of-view. From this FOV, you can find a reference mark on the pavement to determine where to start the calibration lines. The material used for the calibration marks is also important. There are industrial brands of pavement tape available, but the environmental conditions must be

considered. The initial pavement marking effort used pavement striping tape. Following two solid days of rain, there were almost no perceptible traces that the tape had ever been placed on the pavement. Thus, it was obvious that another striping material would be needed. After obtaining permission from the WSDOT, the next pavement marking effort used pavement marking spray paint. These marks withstood the local environmental conditions for a very long time. The process of putting markings down on the pavement is in itself a challenging prospect. Even with the assistance of a WSDOT safety truck, the person placing the tape or paint puts themselves in jeopardy because of working so close to the travel lanes and high-speed traffic. Since this method of determining calibration distances was not very practical for repeated test sites, it was decided that another method should be investigated in the interest of time efficiency and safety.

A Total Station, an electronic surveying instrument, was investigated for determining calibration distances. As Figure 48 below illustrates, it sometimes can be impossible to place these marks without closing a lane of traffic. For this particular FOV, the stripes would need to be placed in the median area, and there is not enough room to work there without closing the HOV lane.



**Figure 48. FOV with very limited area for placing calibration marks**

A notable feature about this FOV are the dark lines (pavement section joints) running perpendicular to the travel lanes. Since these show up very well on the digitized image, knowing the distances between these lines is a good calibration reference.

Additionally, the raised pavement markers provide good measurement points. With the use of a Total Station, these points could just be measured from the overpass above the freeway. For situations like this, the Total Station can be a very effective tool.

The only limitation is that you can only measure what is there; unlike the physical placement of markings, where you can place them anywhere you like, with the Total Station method, you have to measure existing things in the roadway that can be seen within the FOV. Care must also be taken, however, because items that are just barely perceptible from the camera FOV may not be distinguishable once the video is digitized without color and at lower resolution.

The Total Station is a sophisticated piece of surveying electronics. The Total Station allows you to focus on distant and small objects and measure very precise angles from a reference point to those objects. It also allows you to measure the distance between the Total Station and a reflector.

The general method for using the Total Station to measure distances for the test sites is as follows:

1. Setup the Total Station on a tripod at one end of the overpass and a tripod mounted reflector at the other end (for our test sites, there were sidewalks present that we could use to setup on; thus not interfering with arterial traffic either). When determining the position to place both the Total Station and the reflector, it is better to place them as far apart as practicable, as that will contribute to the overall accuracy. However, it is also necessary that the Total Station will be able to “see” all the objects that need to be measured on the roadway surface from both tripod locations.
2. Measure the horizontal distance between the Total Station and the reflector. Once the Total Station is properly aimed at the target (reflector), this is an automatic and immediate process.
3. With the Total Station still aimed at the reflector, set the horizontal angle to zero.
4. Rotate and aim the Total Station at every object you want to measure the distance for, and then obtain the horizontal angle reading for each object.
5. Move the Total Station to the other tripod (the one with the reflector previously mounted on it) and repeat the angle measurements for each of the objects measured from the other location.

6. The final step just involves using all the measurements and simple trigonometry to compute distances between objects. Essentially, every measured object forms a triangle between itself and the two tripod locations, and since you know the distance of one side of the triangle (the distance between the two tripods) and two angles (one measured to each object from both tripod locations), you have all the information you need to solve for the distances to each object.

This procedure was tested for one of the arterial locations, for which distances had already been measured, and the results matched.

## APPENDIX C LICENSE PLATE MATCHING AND SIMULATION PROGRAM DESCRIPTION

The screen capture shown below shows the file details after the upstream and downstream data files have been opened for processing.

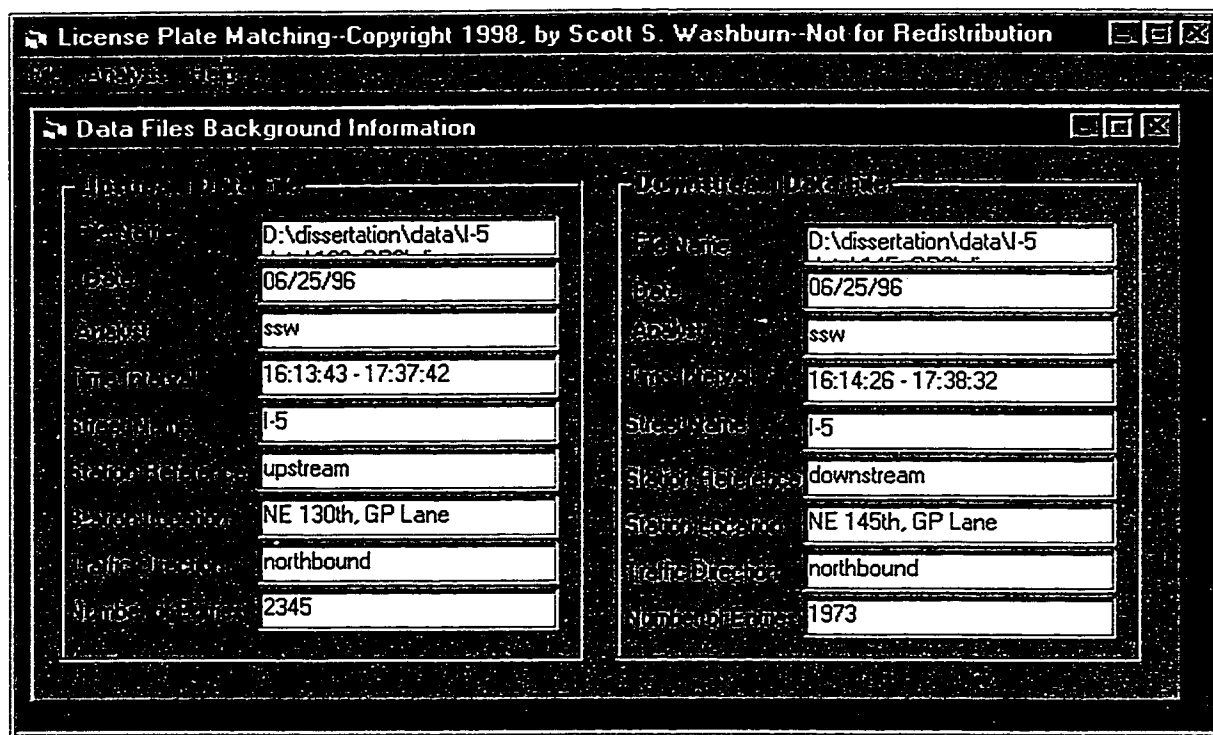
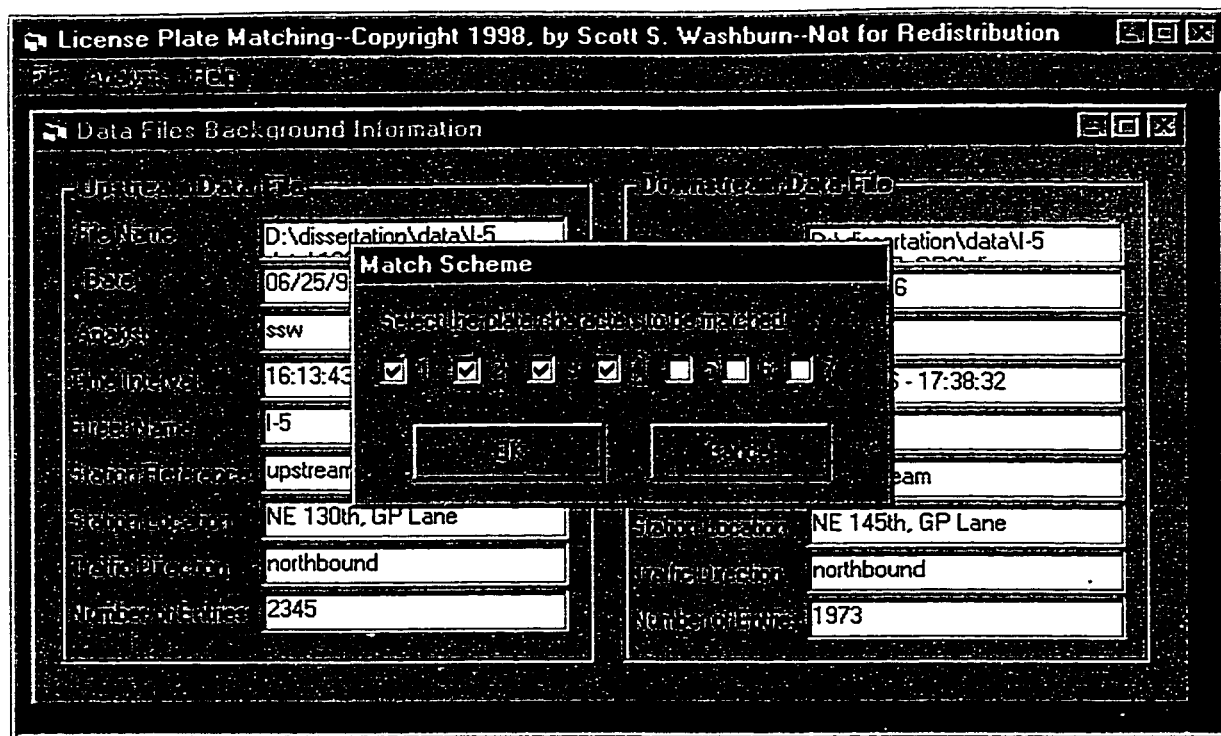


Figure 49. Matching program main screen

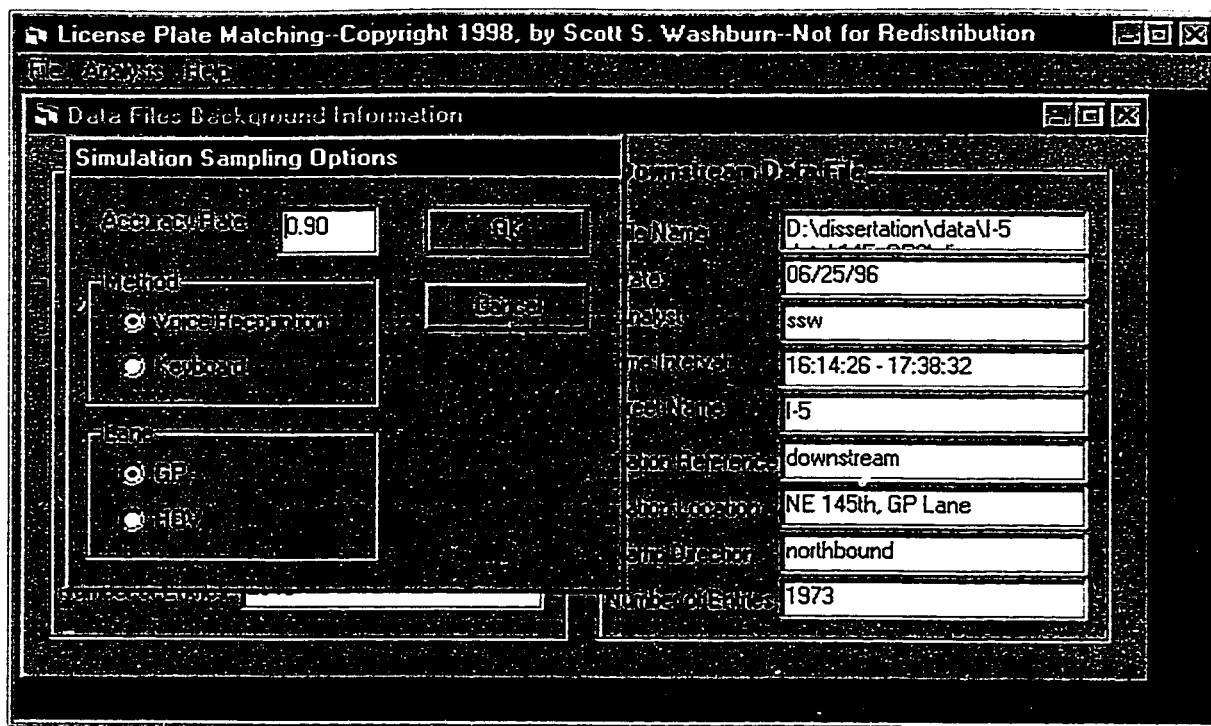
The screen capture shown below displays the dialog box that allows the user to specify which plate characters will be recorded during the simulation. This step is skipped if exact matching of the original strings contained in each of the data files is required.



**Figure 50. License plate character entry/match scheme**

After selecting the plate entry scheme, the upstream and downstream license plate files get converted to the corresponding character strings.

The next step is to specify the entry accuracy rate desired, which data collection method should be simulated, and which lane of data is being processed. These items are specified with the following dialog box.



**Figure 51. Sampling options dialog box**

The last step is the sampling simulation and/or license plate matching. When the sampling simulation gets performed, the probabilistic approach is implemented to determine which vehicles get sampled. And then the program determines whether the plate is entered correctly, through a probabilistic uniform distribution, and based upon the user-specified entry accuracy rate. The results are output to a standard text file that can be loaded into a spreadsheet for further analysis.

## APPENDIX D

### LICENSE PLATE CONSIDERATIONS

---

Washington State Law (RCW 46.16.240) states “The vehicle license number plates shall be attached conspicuously at the front and rear of each vehicle for which the same are issued and in such a manner that they can be plainly see and read at all times...”

Additionally, RCW 46.16.240 states “Each vehicle license number plate...shall be kept clean so as to be plainly seen and read at all times” and “It shall be unlawful...to display upon any vehicle any vehicle license number plate or plates which have been in any manner changed, altered, disfigured or have become illegible. License plate frames may be used on vehicle license number plates only if the frames do not obscure license tabs or identifying letters or numbers on the plates and can be plainly seen and read at all times. It is unlawful to use any holders, frames, or any materials that in any manner change, alter, or make the vehicle license number plates illegible.

Despite these laws a surprising number of vehicles do not have a license plate on the front of the vehicle. A small percentage also have the license plate placed on the dashboard of the vehicle. There are also a large number of vehicles that have license plate frames that cover the bottom part of the letters on the license plate. This makes it almost impossible to distinguish between several letters, like E and F, I and T, and L and J. The use of tinted, full cover license plate “holders” also seem to be getting more popular. While these generally add to the difficulty of reading the plate during typical daytime data collection hours, they undoubtedly make plates more difficult to read during nighttime hours, especially since they tend to neutralize the reflective coating that is placed on the plates for the specific purpose of aiding nighttime visibility. There does not seem to be any good reason for allowing these types of license plate “holders”.

A lot of dirty (4x4 vehicles mainly), mangled/bent, and extremely faded plates were also all too common.

Although it is extremely rare to encounter a vehicle without a rear license plate, this plate position is far from ideal as well. A large percentage of vehicles have trailer

hitches which usually occlude at least one or two plate characters. And trunk-mounted bicycle racks often times block parts of the rear license plate.

**Table 42. Unreadable Plate Percentages**

| <b>Data Set<sup>a</sup></b> | <b>Missing plate (1)</b> | <b>Mangled / Bent (2)</b> | <b>Faded / dirty (3)</b> | <b>Occluded plate<sup>b</sup> (4)</b> | <b>Total (1,2,3)</b> | <b>Total (1,2,3,4)</b> | <b>Total Vehs</b> |
|-----------------------------|--------------------------|---------------------------|--------------------------|---------------------------------------|----------------------|------------------------|-------------------|
| <b>GP Total</b>             | 1.0%                     | 0.1%                      | 0.5%                     | 2.2%                                  | 1.6%                 | 3.8%                   | 5107              |
| <b>HOV Total</b>            | 1.2%                     | 0.1%                      | 0.3%                     | 1.8%                                  | 1.6%                 | 3.4%                   | 4304              |
| <b>Arterial 1</b>           | 1.4%                     | 0.0%                      | 0.1%                     | 2.6%                                  | 1.5%                 | 4.1%                   | 1224              |
| <b>Arterial 2</b>           | 5.2%                     | 1.0%                      | 0.0%                     | 2.0%                                  | 6.2%                 | 8.2%                   | 1225              |

<sup>a</sup> The first two rows refer to rear license plates, the second two rows refer to front license plates.

<sup>b</sup> Occluded plates include those that are not completely readable due to trailer hitches, bicycle racks, and oversized license plate frames.

Out-of-state plates through another wrinkle into the mix. Every state has various plate character arrangement formats. Accounting for all the formats of just one state in the data collection is burdensome enough, without adding the complexity of all the various possible formats from other states. Certainly, one option is to just ignore out-of-state plate. However, you are then ignoring potential match candidates and artificially constraining your vehicle population. Ignoring out-of-state plates in some parts of the country are probably not a significant consideration, but certainly in parts of the East where many different out-of-state plates may be present in any given state at one time can certainly be a significant factor. By not ignoring out-of-state plates, you do run an additional risk of false matches due to overlap of identical plate numbers from other states. In this case, some sort of state identification code may want to be considered for entry along with the plate characters.

#### **From RCW 46.16.305 (Special License Plates)**

Finding--1997 c 291: "The legislature finds that the proliferation of special license plate series has decreased the ready identification of vehicles by law enforcement,

and increased the amount of computer programming conducted by the department of licensing, thereby increasing costs. Furthermore, rarely has the actual demand for special license plates met the requesters' projections. Most importantly, special plates detract from the primary purpose of license plates, that of vehicle identification." [1997 c 291 § 1.]

These issues affect both the manual form of data collection which requires human observation of license plate numbers and the automatic forms of license plate reading using video image analysis systems.

**VITA**

---

Scott Stuart Washburn

University of Washington

1999

**Academic Background**

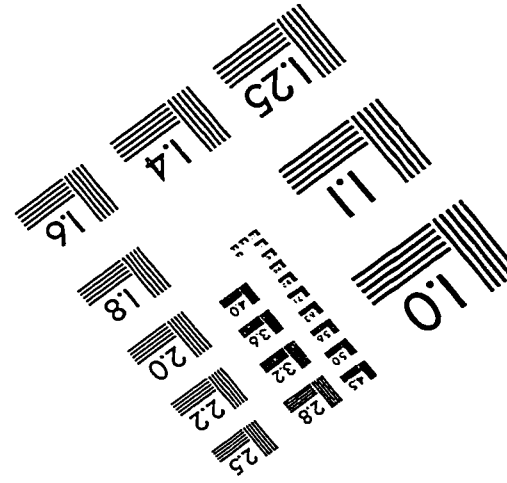
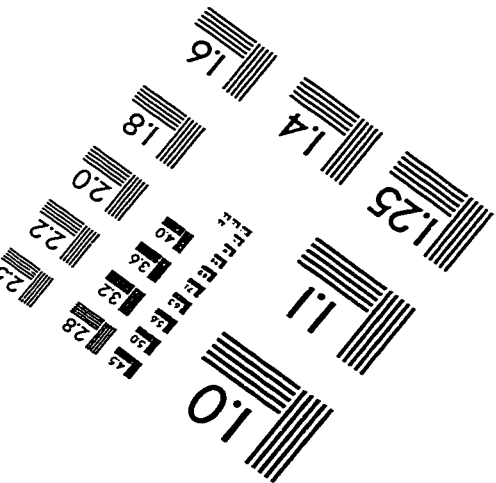
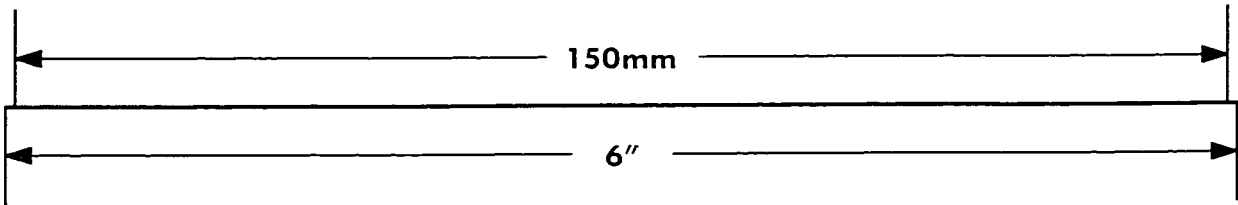
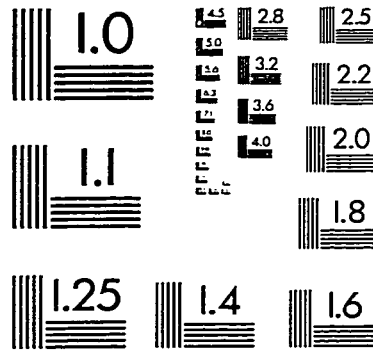
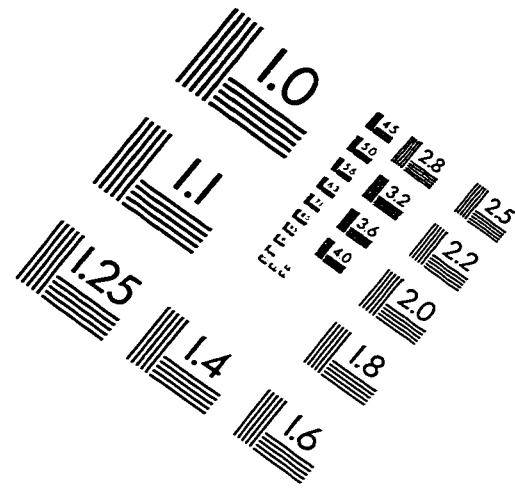
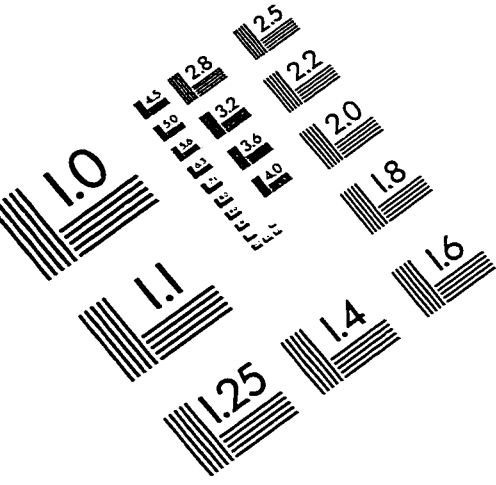
- June 1999      Doctor of Philosophy in Civil Engineering, University of Washington
- March 1993     Master of Science in Civil Engineering, University of Washington
- June 1991      Bachelor of Science in Civil Engineering, University of Washington

**Professional Background**

- 4/99 – 6/99     **Transportation Engineer**, Transportation Solutions, Incorporated (TSI), Redmond, WA—Part Time.
- 1/99 – 3/99     **Instructor**, Department of Civil and Environmental Engineering, University of Washington. Taught CIVE 441 and CETS 511.
- 9/94 – 3/99     **Research Assistant**, Department of Civil and Environmental Engineering, University of Washington.
- 9/97 - 3/98     **Instructor**, Department of Civil Engineering, College of Engineering, University of Washington. Taught CIVE 410, CIVE 441, and ENGR 100.
- 5/97 - 10/97    **Transportation Engineer**, HNTB, Bellevue, WA—Half-time Special Consultant for Tacoma Narrows Corridor Traffic Modeling Project.
- 6/95 - Present **Program Manager**, Software for Transportation Solutions (STS), Kirkland, WA.
- 9/96 - 3/97 }  
9/95 - 3/96 } **Teaching Assistant**, Dept. of Civil Engineering, University of Washington.  
1/95 - 6/95 }
- 9/94 – 9/95     **Transportation Engineer**, Transportation Solutions, Incorporated—Part Time, On-Call work.

- 1/94 - 9/94 **Transportation Engineer**, Transportation Solutions, Incorporated—Full Time.
- 9/92 - 1/94 **Transportation Engineer**, David Evans and Associates, Bellevue, WA.
- 9/91 - 8/92 **Research Assistant**, Dept. of Civil Engineering, University of Washington.
- 2/90 - 8/91 **Computer Systems and Software Consultant/Programmer**, TransNow, U.S.D.O.T. Regional Research Center, Department of Civil Engineering, University of Washington.

# IMAGE EVALUATION TEST TARGET (QA-3)



APPLIED IMAGE, Inc  
1653 East Main Street  
Rochester, NY 14609 USA  
Phone: 716/482-0300  
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved