SHRP 2 Reliability Project L02

Establishing Monitoring Programs for Travel Time Reliability

PREPUBLICATION DRAFT • NOT EDITED

SHRP 2
STRATEGIC HIGHWAY RESEARCH PROGRAM
Accelerating solutions for highway safety, renewal, reliability, and capacity

TRANSPORTATION RESEARCH BOARD
OF THE NATIONAL ACADEMIES

ACKNOWLEDGMENT
This work was sponsored by the Federal Highway Administration in cooperation with the American Association of State Highway and Transportation Officials. It was conducted in the second Strategic Highway Research Program, which is administered by the Transportation Research Board of the National Academies.

NOTICE
The project that is the subject of this document was a part of the second Strategic Highway Research Program, conducted by the Transportation Research Board with the approval of the Governing Board of the National Research Council.

The members of the technical committee selected to monitor this project and to review this document were chosen for their special competencies and with regard for appropriate balance. The document was reviewed by the technical committee and accepted for publication according to procedures established and overseen by the Transportation Research Board and approved by the Governing Board of the National Research Council.

The opinions and conclusions expressed or implied in this document are those of the researchers who performed the research. They are not necessarily those of the second Strategic Highway Research Program, the Transportation Research Board, the National Research Council, or the program sponsors.

The information contained in this document was taken directly from the submission of the authors. This document has not been edited by the Transportation Research Board.

Authors herein are responsible for the authenticity of their materials and for obtaining written permissions from publishers or persons who own the copyright to any previously published or copyrighted material used herein.

The Transportation Research Board of the National Academies, the National Research Council, and the sponsors of the second Strategic Highway Research Program do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of the report.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. On the authority of the charter granted to it by Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

The **Transportation Research Board** is one of six major divisions of the National Research Council. The mission of the Transportation Research Board is to provide leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal. The Board's varied activities annually engage about 7,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation. **www.TRB.org**

www.national-academies.org

Copy No. _____

SHRP 2 Project L02
Establishing Monitoring Programs for Travel Time Reliability

Final Report

Prepared for:
Strategic Highway Research Program 2

(SHRP 2)

January 31, 2013

Institute for Transportation Research and Education

In association with:

Iteris/Berkeley Transportation Systems, Inc.
Kittelson & Associates, Inc.
National Institute of Statistical Sciences
University of Utah
Rensselaer Polytechnic Institute
Joseph Schofer of Northwestern University
Asad Khattak of Planitek

ACKNOWLEDGMENT OF SPONSORSHIP

This work was sponsored by Federal Highway Administration in cooperation with the American Association of State Highway and Transportation Officials, and it was conducted in the Strategic Highway Research Program, which is administered by the Transportation Research Board of the National Academies.

Table of Contents

GLOSSARY

ACRONYMS

APC	Automated Passenger Count
ASOS	Automated Surface Observing System
AVI	Automated Vehicle Identification
AVL	Automated Vehicle Location
AWOS	Automated Weather Observing System
CAD	Computer Aided Dispatch
CDF	Cumulative Density Function
CV	Connected Vehicle
ESS	Environmental Sensor Station
ETC	Electronic Toll Collection
GPS	Global Positioning Satellite
HAR	Highway Advisory Radio
ITS	Intelligent Transportation System
LCS	Lane Control Status
LPR	License Plate Reader
MAC	Media Access Control
PDF	Probability Density Function
PeMS	Performance Measurement System
RFID	Radio-Frequency Identification
SHRP 2	Strategic Highway Research Program 2
TMC	Transportation Management Center
TT-CDF	Travel Time Cumulative Density Function
TT-PDF	Travel Time Probability Density Function
TTRMS	Travel Time Reliability Monitoring System
VMT	Vehicle Miles Traveled

TERMS

Market: A set of users in combination with a route bundle
Monument: A reference point used for travel time measurement

Non-Recurring Event: An event that does not occur regularly during a typical time of day, including traffic incidents, work zones, weather, special events, traffic control devices, and fluctuations in demand. The effect of non-recurring events can be magnified by inadequate base capacity.

Passage Time: A timestamp assigned to a vehicle when it passes a given monument

Regime: The categories of conditions under which the segment, route, or network is operating at a given point in time (or from one time to another). It is effectively the "loading condition" on the system.
Route: A sequence of segments
Route Bundle: A set of two or more routes

Sample Space: The set of raw data that pertain to each context for which a probability density function is being developed, such as those that pertain to a regime (e.g., congested conditions) or to another logical grouping (e.g., 7:00 a.m. to 9:00 a.m.) Also known as an observation set, observation time frame, or sample frame.
Segment: A path between two monuments

Travel Rate: Travel time per unit distance
Travel Time: The amount of time spent traveling over a given segment or route
Trip Time: The door-to-door time for a trip

Use Case: Description of a system's behavior based on user needs
User: People or package making a trip across the network

EXECUTIVE SUMMARY

Within SHRP 2, Project L02 was focused on creating a suite of methods by which transportation agencies could monitor and evaluate travel time reliability. Creation of the methods also produced an improved understanding of why and how travel times vary and the factors that create that variation.

The Final Report provides a brief narrative about what reliability is and how it can be measured and analyzed. A general finding is that reliability is best described by creating holistic pictures like probability density functions (PDFs) and their associated cumulative density functions (CDFs). The PDFs are helpful for identifying multi-modality or the existence of multiple operating conditions within the data being examined (Barkley, *et al.* 2012, Guo *et al.* 2010, Fraley and Raftery 2009). The CDFs are helpful for seeing if progress is being made in making a system more reliable or for comparing the reliability of one system against another.

A survey of the state-of-the-art and state-of-the-practice in travel time reliability monitoring systems worldwide helped guide development of the methods. It showed that Europe and Asia were slightly ahead of the United States at the time the project started. A second survey among potential future uses of the monitoring system helped guide its functional features. The potential users included: 1) system administrators and their staffs, 2) highway system operators, 3) transit system operators, 4) freight service providers, 5) highway system users, 6) transit system users, and 7) freight system users. Each had its own special needs with consistency being evident among the system operators (1, 2, 3, and 4) and the users (5, 6, and 7). The findings from the survey were coalesced into a set of use cases that became the driving force behind the system's functional specifications.

The project's main product is a guidebook which describes how an agency should develop and use a Travel Time Reliability Monitoring System (TTRMS). The guidebook follows the block diagram presented in Figure ES-1 for purposes of describing the TTRMS. Each module is shown as a box, while the inputs and outputs are shown as circles.

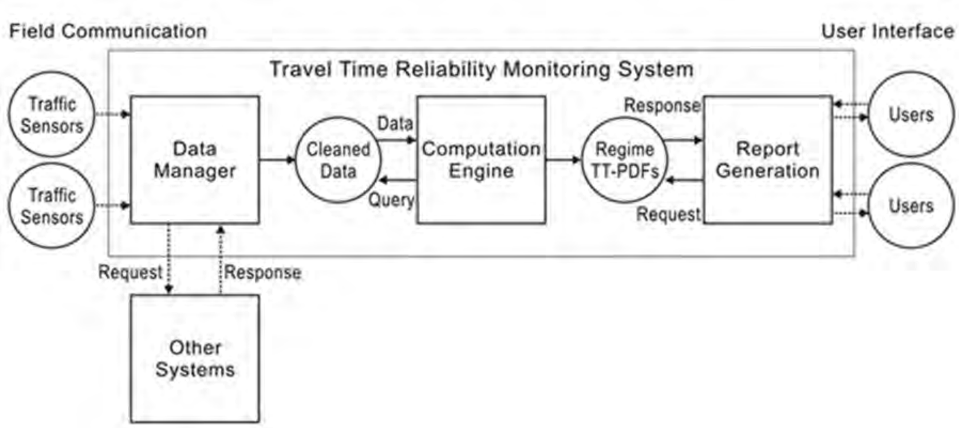


Figure ES-1: Travel Time Reliability Monitoring System Modules

The monitoring system is not intended to be stand-alone. Rather, it is intended to mate up with an existing traffic management system.

The three major modules of the monitoring system are: a data manager, a computational engine, and a report generator. The data manager assembles incoming information from traffic sensors and other systems, such as weather data feeds and incident reporting systems, and places them in a database that is ready for analysis as “cleaned data”. The computational engine works off the cleaned data to prepare “pictures” of the system’s reliability: when it is reliable, when it is not, to what extent, under what conditions, etc. In the exhibit this is illustrated by “regime TT-PDFs”. The report generator responds to inquiries from users—system managers or travelers—and uses the computation engine to analyze the data and provide information that can then be presented back to the inquirer or decision maker.

The Guidebook uses five chapters to describe the monitoring system:

1. *Introduction*: an overview of travel time reliability.
2. *Data Collection and Management*: the types and application of various types of sensors, the management of data from those sensors, and the integration of data from other systems that provide input on sources of unreliability (e.g., weather, incidents). This represents the left side of the figure in Exhibit 1 and includes traffic sensors, other systems, and the data manager.
3. *Computational Methods*: how probability density functions can be derived from the variety of data sources. This represents the center part of the figure in Exhibit 1 and includes the process of generating travel time probability density functions that can be used to derive a variety of reports to users.
4. *Applications*: a discussion about five real-world case studies that were conducted as part of the project as well as a set of use cases that show how the methods can be applied.
5. *Analytical Process:* a beginning-to-end discussion about how the guidebook indicates travel time reliability should be analyzed under various conditions.

The Guidebook is supplemented by four documents that provide additional detail to support the development and application of travel time monitoring systems. These documents are as follows:

A. *Monitoring System Architecture.* This document presents examples of detail data structures for the organization of various data sources. This document provides supporting detail for Chapter 2 of the guidebook.
B. *Methodological Details.* This document presents detailed discussions of the analytical methods that can be used to calculate travel time reliability measures from a variety of input sources. This document provides supporting detail for Chapter 3 of the guidebook.
C. *Case Studies.* This document presents a series of detailed case studies that exercise various aspects of the guidebook, including system architecture, analysis of recurrent and non-recurrent sources of congestion, and the application of a

variety of use cases. This document provides supporting detail for Chapter 4 of the guidebook.

D. *Use Case Demonstrations.* This document illustrates the application of a variety of use cases for a travel time reliability monitoring system. This document provides supporting detail for Chapter 4 of the guidebook.

An executive summary is also provided for the Guidebook. It gives agency managers a description of what a TTRMS is, why it is valuable, and how it can be used.

Travel time reliability has been regarded by the L02 team as the absence of variability. That is to say that a system, segment, or route has reliable travel times if it has consistent travel times for a given operating condition every time that condition arises.

For example, Figure ES-2 shows average travel times on workdays during 2011 for a route on I-5 in San Diego, California. It is clear that the travel times on this route are not always the same; unfortunately, the system is not completely reliable. Not only does the level of congestion have an effect, as shown by the data points for the “None” condition, but non-recurring events have impacts as well.

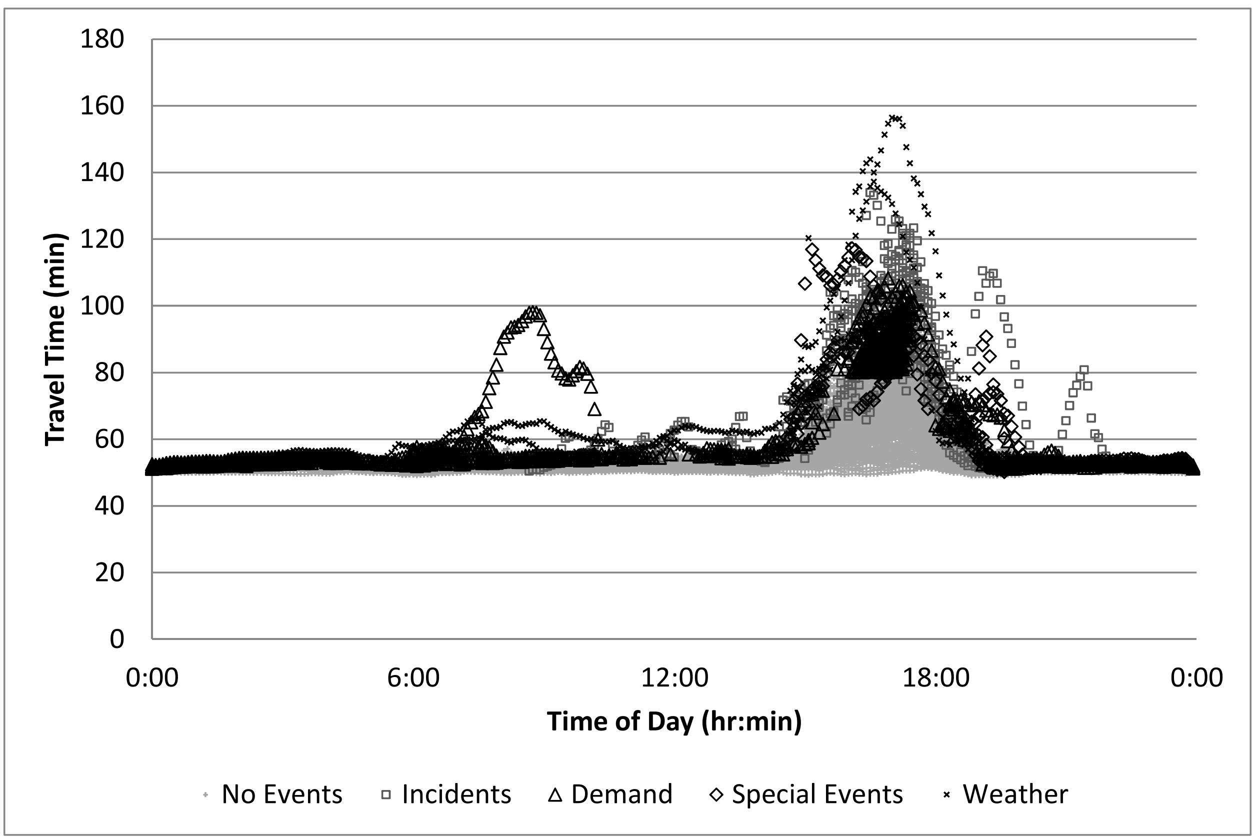


Figure ES-2: Variations in Travel Times by Time of Day Across a Year

Figure ES-2 is valuable for gaining an understanding of how the system is operating, but it does not provide a summary of that performance nor does it provide helpful information in guiding system managers toward actions that might be taken.

Through cumulative density functions (CDFs), the monitoring system takes the data displayed in Figure ES-2 and summarizes it in a fashion that makes the performance of the facility clear and helps stimulate ideas for mitigating actions. Figure ES-3 shows the CDFs for the different operating conditions that existed on a route on I-8 going westbound during a number of months in 2011. It shows the cumulative distributions for the travel times during several different operating regimes. The distributions show what percentage of the time for each operating condition that the travel time was a particular value or shorter. For example, when traffic incidents occur during heavy (recurrent) congestion, one half (50%) of the travel rates (seconds per mile) are up to 70 sec/mile. That is, 50% of the travel rates are this long or *shorter/smaller*. The 90th percentile travel rate is 110 seconds per mile. Or put another way, 9 out of every 10 vehicles is traveling at that rate or faster.

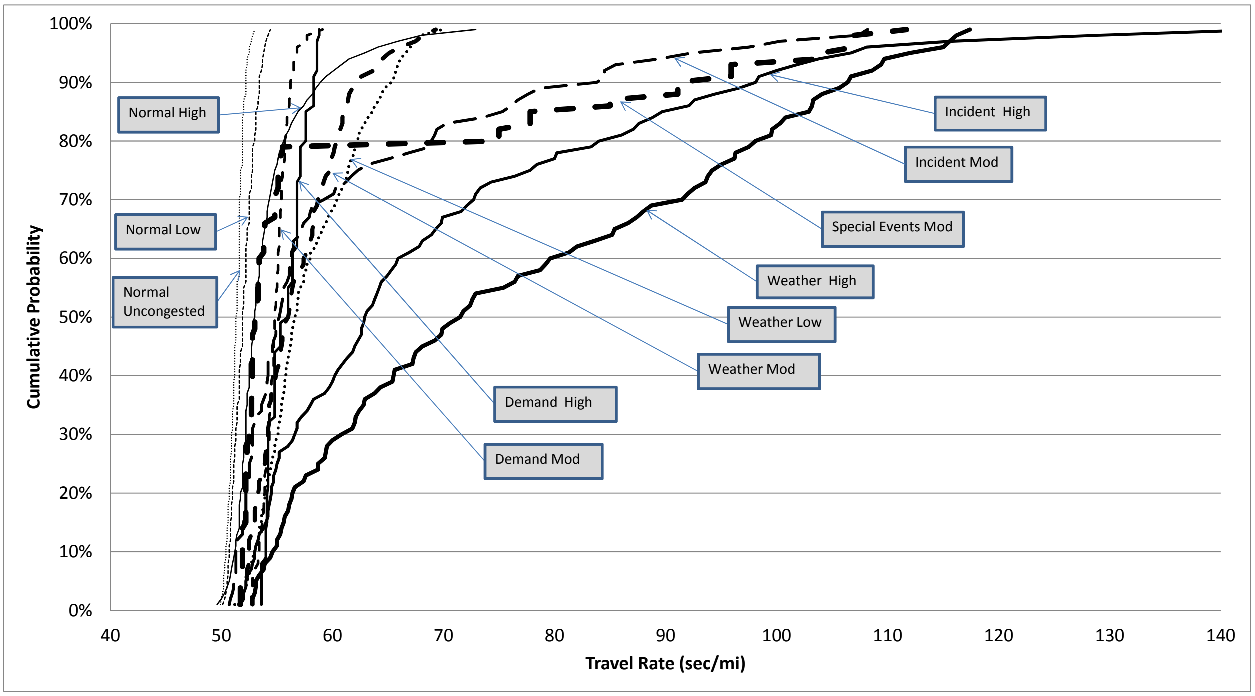


Figure ES-3: How Travel Rates Are Affected by Congestion and Non-Recurring Incidents

With a little experience, an operator can learn how to effectively compare the distributions with one another. For example, he or she can compare the distribution for high recurrent congestion *and* traffic incidents with high recurrent congestion without incidents. Without incidents, 50% of the vehicles are traveling at 58 sec/mi instead of 70 sec/mi—considerably faster. And at the 90th percentile, the difference is even more dramatic: 65 sec/mi versus 110 sec/mi. Not only does this comparison indicate that the difference between the two conditions is dramatic, but it also suggests that taking actions to mitigate these impacts would produce significant benefits in terms of improving reliability. The mitigating actions would be intended to cause the travel times (or travel rates) during incidents to get much closer to those when there are no incidents. Moreover, after the mitigating actions have been taken, the TTRMS can show how reliability improved.

To fulfill its mission as a decision support tool, the monitoring system needs to do four things as illustrated by Figure ES-4 below.

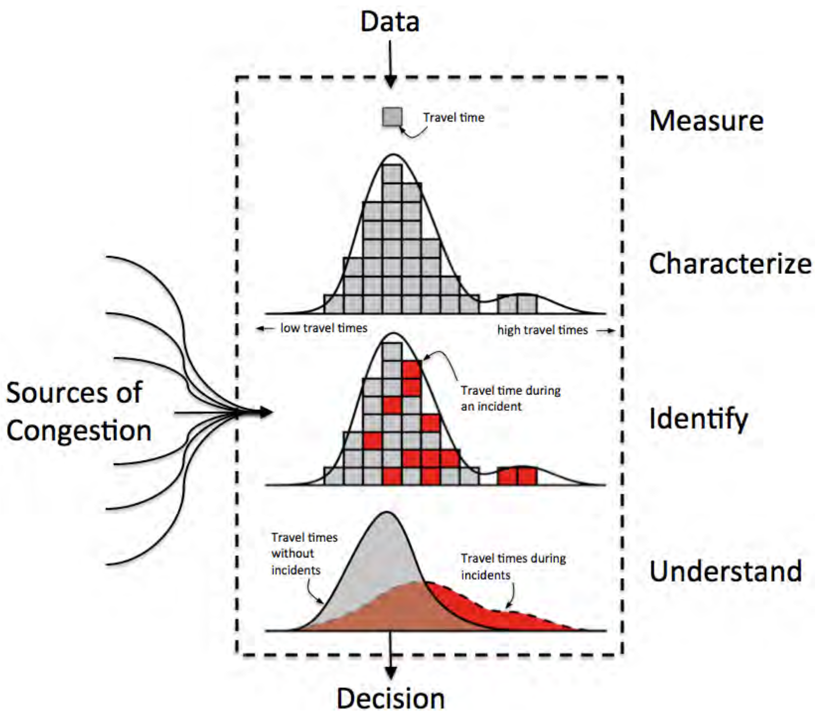


Figure ES-4: Information Flow in the Monitoring System

First, the monitoring system needs to **measure** travel times. This is a complex technical task due to the variability of traveler behavior and the plethora of different measurement sensors. Correctly *measuring* travel times along a given route requires a great deal of systems development effort and statistical knowledge. This guidebook serves as a primer on how to measure travel times, effectively, using available technologies and statistical techniques. Measuring an individual travel time is the foundational unit of analysis for reliability monitoring.

Second, the monitoring system needs to **characterize** the reliability of a given system. This is the process of taking a set of measured travel times and assembling them into a statistical model of the behavior of a given segment or route. The statistical paradigm outlined in this guidebook is that of using probability density functions to characterize the performance of a given segment or route, usually specific to a particular operating regime (a combination of congestion level and non-recurring event). This guidebook gives specific advice on the statistical decisions required to effectively characterize the travel times. Characterizing the reliability of a segment or route is fundamental to making good decisions about what to do to improve the performance of that segment or route.

Third, the monitoring system needs to **identify** the sources of unreliability. Once the reliability of a segment or route has been characterized, transportation managers need to understand what caused the unreliability (and how to "fix" it). The guidebook follows the causal list that FHWA uses to describe why congestion arises, breaking these sources into the seven major influencing factors described previously (two internal and five external, see Federal Highway Administration 2008). It discusses how to pull in data for these influencing factors and effectively fuse them with the travel time data generated in previous steps. Identifying the travel times impacted by these sources of congestion is required preparation for understanding system reliability.

Finally, the monitoring system needs to help operators **understand** the impact of these sources of unreliability on the system. This final step in turning raw data into actionable decisions requires both quantitative and qualitative methodologies: operators need clear visualizations of data, as well as quantifications. This dual approach supports both data discovery and final decision-making about a given route. Understanding reliability is the key to good decision-making about improving system reliability.

A monitoring system that accurately and consistently executes these four steps can be a powerful tool for traffic management. It enables decision makers to understand how much of their delay is due to unreliability, and prompts ideas about how to mitigate that delay. For example, it helps a freeway operator understand whether to deploy more service patrol vehicles (to clear incidents more quickly) or focus their efforts on coordinating special event traffic (to reduce delay from stadium access)? A reliability monitoring system, as outlined in this guidebook, can help an operator understand which of these activities is worth the investment, and what the payoff might be. Such systems add a new, powerful, practical traffic management tool to the arsenal of system operators.

CHAPTER 1: INTRODUCTION

Within SHRP 2, Project L02 was focused on creating a suite of methods by which transportation agencies could monitor and evaluate travel time reliability. Creation of the methods also produced an improved understanding of why and how travel times vary and the factors that create that variation.

A spectrum of future users helped shape the system. This included system operators who would want to take actions that would make the travel times more reliable and system users, like the traveling public, who would want to use the information to avoid travel delays and make sure they arrive at their destinations on time.

Project Context

Reliability is one of four focus areas that comprise the Strategic Highway Research Program (SHRP) 2, authorized by Congress in 2006. The purpose of the reliability focus area is to “reduce congestion and improve travel time reliability through incident management, response, and mitigation” (Transportation Research Board, 2012). Four themes have been established under this focus area:

- Theme 1: Data, Metrics, Analysis, and Decision Support
- Theme 2: Institutional Change, Human Behavior, and Resource Needs
- Theme 3: Incorporating Reliability into Planning, Programming, and Design
- Theme 4: Fostering Innovation to Improve Travel Time Reliability

L02 was part of the first theme, providing guidance to operating agencies about how they can put better measurement methods into practice and understand the relationship that travel time reliability has to the seven major sources of non-recurrent congestion (Cambridge Systematics *et al*. 2003 and Federal Highway Administration 2008):

- Traffic incidents,
- Work zones,
- Weather,
- Special events,
- Traffic control devices,
- Fluctuations in demand, and
- Inadequate base capacity.

Work Products

The primary work product from L02 is a Travel Time Reliability Monitoring System (TTRMS) Guidebook. It is intended to be used by operating agencies to create, operate and maintain a TTRMS. The Guidebook is a stand-alone document and is not included as part of this final report.

The purpose of this final report is to describe the process that led to the development of the guidebook: the steps that were taken, and the materials that were developed. The information is presented in an order that makes it clear why the TTRMS is designed the way it is. In addition, this first section provides an overview of the study and a guide to the final report.

Guide to the Report

Chapter 1 provides a brief narrative about what reliability is and how it can be measured and analyzed. The description emerges from the findings of researchers worldwide as well as the developments from the project. A general finding is that reliability is best described by creating holistic pictures like probability density functions (PDFs) and their associated cumulative density functions (CDFs) to portray the reliability performance of segments, routes, sub-networks, or systems. Tracking single values does not seem to be sufficient. This discussion is intended to help the reader understand why the TTRMS needed to be designed (as reflected in the functional specifications) as it was, what data it needed to collect, and how it needed to be prepared to respond to user inquiries. After reading this section, the remaining sections should seem to be logical, intuitive extensions of the ideas presented.

Chapter 2 reports the findings from a survey of the state-of-the-art and state-of-the-practice in travel time reliability monitoring systems worldwide. It shows that Europe (see, for example, Transportation Research Center 2010) and Asia had made substantial progress, while the systems in the United States were closer to being in their infancy than being mature, and that almost all of the systems in the US were focused on reporting single value statistical measures like the buffer time index. In addition, Chapter 2 summarizes a second survey that was conducted to determine the needs of TTRMS users. The needs seemed to be similar within specific groups: 1) system administrators and their staffs, 2) highway system operators, 3) transit system operators, 4) freight service providers, 5) highway system users, 6) transit system users, and 7) freight system users. Each had its own special needs with consistency being evident among the system operators (1, 2, 3, and 4) and the users (5, 6, and 7). The findings from the survey were coalesced into a set of use cases that became the drivers for the TTRMS functional specifications.

Chapter 3 describes the resulting functional specifications that were developed for the TTRMS. They were developed in response to the cases that emerged from the second survey (described in Chapter 2) and reflect an advance in the state-of-the-art that should serve the user community for several decades. The specifications focused on three main functions that the TTRMS needed to be able to perform: 1) data collection, assembly, and quality control, 2) computation of basic reliability descriptions for segments and routes in the network, and 3) responses to user requests.

Chapter 4 discusses data collection and quality enhancement activities that need to be part of the TTRMS. Most of this material is in the Guidebook and is not repeated to avoid redundancy. The main point is that high quality data must be available for a TTRMS to work effectively. Hence careful quality control on the incoming data is important. Imputation methods have value, and careful decisions are needed about how much data to archive.

Chapter 5 presents an analysis of recommended sensor spacing and sampling rates. A method that focuses on accuracy in reproducing the actual vehicle trajectories reveals that half and quarter-mile spacings have great value; and sampling rates near 30 seconds to a minute are very valuable.

Chapter 6 presents the suite of methods that were developed for the TTRMS so that it could create travel time reliability information from the data assembled. The main objective is to create probability density functions for highway segments; and from these density functions for routes. Four types of data feeds are given heavy emphasis: single loop detectors, double loop detectors, AVI-equipped vehicles and AVL-equipped vehicles.

Chapter 7 gives a summary of the validation efforts that were conducted to ensure that the TTRMS was hitting the right targets. The main aspect of this validation was a set of five case studies where prototypes of the TTRMS were put in place. The locations were San Diego, Sacramento/Lake Tahoe, Northern Virginia, Atlanta, and New York City. Each one took advantage of data feeds and data sources that happened to be available for those areas. A secondary aspect of the validation was a set of use case studies that demonstrated that the TTRMS could respond to all the various inquiries identified by the user community surveys.

Chapter 8 concludes with a summary of the findings from the project and a description of the lessons learned from the project. It is clear that much has been learned from the L02 effort and that many unanswered questions still need to be addressed.

Travel Time Reliability

This section provides an overview of the definition of travel time reliability and how it has been measured. Fundamental ways to measure travel time reliability are introduced. The intent of the discussion is to prepare the reader for the material that follows in the remaining sections of the report.

Concepts

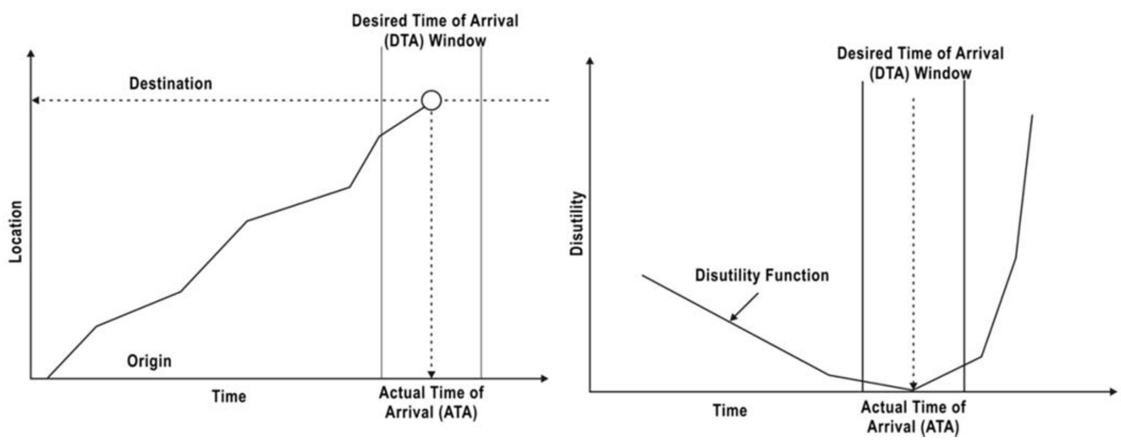
Consistent with Transportation Research Centre 2010, travel time reliability can be thought of as the absence of variability in travel times. If a system is reliable, people know how long it will take them to get to make a trip, whenever they want to leave. This might be unconditional, or dependent on something they can observe, like the weather conditions. If a freeway is perfectly reliable, then its travel time is always the same. It is either always the same under similar conditions, or more ideally it is the same regardless of the conditions that exist. These ideas are similar to the way we think about vehicle reliability. When the key is turned to the “on” position, if the vehicle always starts, then it starts reliably. In addition, when the road is dry, its performance is always the same. When the road is wet or snow covered, the performance is slightly different, but it is always the same under similar conditions.

A difference does seem to exist between the way reliability was originally defined by Ebeling (1997) and the manner in which the term is presently being used in the transportation context. As Elefteriadou and Ciu (2007) point out, Ebeling (1997) suggested that reliability should be defined as “the probability that a component or system will perform a required function for a given period of time when used under stated operating conditions. It is the probability of a non-failure over time.” This is slightly different from the idea of consistency, which has to do with the absence of variability.

Brought into the transportation network context, Ebeling's definition implies that the system would be deemed reliable (formally speaking) if each traveler or shipper experienced actual times of arrival (ATA) which matched desired times of arrival (DTA) within some window, as shown in Figure 1-1a). Depending on the utility (disutility) function that pertained to the trip, in some cases the difference between the ATA and DTA would be extremely important; in other cases it would be less so. For example, the disutility function for a trip to catch a plane would be sharply defined while the one for a trip to the store might be less so.

If the ATA lies outside the DTA window, especially if the ATA is after the DTA, a reliable trip was not completed. Hence, the transportation system is reliable, technically speaking, if the ATAs all lie within their DTA window. Otherwise, the system has "failed" or not performed reliably. As Elefteriadou and Ciu (2007) point out, such a definition of reliability becomes well defined.

In a more general sense, the reliability of the system can be measured using utility theory, as described for example by Hansson (2005) and discussed by many researchers including Vickrey (1969), Lam and Small (2001), Noland and Small (1995), and Bates et al. (2001). Utility (of the trip) is maximized if the ATA is inside the DTA window. Conversely, disutility is greater if the ATA lies outside the DTA window; and the aggregate disutility for all trips among all users is the "societal cost" of the system's unreliability. The function that evaluates the disutility may be symmetric or asymmetric depending on the situation, as shown in Figure 1-1 b). Truckers incur significant penalties if they are either late or early in delivering shipments to the receivers. Individual travelers can be late for appointments or miss the opportunity to insert additional tasks like stopping for coffee or sleeping later if they are early.



a) DTAs and ATAs **b) Disutility based on the ATA**

Figure 1-1: Basic Reliability Concepts: Desired Times of Arrival (DTA) and Actual Times of Arrival (ATA) and the associated disutility functions

If the DTA windows for trips were known today, it would be possible to assess the system reliability on the basis of the percent of ATAs that fall within their DTA windows. This would be a useful metric both for the entities making the trips as well as the organizations providing the service (e.g., the Transportation Management Center or transit system operator). The aggregate disutility could also be computed by summing the disutility values for each trip. Obviously, this world does not presently exist.

What can be observed today, at least in part, are travel times on segments and routes in the network. Some TMCs can monitor probes, vehicles equipped with tags in areas where toll roads exist, and others can generate speed distributions at specific point locations in the network where sensors (speed traps) are installed.

Implementable Ideas

To implement these ideas, the TTRMS can establish desired travel times (DTT), or better yet, desired travel rates (DTR) in seconds per mile so that the length of the facilities does not interfere; performance levels to be achieved on the segments and routes, consistent with Ebeling 1997. These DTRs can be dependent upon the regime under which the system is operating (combination of the influencing factors) and they can be adjusted over time as the network conditions change – demand grows and/or improvements are made.

A segment or route can then be deemed as performing reliably if its actual travel rate (ATR) lies within the acceptable DTR window given the regime under which the segment or route is operating. The TMC team can monitor the number of segments or routes whose ATR lies within the DTR window; they can see how that number varies based on the network, segment, or route operating conditions (e.g., an incident during high congestion); and actions can be identified to increase the number of segments or routes whose ATR is within its DTR window.

This paradigm can also be extended to the system users. Trips can be considered successful if their actual travel rate (ATR) falls within an allowable DTR window based on the conditions under which the trip was made. Reliability can be measured by the percentage of trips whose ATR fall within the allowable DTR window. By extension, the aggregate disutility experienced by the travelers or shippers can be assessed, in principle, using disutility functions which compare the ATRs one-at-a-time with their corresponding DTRs and then sums the results.

Service providers want to see if different ways to operate the system would be likely to produce better alignment between the ATRs and the DTRs (or if capacity investments are needed). Naturally, this decision making is aimed at variance reduction and shifts in the mean values either lower or higher so that the requisite confidence interval objectives are met given the DTT windows.

Decisions made by the team using the TTRMS become akin to the mean-variance tradeoff analyses so prevalent in financial planning (see, for example, Maginn et al. 2007). In this instance, the tradeoff is between minimizing the mean (or median travel times), as in building new network links or adding capacity (to reduce the mean or median travel rates), versus taking actions like improving incident response or managing the impacts of weather (see Wang *et al.* 2009, Wang *et al.* 2011, Leng *et al.* 2009, Hainen *et al.* 2012) better so that the variation in the travel rates is reduced, getting more of the ATRs within their DTR windows.

Reliability Measures

Although many reliability measures are in popular use today, like the travel time index and the buffer time index, the project team found it most fruitful to measure and assess reliability—

actually consistency—through probability density functions (PDFs). They portray the entire distribution of travel times (or travel rates) that arise across time, among vehicles, or on some other basis (e.g., among seasons).

An example of a PDF with which almost everyone seems to be familiar is exam scores, say based on qualifying exams like the professional engineer's exam. People are ranked according to their percentile position (90% have lower scores; 10% have higher scores) and the objective is to be in a top percentile. These same ideas pertain to travel time reliability. The distribution of travel times is the performance metric that should be monitored and performance improves when any or all of the percentile positions have lower travel times (or travel rates).

Probability distributions are often presented three ways (see Karr 1993, for example). The first is via a histogram, where bar heights are used to represent the relative frequency with which specific conditions pertain. Figure 1-2 shows histograms of travel times for bus route #20 in San Diego during the midday peak for various operating conditions: when everything was normal or when the system was being affected by a special event, an incident, or high demand.

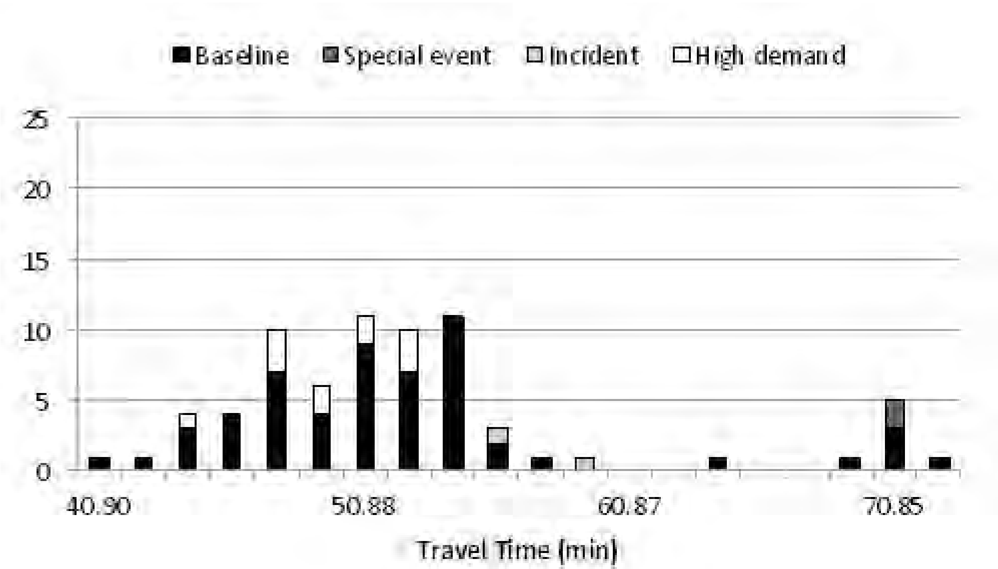


Figure 1-2: Example PDFs for Various Event Conditions

The second way to present these distributions is via a probability density function (PDF). A PDF portrays the same information as a histogram except that the bar heights have been normalized so that their sum equals 1.0 or 100% (this is the same thing as the area under the PDF equaling 1.0). Figure 1-3 shows PDFs for I-8 in San Diego under three conditions.

In the PDF, as with the histograms, it is possible to see that some travel times are more common than others, and that the distribution of the travel times is different for the various operating conditions (Barkley, *et al.* 2012, Guo *et al.* 2010, Fraley and Raftery 2009). The various common operating conditions are often called "modes" in a statistical sense, and the PDF helps the analyst to spot these various modes because they stand out as high points in the PDF.

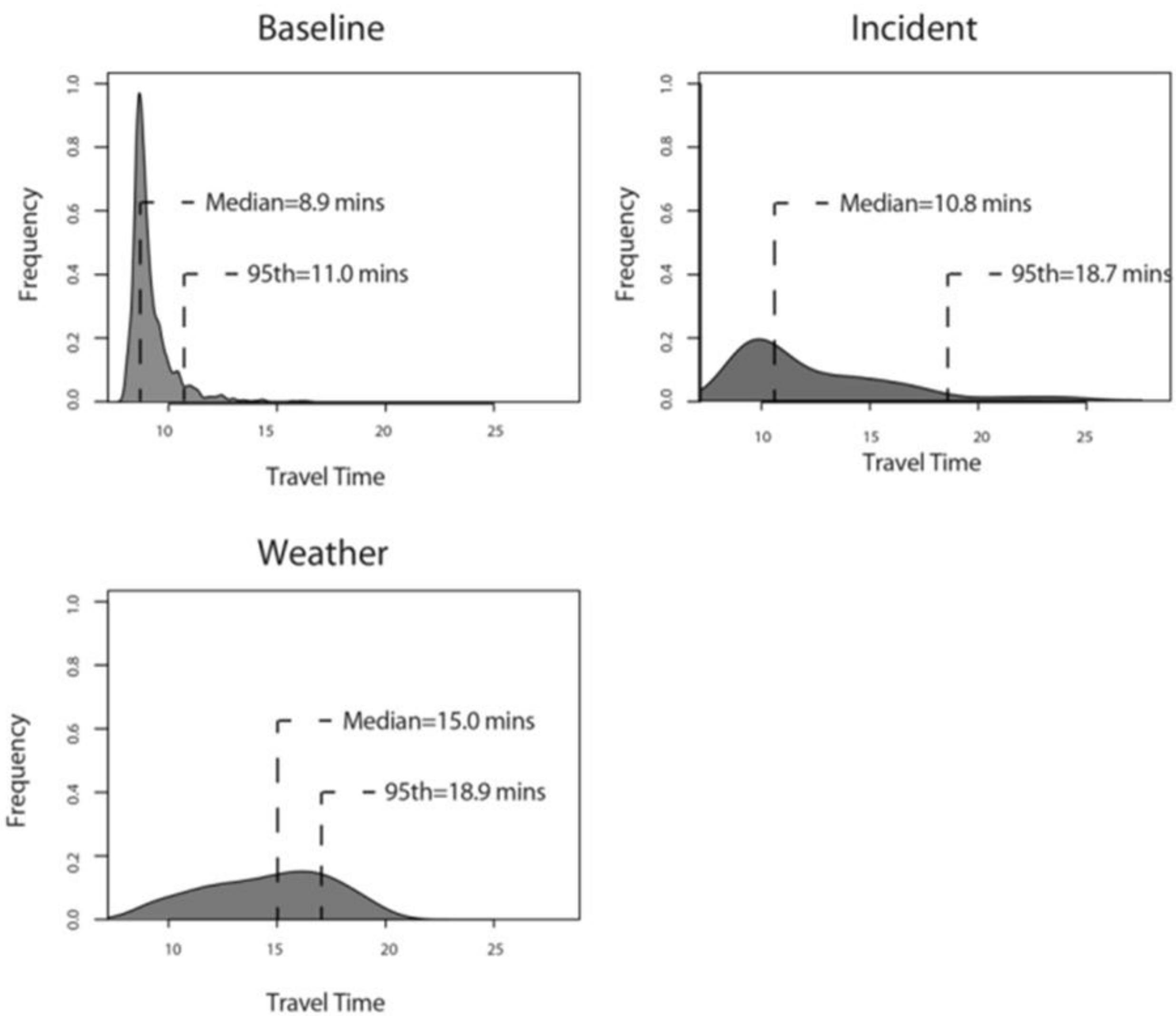


Figure 1-3: Example PDFs for Various Event Conditions

The third way to present these distributions is via a cumulative density function (CDF). The CDF is based on the PDF in that the value shown in the CDF at any point in the graph is the integral of the PDF up to that point (i.e., the area enclosed within the PDF above the horizontal axis). A property of the PDF is that its area sums to 1.0 which means the CDF ultimately rises to a maximum of 1.0. Figure 1-4 shows the CDFs for the various regimes associated with the performance of a different facility in San Diego (Interstate 5 from the junction with I-805 to the exit for 8th Street in National City). As with the PDF, one can clearly see differences in the distribution of the travel *rates* (as compared to *times*) and that the distribution of rates for some regimes is much different than for others. It is through the use of these tools—the histogram, PDF, and CDF—that the case studies and use cases reach conclusions about the influence of various factors on the travel times and travel rates.

Both Figures 1-3 and 1-4 show that the distributions of travel times are often multimodal. That is, in a statistical sense, they have several local maximums (or in the case of the CDFs, multiple inflection points where the slope gets smaller and then larger again). For individual vehicle travel times this multimodality can arise when the observations come from two different traffic streams such as HOV and non-HOV vehicles (or tagged and non-tagged vehicles), or cars and trucks where the trucks have different speed limits, even when the data are collected at the same point in time. Another example is individual vehicle travel times for an arterial, whereby some vehicles progress between traffic signals without stopping while others do not. Multimodality can also arise when the data come from different operating conditions, as in a set of average travel times for the same 5-minute time slice across a year. In fact, this multimodality should be expected. Absence of this multimodality would indicate that the operating conditions do not matter; that is, the travel time is consistent regardless of incidents, the weather, etc. At least today, this is not the case.

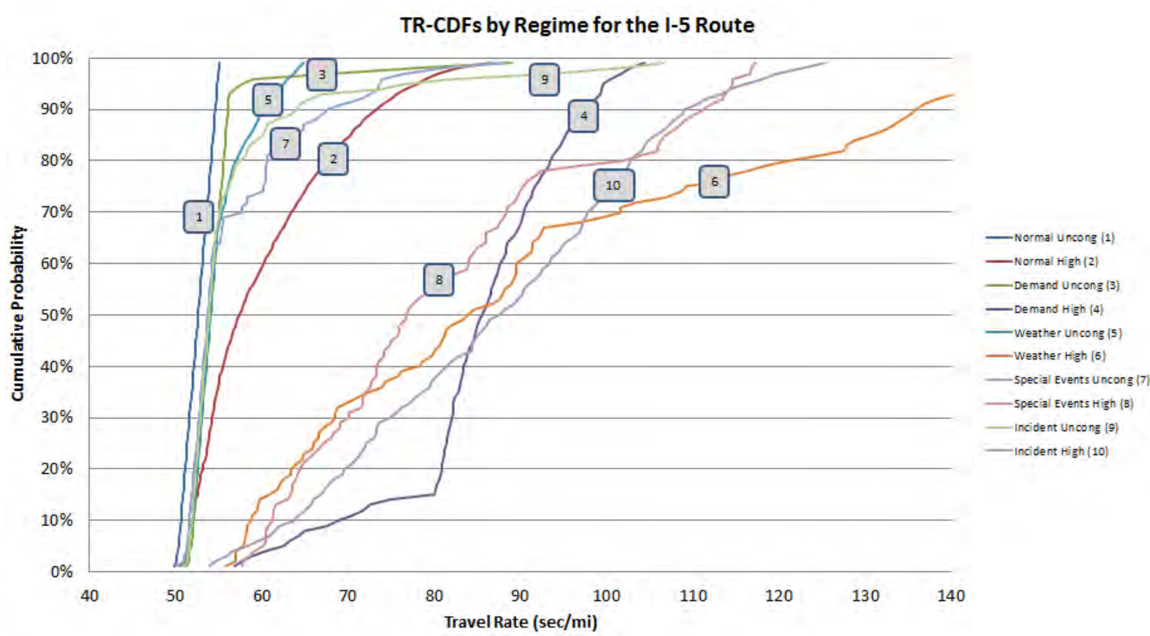


Figure 1-4: Example CDFs under Various Regimes (Operating Conditions)

Since the word "mode" is used in other ways in transportation, the word regime is used elsewhere in this document, instead of mode, to describe these various operating conditions (or or variations of a given condition). Moreover, common traffic engineering terms are used to describe these modes like "congested", "uncongested", "transition", "incident", "weather", etc. The regimes help enhance the quality of the PDFs. It keeps them from being noisy, and it helps maximize the incremental value derived from the data acquired every day.

The last concept—and an important one—is that all the reliability metrics of interest can be derived from these probability density functions (PDFs). The PDFs completely describe the travel times or travel rates (travel times per unit distance). Hence, the typical metrics of interest for characterizing reliability—planning index, buffer index, average, median, 95th percentile, or others—can be computed based on the PDFs. As a result, these PDFs, supplemented by ancillary data about the environment that does (or will exist) in the timeframe of the analysis (e.g., weather, incidents), represent sufficient information to answer the questions about measuring reliability (see also Tu *at al.* 2008).

CHAPTER 2: SURVEYS OF EXISTING SYSTEMS AND USER NEEDS

This chapter presents two components of data gathering: a survey of existing travel time reliability monitoring systems, and an assessment of user needs for information that might be produced by those systems.

Survey of Existing Systems

One task in the overall project was to determine what reliability monitoring systems already existed worldwide including in the United States. The results were current as of 2010. Of particular interest was the capability of these systems to monitor travel times and assess reliability; and how, and to what extent, this information is disseminated to various constituencies. Also of interest was 1) plans for expanding such systems, 2) the manner in which the needs of the end users are solicited and incorporated into the plans for future enhancements, and 3) the ways in which reliability data are being used to make operational, tactical, and strategic decisions about managing the performance of the system (or why this is not occurring).

It quickly became apparent that Europe and Asia were somewhat ahead of the United States in addressing the issue of travel time reliability monitoring. A good example of this from Europe is the study on improving reliability on surface transportation networks (see Transportation Research Center 2010).

Insofar as the United States is concerned, the top 25 major metropolitan areas were studied. The travel monitoring websites were visited to see what information is currently provided related to travel times and travel time reliability. Based on this investigation, a list of commercial service providers was also created, including those that operate behind the scenes as well as those that convey information to end users.

Findings

The survey showed that the collection, processing, and dissemination of traffic content has become a sizable business over the past decade. The focus of this task report is on how travel time and travel time reliability are measured and conveyed by traffic monitoring systems; hence a brief overview is given to orient the reader in the traffic content business world.

Listed in Table 2-1 are seven types of data collected for use in traffic content applications. For example, multiple companies collect incident data so they can calculate delays (due to demand exceeding capacity), disseminate delay information, and offer drivers alternate routes. One company, TransGuide, explicitly records the number of non-arriving vehicles that crossed an upstream data collection point but did not cross the downstream point of interest.

Table 2-1: Traffic Data Collection – What is Collected

	Navteq [V]	Inrix [V]	TRANSCOM [U]	TransGuide [U]	TranStar [U]
Speeds			X		
Travel times			X	X	
Number of non-arriving vehicles			X		
Incident data	X		X		X
Construction/ work zone data			X		X
Event data	X				X
Historical data	X	X	X		

V = private company
U = public agency or consortium

With available technology there are a host of possible data collection technologies, everything from probe vehicles to video cameras to loop detectors, as shown in Tables 2-2 and 2-3. Some entities specialize in a single method of data collection. For instance, TRANSCOM, TrafficGauge, AirSage, SpeedInfo, and Traffax use tag readers, proprietary devices, cellular phones, solar-powered radar sensors, and Bluetooth device MAC address readers, respectively. There are companies such as Inrix and Traffic.com that collect transportation data via multiple methods then carefully fuse the data to create a more comprehensive picture of current and future traffic conditions.

Table 2-2: Data Collection – Methods

	Navteq V	**Inrix V**	**TRANSCOM U**	**TransGuide U**	**TranStar U**	**Navteq V**	**Airsage V**
Probe vehicles							
• **EZ Pass tag readers**			X				
• **GPS fleets**	X	X					
• **phone data^**	X						
o **GPS-enabled**							X
o **Triangulation**							X
• **Bluetooth data**						X	
Proprietary sensors	X						
Government sensors	X	X*					
Incident data	X						
Event data	X						
Historical data	X	X					
Highway-embedded sensors		X					
Video monitors / cameras		X		X	X		
FM radio stations		X					
Local traffic monitoring centers		X					
Speed sensors					X		

^ often called "crowdsourcing"
* via SmartDust network
V = private company
U = public agency or consortium

Table 2-3: Data Collection by Select DOTs – Methods

	Rural		**Mid-sized**		**Large**		
	WSDOT (I-90, I-5)	Caltrans D3	Albany, NY	Orlando, FL	Los Angeles, CA	San Francisco, CA	Atlanta, GA
Probe vehicles							
• **Tag readers**		X	X	X		X	
• **GPS fleets**	snow plows						
• **phone data**						X	X (ending)
• **Bluetooth data**						X	
Highway-embedded sensors	X	Loops & WIM	X	Loops & radar	Loops & WIM	Loops & WIM	
Video monitors / cameras							X
Incident data	X		X	X			X
Event data	weather		X	X			X

Source: BTS, 2010

Whether a single method or multiple ones are implemented, there are approximately six basic steps that agencies can follow.

- Decide upon the raw data to collect. Dataset 1, dataset 2, dataset n.
- Collect the data with sensors placed in the network. The sensors can be fixed, partially mobile, or fully mobile. The sensors can be owned, bought from others, or shared.
- Record metadata (i.e., catalog information) on the sensor operations. This includes level of sensor quality, malfunctions, and level of (sensor) reliability.
- Run algorithms that combine data, choose the most reliable pieces of data (when multiple ones exist), and impute data for spots that are missing or potentially erroneous.
- Convert data into usable traffic information; augment with non-traffic information such as weather, and parking.
- Disseminate traffic information to users. Sell to other companies, give to consumers, and so forth.

Tables 2-4 and 2-5 show what information is disseminated and how it is disseminated. Everything from work zone information to weather updates is disseminated in addition to travel content. One challenge with reaching those who choose to drive is to convey enough detail about traffic conditions and route alternatives to enable drivers to make informed choices. Therefore, some companies convey general information about specific road segments, such as freeway links (i.e. 5 minute delay between Exits 2 and 3), while others cater to individuals desiring information and guidance tailored to their needs e.g., congestion ahead, take Exit 2, turn left onto Main St. to avoid).

Lastly, once a traffic content message is packaged, it can be delivered in a variety of ways, as can be seen in the following table. The most common method is via an internet website. Several providers augment a website with mobile applications (Twitter, SMS text alerts sent to cell phones) while others, especially departments of transportation, convey information directly on or near affected facilities (via variable message signs and highway advisory radio). The most personalized type is as mentioned above, guidance, specifically through navigation devices that offer real-time traffic updates. The following section covers detailed examples of traffic content dissemination using websites; however, much of this content can be conveyed using other methods discussed.

Table 2-4: Information Dissemination (Selected Content) – What is Disseminated

	Navteq [V]	**Inrix** [V]	**TRANSCOM** [U]	**TransGuide** [U]	**TranStar** [U]
Speeds	X				
Spot speeds					
Average speeds		X			
Travel times			X		
OD			X		
Path			X		
Expected					X
Average					
Personalized updates					X
Map data	X	X			X
Incident info	X	X	X	X	X
Construction/ work zone info	X		X	X	
Congestion/ flow info	X	X		X	
Weather info					X
Real-time traffic info (?)	X				X
Historical data					
To emergency personnel:					
Incident location					X
Quickest route to incident					X
Stalled vehicle locations					X

[V] = private company
[U] = public agency or consortium

Table 2-5: Information Dissemination (Selected Content) - Methods

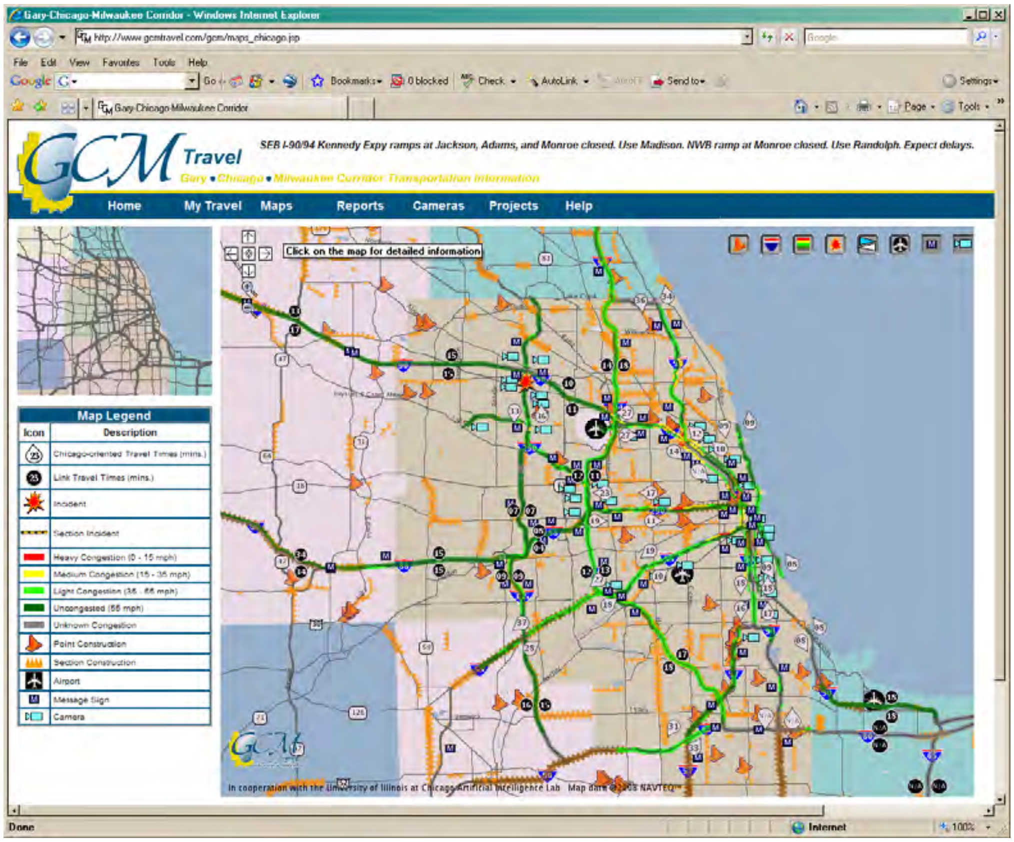
	Navteq [V]	Inrix [V]	TRANSCOM [U]	TransGuide [U]	TranStar [U]
Internet	X		X	X	X
RSS feed					X
Twitter					X
E-mail					
Cell phone / mobile alerts					X
Low-power TV stations				X	
HAR					X
DMS / VMS			X	X	X
AM/ FM radio	X				
Satellite radio	X				
Broadcast & cable TV	X				
Wireless applications	X				
GPS Navigation device for dynamic rerouting*		X			
In-car service for dynamic rerouting^		X			

* Dash Express service for example.
^ BMWs, MINIs for example.
[V] = private company
[U] = public agency or consortium

User Interfaces

Nearly every major metropolitan area in the United States has a travel time monitoring and reporting system. While there are multiple communication means that can be employed to disseminate reliability information, such as radio announcements, variable message signs, and smart phones, websites are commonly used to present information, in large part because of the amount of information that is communicated. Four or five highway-oriented traffic websites in a single metropolitan area is common (and there may be additional public transportation websites). One of these is often maintained by the government and operated in affiliation with the local transportation management center. The others are mainly private sites maintained by service providers. The websites are generally difficult to find presently when a general search engine like Google or Yahoo is used. The best way to find them is to go to the FHWA website (http://www.fhwa.dot.gov/trafficinfo/index.htm) and drill down through the sub-pages to find the website for the metropolitan area of interest.

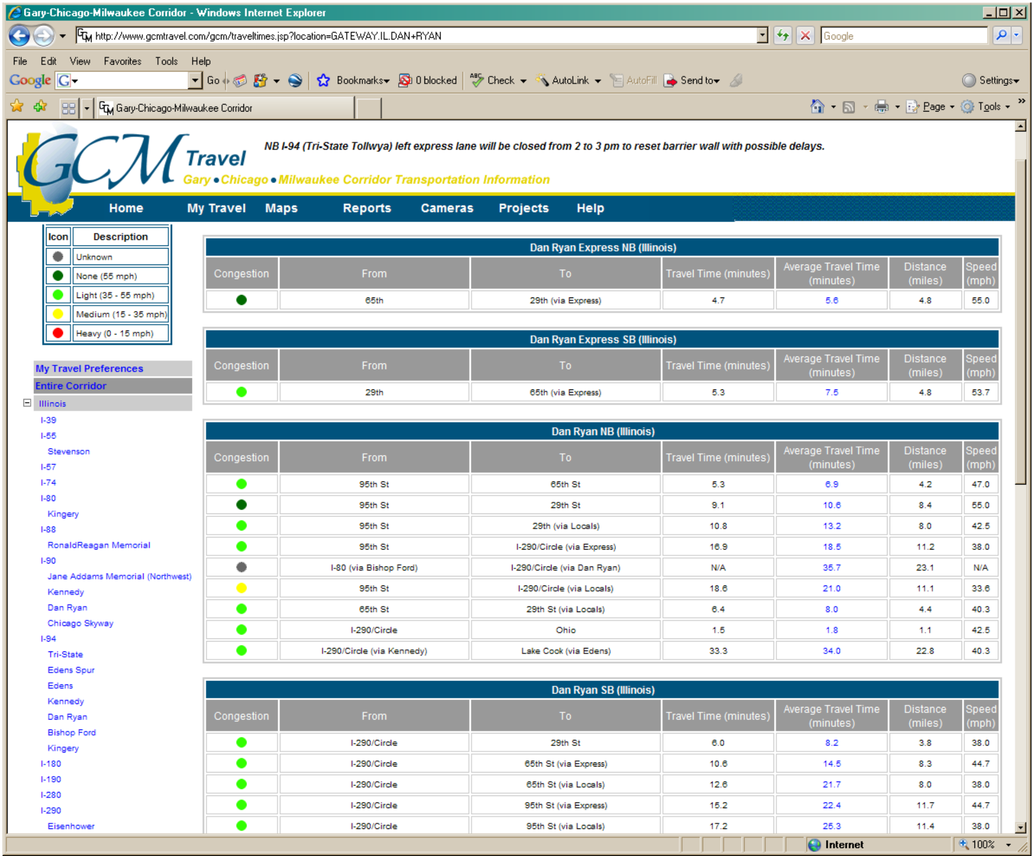
A good example of a public website is the one for the Chicago area. The main map page is shown in Figure 2-1. Color-coded maps are common, with the colors depicting speeds on individual highway segments, periodically updated. Incidents and construction areas are also almost always shown along with other significant landmarks, like airports.



Source: www.travelmidwest.com (7) accessed on 6/22/2009

Figure 2-1: Traffic Speeds Map for the Greater Chicago Area

The maps are sometimes supplemented by tables, like the one shown in Figure 2-2 that depicts travel times, speeds, and distances for the instrumented highways. In this case, the information includes current travel time, average travel time, distance, and current average speed. The speeds and travel times currently come from point sensors. (See the Inrix discussion for more details about other sensors.) The level of congestion is also identified with a green, yellow, or red dot, except for the segments that are not instrumented.

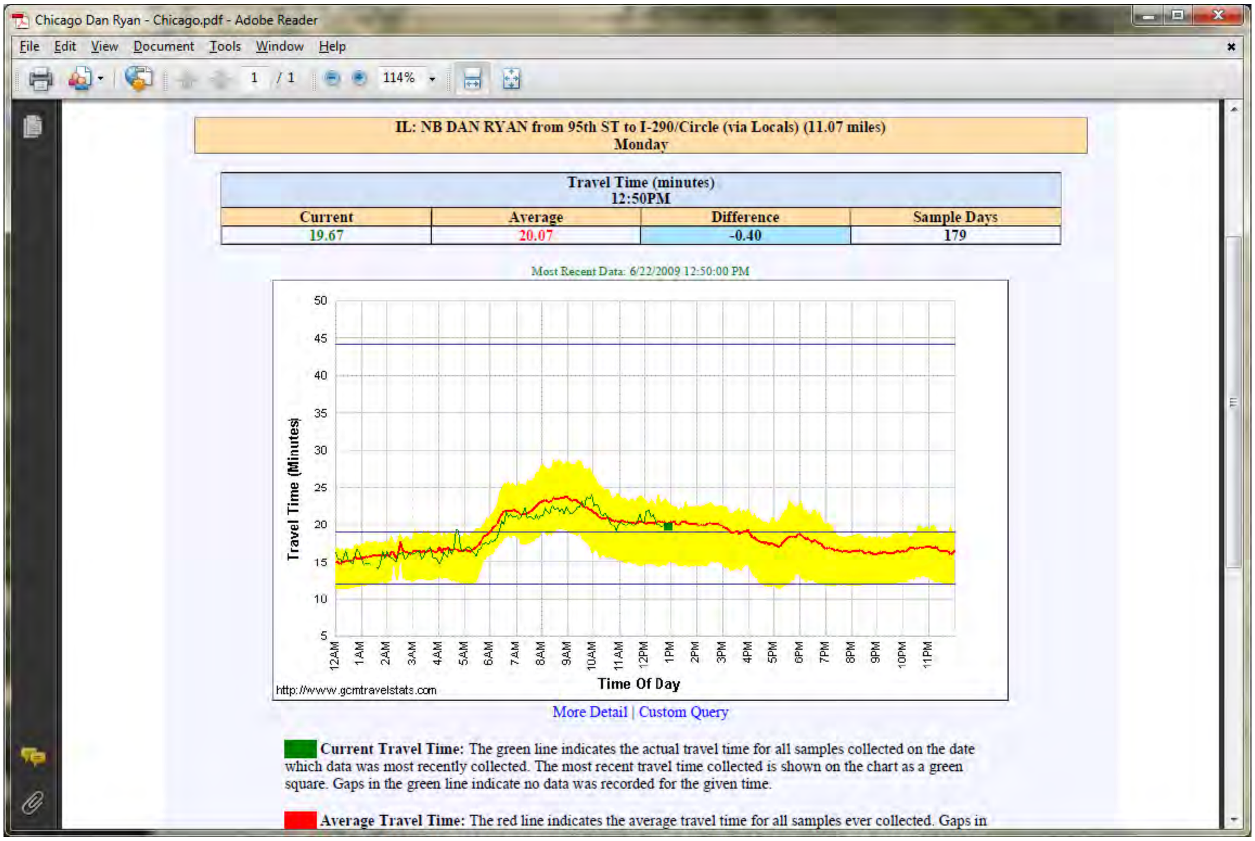


Source: www.travelmidwest.com (7) accessed on 6/22/2009

Figure 2-2: Current Congestion and Travel Times for a Freeway Segment

For this Chicago website, drilling down into the average travel time field yields a more detailed picture, and one that is useful in terms of travel time reliability. Figure 2-3 shows that for this freeway segment and direction, the current travel time is 10.88 minutes, the average is 13.17, the difference is -2.29 minutes, and the average is based on 186 sample days. The time-of-day trend shows high travel times in the AM peak that start to rise about 5:00 AM and return to nominal night-time, free-flow conditions by about 3:00 PM. On the day when the website was visited (7/28/09), unlike most days, there was a major spike in travel time at 2:30 PM, most likely caused by an incident. The yellow band shows the normal range of travel times (apparently plus or minus one standard deviation as evidenced by the reference to 68%) and the blue lines indicate travel times at free flow speed (55 MPH), medium traffic congestion (35 MPH), and heavy congestion (15 MPH).

These graphs provide travel time reliability information, that is probably the reason they were created, but there is no evidence that they are being integrated into the travel time information presented in the higher-level maps. Nor is there a quick, easy, and obvious way to reach these reliability graphs from the travel time map.



Source: www.travelmidwest.com (7) accessed on 6/22/2009

Figure 2-3: Travel Time Reliability Trends for a Freeway Segment

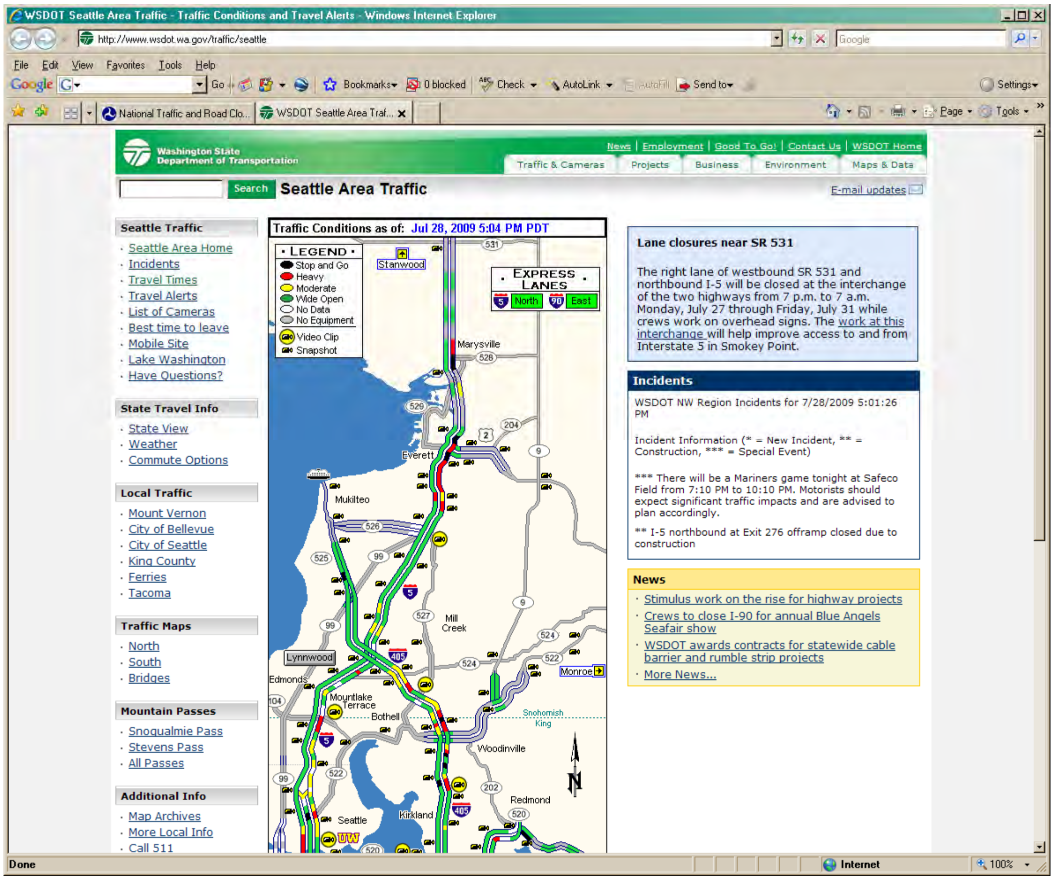
Table 2-6 summarizes the characteristics of the travel time websites in each of the 25 largest metropolitan areas in the US. Many of them have a website that was developed locally and is maintained by a local staff. For example, this is true for Chicago, Dallas-Fort Worth, Atlanta, and Detroit. Others, including those in California and Texas, use a website shell that was created for statewide use. Some cities use commercial providers, e.g., Philadelphia, Washington D.C., Pittsburgh, and Sacramento.

It is also apparent that some of these websites are partnerships, with one entity maintaining the website and another, in the background, doing the data assembly and data processing. A good example is New York City where BeatTheTraffic is a commercial vendor providing travel time information and Inrix is responsible for assembling and processing the data behind the scenes.

Table 2-6: Travel Time Information for the Top 25 Metropolitan Areas

Rank	Metropolitan Area	Population	Website(s) and Features
1	New York - Northern New Jersey - Long Island	19,006,798	http://www.trips123.com/traffic_main.asp
	Edison, NJ	2,325,224	This component of the Trips123.com website is currently under construction and is coming soon.
	Nassau-Suffolk, NY	2,863,849	Event list does exist.
	Newark-Union, NJ	2,121,076	http://www.beatthetraffic.com/ajax/traffic/map.sapx?regionid=15&viewname=New+York+City (Inrix)
	New York - White Plains - Wayne, NY-NJ	11,696,649	Color-coded speed maps, travel times on segments
2	Los Angeles - Long Beach - Santa Ana, CA	12,872,808	http://caltrans511.dot.ca.gov/
	Los Angeles - Long Beach - Glendale, CA	9,862,049	http://map.commuteview.net/CommunityView/html/es_main.html?7
	Santa Ana - Anaheim - Irvine, CA	3,010,759	Travel times, average speeds, time of update
3	Chicago, IL-IN-WI	9,569,624	http://gcmtravel.com/gcm/maps_chicago.jsp
	Chicago - Naperville - Joliet, IL	7,990,248	Current travel time, average travel time, average speed
	Gary, IN	702,458	
	Lake County - Kenosha County, IL-WI	876,918	
4	Dallas - Fort Worth - Arlington, TX	6,300,006	http://www.trans-vision.org
	Dallas - Plano – Irving, TX	4,226,003	Shows speeds with colors for ranges
	Fort Worth – Arlington, TX	2,074,003	
5	Philadelphia - Camden - Wilmington, PA-NJ-DE-MD	5,838,471	http://www.traffic.com/controller/myTraffic
	Camden, NJ	1,250,569	Gives route, travel time at speed limit, current travel time, delay, average speed
	Philadelphia, PA	3,892,194	
	Wilmington, DE-MD-NJ	695,708	
6	Houston - Sugar Land - Baytown, TX	5,728,143	http://traffic.houstontranstar.org/layers/
			Speed map, speed charts for specific segments, build route
7	Miami - Fort Lauderdale - Pompano Beach, FL	5,414,772	http://www.beatthetraffic.com/traffic/map.aspx?regionid=27&viewname=Miami
	Fort Lauderdale - Pompano Beach - Deerfield Beach, FL	1,751,234	Distance, sensor %, current trip time, ideal trip time, average speed
	Miami - Miami Beach - Kendall, FL	2,398,245	http://www.traffic.com/Miami-Traffic/Miami-Traffic-Reports.html
	West Palm Beach - Boca Raton - Boynton Beach, FL	1,265,293	Gives route, travel time at speed limit, current travel time, delay, average speed
8	Atlanta - Sandy Springs - Marietta, GA	5,376,285	http://www.georgianavigator.com/perl/trips
			Point-to-point travel times by highway, not chained
9	Washington - Arlington - Alexandria, DC-VA-MD-WV	5,358,130	http://traffic.yahoo.com/maps_result?csz=washington,_DC&country=us&trf=1
	Bethesda - Gaithersburg - Frederick	1,176,401	Color coded map for speeds, plus incidents; path distance and travel time, but not delay
	Washington - Arlington - Alexandria, DC-VA-MD-WV	4,181,729	
10	Boston - Cambridge - Quincy, MA-NH	4,522,858	http://www.boston.com/traffic
	Boston - Quincy, MA	1,884,659	(boston globe) limited "slowness" qualitative
	Cambridge - Newton - Farmington, MA	1,482,478	http://www.beatthetraffic.com/traffic/map.aspx?regionid=52&viewname=Boston
	Peabody, MA	736,457	Powered by Inrix
	Rockingham County - Stafford County, NH	419,264	http://www.smarttraveler.com/scripts/bosmap.asp?city=bos&cityname=Boston
			Travel times and updating timestamp (seems to be current time), table of travel conditions
11	Detroit - Warren - Livonia, MI	4,425,110	http://mdotwas1.mdot.state.mi.us/public/drive/rtt.cfm
	Detroit - Livonia - Dearborn, MI	1,949,929	Color-coded congestion map, speed brackets, table of average speed trends for some locations
	Warren - Troy - Farmington, MI	2,475,181	
12	Phoenix - Mesa - Scottsdale, AZ	4,281,899	http://www.az511.com/RoadwayConditions/index.php
			Color-coded average speeds on major links
13	San Francisco - Oakland - Fremont, CA	4,274,531	http://traffic.511.org/traffic_map.asp?
	Oakland - Fremont - Hayward, CA	2,504,071	Travel times, average speeds, time of update
	San Francisco - San Mateo - Redwood City, CA	1,770,460	Predict-a-trip, different routes, but not reliability
14	Riverside - San Bernadino - Ontario, CA	4,115,871	See Los Angeles
15	Seattle - Tacoma - Bellevue, WA	3,344,813	http://www.wsdot.wa.gov/traffic/seattle
	Seattle - Bellevue - Everett, WA	2,559,174	Color-coded speed map, average and current travel times by highway segment
	Tacoma	785,639	95% TT estimator, also found HOV performance comparisons
16	Minneapolis - St. Paul - Bloomington, MN-WI	3,229,878	http://www.dot.state.mn.us/tmc/trafficinfo/map/refreshmap.html
			Color-coded speed map
17	San Diego - Carlsbad - San Marcos, CA	3,001,072	http://www.dot.ca.gov/dist11/d11tmc/sdmap/showmap.php
			Shows speeds
18	St. Louis, MO-IL	2,816,710	http://www.traffic.com/St-Louis-Traffic/St-Louis-Traffic-roads.html?AWOPARTNER=GATEWAYGUIDE
19	Tampa - St. Petersburg - Clearwater, FL	2,733,761	http://www.511tampabay.com
			Color-coded speed map, estimated travel times
20	Baltimore - Towson, MD	2,667,117	http://www.chart.state.md.us/travinfo/speedData.asp
			Color-coded map, table of current speeds
21	Denver - Aurora, CO	2,506,626	http://www.cotrip.org/speed.htm
			Color-coded speed map, travel times
22	Pittsburgh, PA	2,351,192	Commercial vendors, color coded maps, travel times
23	Portland - Vancouver - Beaverton, OR-WA	2,207,462	http://www.tripcheck.com
			Color-coded speed map, delay indicators
24	Cincinnati - Middletown, OH	2,155,137	http://www.artimis.org
			Color-coded speed map, normal times, current times, delays
25	Sacramento, CA	2,109,832	http://BeatTheTraffic.com

A travel time website that directly addresses travel time reliability (really consistency) is the one used in Seattle. While the color-coded map of traffic conditions looks typical of most sites, as shown in Figure 2-4, there are lower levels that provide additional detail.

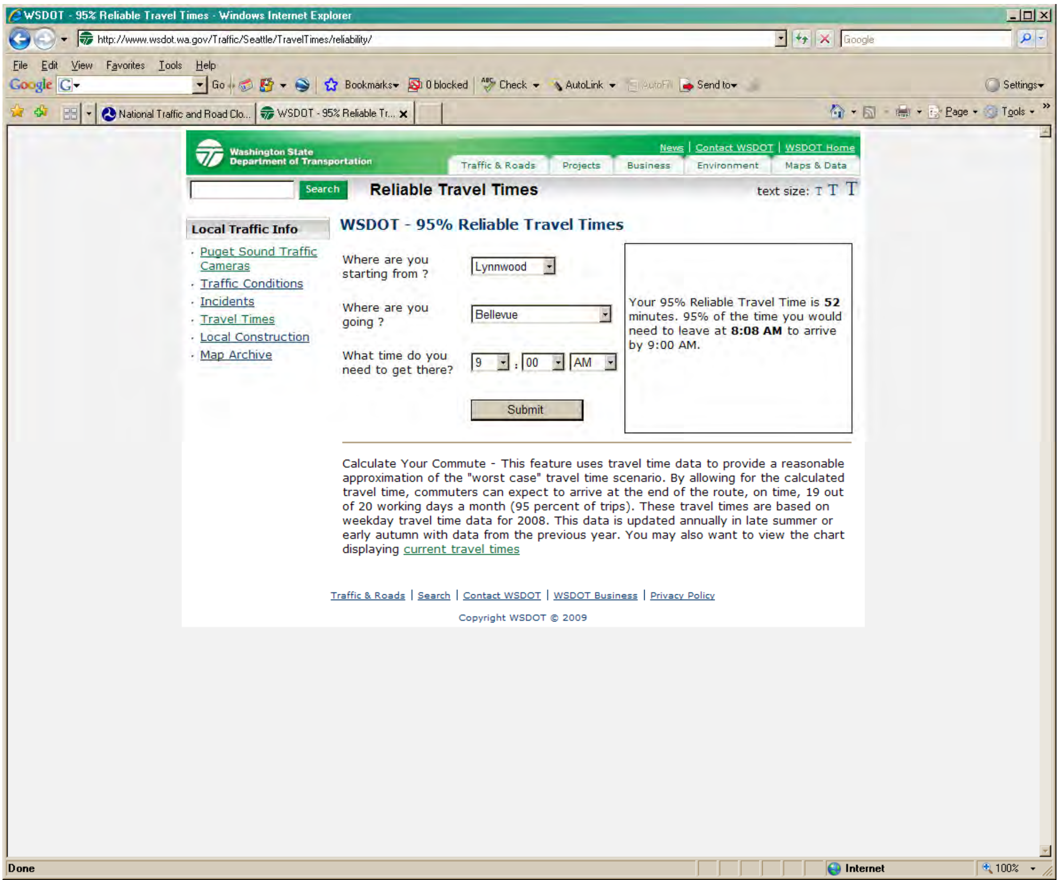


Source: www.wsdot.com/traffic/seattle/default.aspx, accessed on 6/22/2009

Figure 2-4: Seattle Area Traffic Conditons Map

Clicking on the “Best time to leave” hotlink on the lefthand side leads in two clicks to the tool shown in Figure 2-5. It allows the traveler to specify an origin and a destination and receive an estimate of the time one needs to allow to ensure that for 19 out of 20 trips (95 percent of the time) the destination will be reached on time. In the example window, a trip from Lynnwood to Bellevue is to be completed by 9:00 AM. The website reports back that the traveler needs to leave Lynnwood at 8:08 AM and allow 52 minutes for the trip to ensure that the destination will be reached by 9:00 AM. However, it should be noted that the resultant text shown in Figure 2-6 can be confusing or misleading to the average driver. It states in the dialogue box, “Your 95% Reliable Travel Time is **52** minutes. 95% of the time you would need to leave at **8:08 AM** to arrive by 9:00 AM.” The WSDOT text may be misinterpreted to mean that if you leave after 8:08 AM, then 95% of the time you will be late. This potential confusion due to the verbiage chosen points to the need for a standard and the crucial role of the L14 project to determine the best lexicon to convey reliability thoughts to various user groups (in this case, drivers).

No other transportation website was found to provide this functional capability. Querying the website regarding “reliability” leads to this webpage and several others. One of interest is shown in Figure 3-5. It talks about reliability in the context of the difference between travel times in the high-occupancy vehicle (HOV) lanes and the regular (mixed flow) freeway links and asserts that this is a "reliability" result In a sense, it is true that restricted use lanes have more reliable travel times, and if more people use them, reliability improves. The HOV lanes are likely to have more reliable travel times because there is less traffic. The travel times during peak hours are much lower as well, which one would expect.



Source: www.wsdot.com/traffic/seattle/default.aspx, accessed on 6/22/2009

Figure 2-5: 95% Reliable Travel Time Calculator

Other websites can provide travel times for specific trips, but none specifically addresses reliability-adjusted travel times in a direct manner.

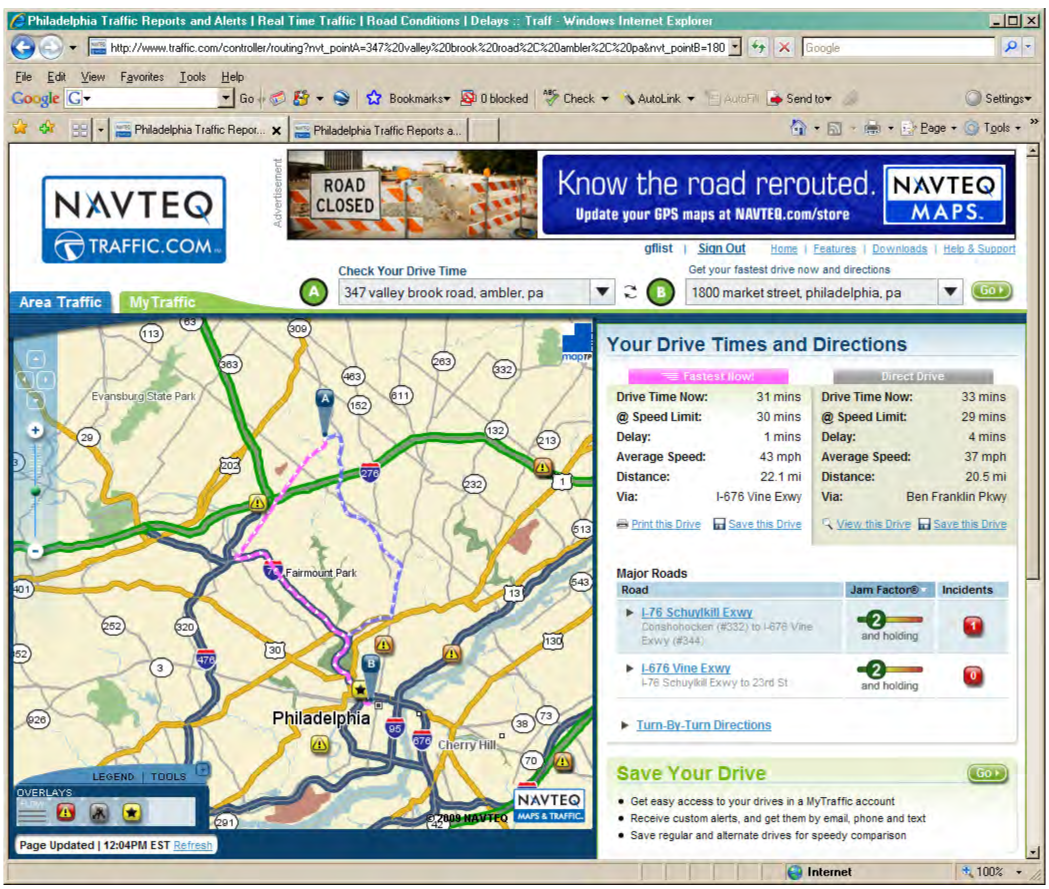
The "common" set of reliability measures (see http://ops.fhwa.dot.gov/publications/tt_reliability/ for example) discussed in the literature are:

- Buffer Index: Computed as the difference between the 95th percentile travel time and the average travel time, normalized by the average travel time
- Planning Time Index: Computed as the 95th percentile travel time index divided by the free-flow travel time index
- Skew Statistic: Computed as the ratio of (90th percentile travel time minus the median) divided by (the median minus the 10th percentile)

- Misery Index: Computed as the difference between the average of the travel times for the (0.5-5) percent longest trips and the average travel time, normalized by the average travel time (useful primarily for rural conditions)

- Failure/On-Time Measure: Computed as the percent of trips with travel times less than a threshold (Calibrated Factor (e.g., 1.3) * Mean Travel Time)

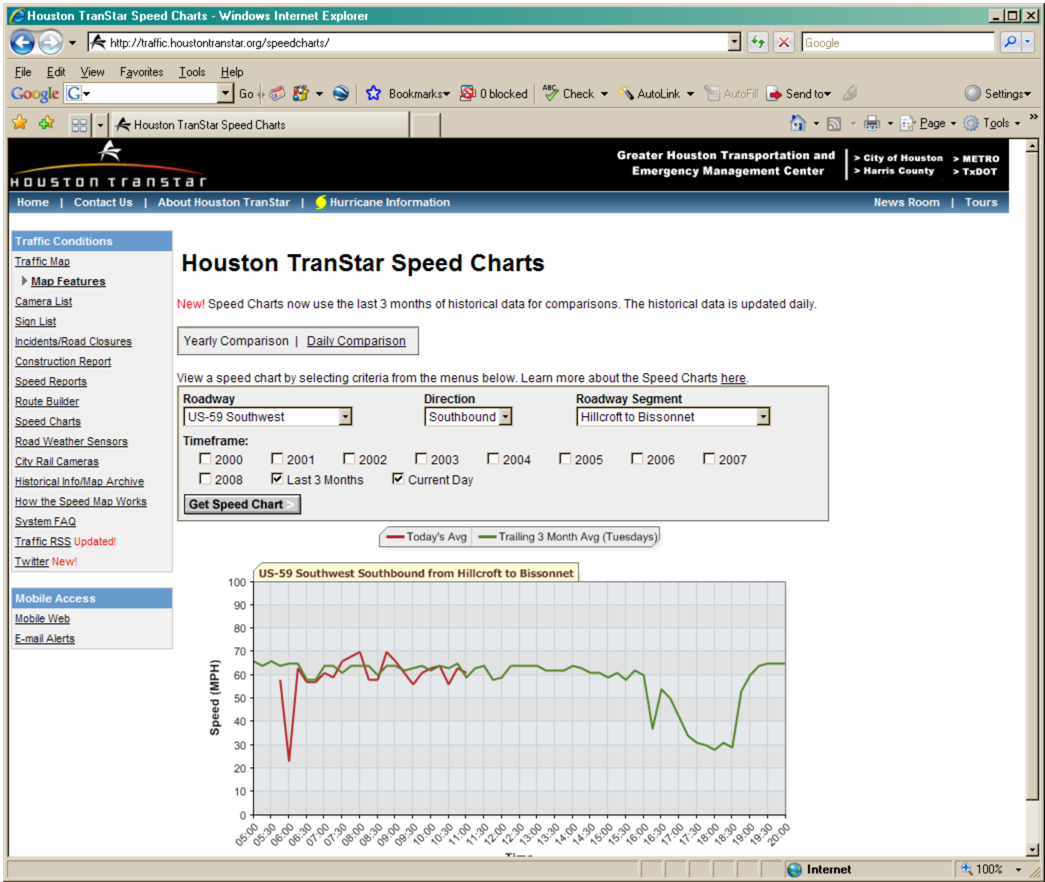
It is not obvious that any of the current developers of websites are actively pursuing the use of these measures on their websites. The limit seems to be travel time trends and comparisons of current travel times with averages. An example would be Traffic.com's function where the traveler can get directions and driving times for one or more routes, including the current level of delay, shown in Figure 2-6. It is not clear that the "delay" is intended to provide a particular likelihood that the trip will be completed in the time listed, but the delay is being estimated and reported.



Source: www.traffic.com, accessed on 6/22/2009

Figure 2-6: An Example of Conveying Travel Time Trends

The Houston TranStar website provides speed charts for specific freeway links as shown in Figure 2-7. The average from the current day (shown in red) is compared with the trailing 3-month average based on the day of the week (shown in green). In the case of the specific link queried, there was a significant drop in speed early in the morning that was strikingly different from the 3-month average.



Source: http://traffic.houstontranstar.org/speedcharts/, accessed on 6/22/2009

Figure 2-7: A Speed Chart for a Link in the Houston Network

A number of commercial companies provide travel time information for metropolitan areas throughout the country. The most common ones seem to be: Traffic.com, BeatTheTraffic, Iteris, TrafficGauge, traffic.yahoo.com, and SmarTraveler.

The format of the information provided by these companies is virtually the same for all locations, populated, of course, with local information. The credit given for the source of the travel time information varies by location. More comprehensive data can be found in the Appendices about these and other companies.

The list of companies assembling and processing the travel time information is more difficult to discern, but a partial list includes: PeMS, especially for California and along the West Coast; Inrix, especially on the East Coast; OpenRoads, especially in Virginia; Iteris; and Highway Information Systems, especially in North Carolina.

Websites are able to communicate a large amount of information that is useful to a traveler before beginning a trip. Other methods of communicating travel time reliability, such as radio announcements, VMS signs, and Smart Phones, are useful for travelers who do not have access to a computer, particularly once they are en route to their destination.

Radio broadcasts usually provide a range of expected travel times or compare the current travel times to a normal condition. For example, a broadcast might say the travel time between two locations is three minutes longer than the average travel time.

VMS disseminate similar types of travel time reliability information as radio broadcasts. While radio broadcasts are accessible throughout the service area of a particular radio station, VMS are permanently located on specific roadways in the network. Some agencies, such as the Maryland Department of Transportation, also show the messages that are currently being displayed on their VMS on their website. Travelers can also access travel time and reliability information through applications on Smart Phones. For example, the Google Maps application for Blackberry phones has a Traffic option that shows the relative speeds on major roadways based on either current traffic conditions or historic data.

Assessment of User Needs

A functioning reliability monitoring system must meet the needs of many different types of users. That is because different users perceive and value deviations from the expected travel time in different ways. In this research, users are classified into the following broad groups:

- Passenger travelers,
- Freight movers,
- Policy makers,
- Roadway system managers, and
- Transit system managers.

Understanding the differences in needs is fundamental to laying the framework for an effective monitoring system (see also Xiong *at al.* 2007). Passenger travelers think about reliability in terms of either: 1) deviation in relation to the total trip time, or 2) how often they are able to arrive within a particular time window. Freight movers think about reliability in terms of whether trips are taking longer or shorter than expected (Morris *et al*. 1998). Policy makers are typically performing high-level evaluations of "output" measures and responding to concerns as to whether their agency is meeting expectations and satisfying benchmarks. System managers are directly responsible for protecting and improving reliability on their network and are most affected by the issues limiting the effectiveness of an agency in providing reliable travel.

There are several factors, internal to the system, which are associated with time-varying relationships between demand and capacity as well as roadway incidents. External factors include weather, special events, and infrastructure failures, and performance of complementary and competing modes. Travelers, operators (carriers), shippers, and other network users gather and use reliability information in travel and shipping decisions. What they learn affects departure times, mode choice, path choice, and even destination and location choices. Businesses and families make some location decisions (residential and work location choices) partly based on expected network reliability. This indicates that reliability information is useful in a variety of different time frames, from near-real time to long term decisions and trends.

Different users need and use different kinds of information on system reliability. Highway managers need technical, quantitative information, both (near) real-time data for operations management and archived historical trend data for strategic and investment planning. Travelers use qualitative, anecdotal and objective, quantitative information on reliability for trip planning.

Users of reliability information receive it from two original sources – direct experience and reports gathered through organized monitoring processes. Information moves in complicated ways. Anecdotal (experiential reports) may move into formal monitoring systems through word of mouth to travelers or locators, through the media to the community and policy makers, etc.

Improved reliability data – data that are accurate, timely, and comprehensive – help facilitate better decisions by all users. These data will (1) help travelers get the best use out of the network, (2) help managers improve reliability, and (3) guide decision makers to using more cost-effective measures that enhance and protect system reliability.

The SHRP 2 Project L11, "Evaluating Alternative Operations Strategies to Improve Travel Time Reliability," conducted an extensive literature review, including traveler and shipper modeling efforts that have explicitly used or tried to use reliability measures. These are important because they represent empirical, analytic tests of the relationships between the behavior of travelers and shippers and reliability. Such modeling studies can tell us what kinds of reliability measures are associated with traveler and shipper behaviors. Such associations are not necessarily indications of causality, but by combining these results with the strong understanding of traveler and shipper behaviors gained in the interview process, we begin to provide a useful basis for sorting out the most important reliability measures from the user perspective.

At the time of L02, L11 was an on-going project intended to provide both a short-term and long-term perspective on innovative ideas leading to practical tools that can be implemented on a system to improve the travel time reliability of that system. L11 emphasized travel time reliability from the standpoint of the everyday users including those engaged in freight as well as passenger transport in both urban and rural areas.

Tasks 1 through 3 of L11 focused on a review of the current and future travel time reliability needs of the users and identified goals for improving reliability. Users, in this case, was a broadly defined group encompassing the following subgroups:

- Passenger travelers
- Freight movers
- Policy makers
- Roadway system managers
- Transit system managers

The user needs identified by L11 were summarized in "SHRP 2 L11 Technical Memorandum 1" and "SHRP 2 L11 Technical Memorandum 2." Travel time reliability is defined in these memoranda as the *variation in travel time for the same trip from day-to-day*. Through focus group interviews, L11 found that most roadway users have the same "desire" for reliable roadway performance—"free flow travel all of the time"—but that this is unrealistic and that users instead plan their lives and businesses around “expected” conditions. Therefore, “unexpected” conditions can degrade the user’s confidence in the overall reliability of the system and can increase costs of travel.

It is important to note that embedded within the work of L11 was a more comprehensive literature review of travel time reliability user needs, including the following preceding efforts:

- *Guide to Effective Freeway Performance Measurement* (NCHRP Research Results Digest 312)
- *Cost-Effective Performance Measures for Travel Time Delay, Variation, and Reliability* (NCHRP Report 618)
- *Identification and Analysis of Best Practices* (SHRP 2 Project L01)
- *Analytic Procedures for Determining the Impacts of Reliability Mitigation Strategies* (SHRP 2 Project L03)
- *Institutional Architectures to Advance Operational Strategies* (SHRP 2 Project L06)
- *Archive for Reliability and Related Data* (SHRP 2 Project L13)
- *Measuring Performance Among State DOTs* (AASHTO)
- *Statewide Incident Reporting Systems* (NCHRP 20-07 – Task 215)
- *Guide to Benchmarking Operational Performance Measures* (NCHRP 20-07 –Task 202)
- *Traffic Incident Management Self-Assessment National Executive Summary Report* (FHWA)
- *Freight Data from Intelligent Transportation System Devices* (Washington State Department of Transportation)

The L02 project team coordinated its efforts with L13 and L14, “Traveler Information and Travel Time Reliability”. L14 started in 2010 and was slated to take two years. Its focus was on identifying the right combination of words, numbers, symbols, etc. to communicate information about travel time reliability to travelers. While the topics are very similar, the focus of L02 and L14 are unique: L02 is focused on data monitoring, and L14 is focused on communicating information to travelers. The interface between the two projects lies in the exchange of information that is gathered from the monitoring system developed in L02 with the communication strategy recommended in L14. It was critical that the information needs for the messages developed in L14 could be accommodated through the performance measures collected/imputed/calculated as part of L02. At the time of this final report, it did not appear that the L14 project would identify performance measures that required data not already collected or calculated as part of L02.

The Needs of Passenger Travelers and Freight Movers

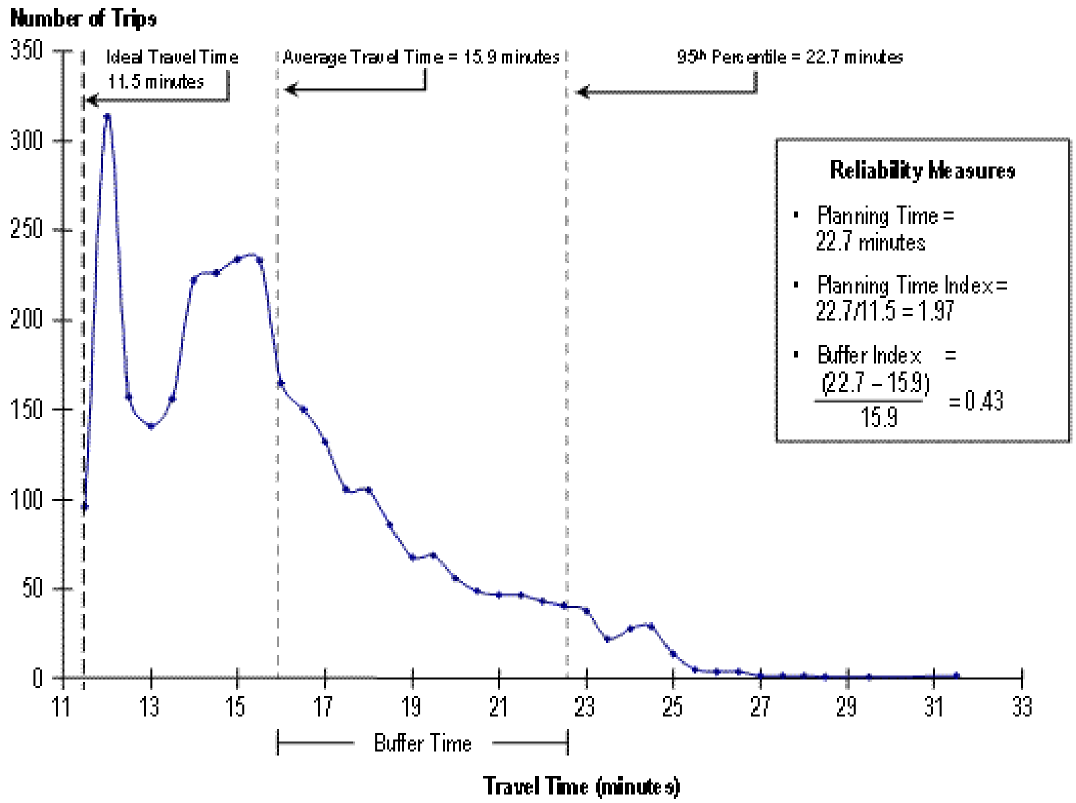
The needs of highway users, in relation to travel time reliability and the factors influencing reliability, were identified for both passenger travelers and freight movers by the L11 project team.

Passenger Travelers

The group of users comprising passenger travelers represents individual vehicle users that drive to work, recreational centers of activity, school, or other types of individual destinations. As Khattak *et al.* 1994, Carrion and Levinson 2010, Tilahun and Levinson 2010, Small *et al.* 2005, Fosgerau and Engleson 2011,Fosgerau and Karlstrom 2010, Jenelius *et al.* 2011, and Batley and Ibanez 2009, Higati *et al.* 2009) demonstrate, individual travelers are known to highly value travel time reliability (especially information about unexpected events), as they can save travel time and avoid schedule delays, i.e., late arrival at destination. In addition, Khattak *et al.* 2003 point out that a substantial portion of travelers seem willing to pay for personalized dynamic information.

The passenger-traveler focus-group interviews indicated that if the travel time actually experienced matches the expectation of the passenger traveler, then the travel time is considered reliable. Deviations from the expected travel time are perceived differently by users according to the context of their trip. For example, travel time deviations when travelers undertake non-discretionary work trips are considered more onerous than deviations on discretionary trips. Therefore, the performance measure(s) used to describe reliability were selected based upon a trip's purpose and the frequency and flexibility of that trip.

Four reliability measures were identified based on the defined user categories for passenger travelers and freight movers. Figure 2-8 provides an illustration of these measures, and how they are calculated, on a travel time chart. (For comparisons of various possible measures see Lomax *et al*. 2003 and Pu 2011.)



Source: (Cambridge Systematics and Texas Transportation Institute 2005), available at http://ops.fhwa.dot.gov/congestion_report/.

Figure 2-8: Graphical Illustration of Reliability Measures

- Planning Time (95th Percentile Travel Time) - Average trip duration in minutes and seconds for 95 percent or less of all trips. This measure estimates the extent of delay during the heaviest traffic days.
- Buffer Index - The difference between the 95th percentile travel time and the average travel time, divided by the average travel time. This represents the extra time (in minutes or as a ratio) that travelers must add to their average travel time when planning trips to ensure on-time arrival. The buffer index increases as reliability worsens.
- Planning Time Index - The 95th Percentile Travel Time divided by the free-flow travel time index. The planning time index can also be understood as the ratio of travel time on the worst day (2 days) of the month compared to the time required to make the same trip at free-flow speeds. Consequently, the planning time index represents the total travel time that should be planned when an adequate buffer time is included.
- Travel Time Index - The ratio of the average travel time in the peak period to the travel time at free flow conditions.

Table 2-7 provides a summary of the recommended performance measures for passenger travelers, grouped according to trip purpose for daily travel, and accompanied by a general assessment of the relative importance or severity of a reliability issue.

Table 2-7: Needs and Travel Time Reliability Performance Measures for Passenger Travelers

Broad Classification by Trip Purpose	Detailed Classification by Trip Purpose	Importance / Severity of Reliability	Primary User Information Need	Recommended Reliability Measure
Daily, Constrained Trips	Work	High	Delay during heaviest traffic days.	Planning Time Index
	Pick-up & Drop-off Children	High		
Daily, Unconstrained Trips	Shopping	Low	Additional time necessary to generally arrive on time.	Buffer Index
	Return home	High-Medium		
Occasional, Constrained Trips	Appointments	High	Travel time during peak period versus off-peak period.	Travel Time Index
	Leisure	Medium-Low		
Occasional, Unconstrained Trips	Leisure	Low	Additional time necessary to generally arrive on time.	Buffer Index

Source: Adapted from SHRP 2 Project L11 Technical Memorandum 1, Exhibits 2 and 4

“Daily, constrained” trips are those where the user experiences day-to-day variability in travel time (due to recurrent and incident/non-recurrent congestion) and desires to arrive at the destination at a fixed time (or within a small time window). Reliability can be defined for these trips as the invariability in desired (or required) arrival time at the final destination from day to day. Travelers can incur “schedule delay” costs, which is a penalty for early or late arrival (discussed in detail in NCHRP 431). Typically, late schedule delay costs are higher than early schedule delay costs. Furthermore, due to travel time uncertainty, travelers may not be able to properly plan their daily activities. Results of the focus group interviews showed that unreliability has the most severe impact on “daily, constrained” trips and yields the heaviest potential consequences, e.g., showing up late to work, stress on others relying on the delayed traveler, and potential monetary losses. It is critical that users performing “daily, constrained” trips plan a total travel time (including a buffer) that assumes a general worst-case scenario so that they can schedule their departure to ensure an on-time arrival. The planning time index estimates how bad delay will be during the heaviest traffic days and conveys to the user the total travel time that should be planned when an adequate buffer time is included; therefore, it is recommended as the reliability measure for “daily, constrained” trips. Examples of “daily, constrained” trips include work commutes with fixed arrival times and picking up children from day care, where parents incur monetary fines for late arrival.

“Daily, unconstrained” trips are those where the user experiences the day-to-day variability in travel time, but there is no fixed arrival time requirement against which a measure of “schedule-delay” can be calculated. Reliability can be defined for these trips as the invariability in travel time from day to day. Consequences of unreliability for these types of trips are typically less severe than “daily, constrained” trips and users’ generally only desire to know how much time they should add to their average travel time to arrive generally on time, which is why the buffer index is the recommended performance measure. Examples of "daily, unconstrained" trips include shopping or returning home from work.

"Occasional, constrained" trips are those where the user does not experience the day-to-day variability but does have temporal constraints on arrival time. Reliability for these trips can be defined as the ability to reach the destination on time. Consequences of unreliability in these types of trips are similar in severity to "daily, constrained" trips and may involve significant inconveniences or monetary losses for the user. These trips often occur during or adjacent to peak periods; therefore, it is of interest to the user to know the ratio of the travel time during the peak period to the travel time under primarily free-flow conditions (a travel time the "occasional" user can most easily relate to). The travel time index provides this ratio. Examples of "occasional, constrained" trips include appointments or leisure trips to scheduled events.

"Occasional, unconstrained" trips are those where the user does not experience the day-to-day variability and the user does not have a fixed arrival-time requirement against which a measure of "schedule-delay" can be calculated. Reliability for these trips can be defined in terms of how close the experienced travel time is relative to the expected travel time. The severity of unreliable trips in this category is typically low, due to the flexibility in arrival time. As these trips are occurring during off-peak hours, the basis for trip planning is the average travel time to the destination; the user's primary interest is how much time to add to this average to generally arrive on time. Similar to "daily, unconstrained" trips, the buffer index is the recommended performance measure for this category. Examples include leisure trips that do not involve scheduled events.

Freight Movers

Freight movers represent an important subset of users of the transportation network. Owing to the higher cost of operating and maintaining commercial vehicles, carriers typically put a higher value on travel time and late schedule delays. As Khattak *et al.* 2008 point out, carriers who ship high value and perishable goods are willing to pay to avoid travel time uncertainty and associated costs. Most freight drivers accumulate their own information on travel times and reliability through experience and peer-to-peer information sharing. Others get reliability information through intermediaries or vendors who gather information from primary sources before packaging and marketing it. Also, many carriers use route guidance devices on their shipping vehicles and use information technology to track shipments.

Based on focus group interviews, freight movers generally perceive reliability in terms of their ability to predict trip times. This differs somewhat from passenger travelers, who generally think about reliability in terms of either: 1) deviation in relation to the total trip time, or 2) how often they are able to arrive within a particular time window. For freight movers, if the frequency of trips taking much longer than expected begins to increase, they will see the system as unreliable. This will result in actions such as moving of times and routes to when reliable travel is available (carriers often travel during off-peak times in congested urban areas), widening time windows for delivery, increasing prices for services, and spreading congested delivery routes across vehicles so late deliveries do not compound as severely throughout the day.

The L11 focus group interviews found that most shippers are relatively insensitive to reliability problems and commonly provide carriers with little flexibility. Also, the interviews suggested that travel time reliability is not a key concern of shippers and is not an issue that has made it to their strategic level of operations planning. This result needs further investigation and clearly cannot be generalized. For shippers that carry perishable goods or time-sensitive goods, travel time reliability is expected to be critical. Furthermore, it seems that shippers incur additional costs due to traffic congestion, mainly due to incident congestion.

Freight movers were classified into one of eight groups in order to try and provide a manageable matrix for which to define the effects of reliability and the needs of this particular subset of users. Table 2-8 displays this classification scheme along with the criteria upon which the classifications were made: 1) level of schedule flexibility, 2) level of operational adaptability, and 3) cost of variability. An example of a type of freight moving company is displayed under each category to provide some context as to the type of user represented.

A brief description of the classification criteria is as follows:

- Level of Schedule Flexibility:
    a. Flexible – Carrier can change schedule (departure times) to less congested times or wider time windows with little consequences; and
    b. Inflexible – Carrier meeting another outgoing vehicle, limited timing flexibility, and narrow delivery windows.
- Level of Operational Adaptability:
    c. Complete – carrier can change route, has many deliveries, has large fleet of interchangeable vehicles; and
    d. None – small fleet, many deliveries, few route choices.
- Cost of Variability:
    e. High - carrier experiences significant costs from travel time variability due to high inventory, carries burden of variability; and
    f. Low - cost of variable travel times is small.

Table 2-8: Classification by Characteristics and Needs of Freight Movers

Group Number	Level of Schedule Flexibility	Level of Operational Adaptability	Cost of Variability	Primary User Information Need	Example Company
1	Flexible	Complete	High	1. Travel time variability on preferred routes and alternate routes throughout the day. 2. Estimate of on-time delivery reliability.	Refrigerated carrier. Carrier that operates in a very congested arterial network. For example, grocery store deliveries by large company.
2	Flexible	None	High	1. Travel time variability on preferred routes throughout the day. 2. Estimate of on-time delivery reliability.	Carrier that pays drivers by the hour.
3	Inflexible	Complete	High	1. Travel time variability on preferred routes and alternate routes during specific delivery time windows. 2. Estimate of on-time delivery reliability.	Carrier required meeting tight time windows for delivery, for example delivery companies like Fed-Ex, or residential moving company.
4	Inflexible	None	High	1. Travel time variability on preferred routes during specific delivery time windows. 2. Estimate of on-time delivery reliability.	Carrier that moves air freight, or fresh seafood and must deliver in tight time window.
5	Flexible	Complete	Low	1. Travel time variability on preferred routes and alternate routes throughout the day. 2. Estimate of mean travel time.	Carrier moves bulk natural resources.
6	Flexible	None	Low	1. Travel time variability on preferred routes throughout the day. 2. Estimate of mean travel time.	Carrier has no delivery time windows.
7	Inflexible	Complete	Low	1. Travel time variability on preferred routes and alternate routes during specific delivery time windows. 2. Estimate of mean travel time.	Moving companies.
8	Inflexible	None	Low	1. Travel time variability on preferred routes during specific delivery time windows. 2. Estimate of mean travel time.	Small, temperature-controlled trucking company.

Source: Adapted from SHRP 2 Project L11 Technical Memorandum 1, Exhibit 3

Freight shippers and carriers incorporate expected roadway conditions into their equipment and staffing decisions. Included in those decisions is the importance of on-time delivery reliability, which is extremely important for some freight movers (e.g., as part of a just-in-time manufacturing activity), while it is of only modest importance for others (e.g., delivery of garbage to a landfill). For freight movements for which on-time delivery is extremely important, additional time is often built into the delivery schedule (which requires the carrier to supply additional equipment and resources) to account for unreliable travel conditions. This increases the cost of those deliveries, but at a lower cost to the carrier (and ultimately the shipper) than if the delivery was late. Shipments of goods with lower monetary “late penalties” are scheduled with less “give” in the schedule, allowing carriers to maintain fewer redundant vehicles and drivers. This means that carriers charge lower overall shipping costs for these movements, but they are delivered less reliably. Consequently, trip reliability and its importance are directly (but not necessarily fully) accounted for in the price of the freight transportation service. Because of the competitive environment of the trucking industry, significant changes in roadway reliability are consequently reflected in the price of trucking services. That is, a more reliable roadway network will result in lower costs to the carriers, who will typically pass those savings along to the shippers in order to remain competitive.

The L11 team conducted detailed interviews of different users within the freight industry and found that there are some differences between the needs of freight movers and passenger travelers, and that even within the group defined as freight movers, there are different needs amongst planners and policy makers versus truck drivers and dispatchers. Planners and policy makers are generally more interested in forecasting travel time and reliability for use in long-term route planning and route cost estimation. Truck drivers and dispatchers are typically more interested in real-time data due to their need to adjust routes in-progress to meet schedules and deadlines in the near-term.

Within the category of freight movers, the type of trip also dictates the type of reliability information that is useful for the freight mover. The type of trip falls into two broad categories: Full Truckload trips, where an entire trailer full of merchandise is picked up at one location and delivered to another, and Less than Truckload trips, where trucks make a series of pick-ups and drop-offs along a route. Full Truckload trips will require information between one origin-destination pair, while Less than Truckload (LTL) trips will require more complicated trip-chaining capabilities.

Time-sensitivity is generally more of a focus in the trucking industry due to a number of factors. Therefore, truck drivers and dispatchers will sometimes make use of real-time communication technologies such as direct-connect units that allow dispatchers to instantly communicate to drivers information that may affect travel time and give the driver alternatives for managing his/her route and travel time. Also, trucking companies make use of satellite tracking technology (similar to OnStar-type systems contained in passenger cars) so that dispatchers can receive real-time information on the location of vehicles and data regarding each vehicle’s behavior (i.e., speeds, heading, and braking information).

Needs of Agencies

The current needs of transportation agencies, in relation to travel time reliability and the factors influencing reliability, were identified in L11 for policy makers and highway system managers. The ability of transportation agencies to provide reliable travel on the transportation system is typically limited by one or more of these factors:

- Limitations due to availability of resources and jurisdictional boundaries;
- Ability to predict the occurrence of disruptions, e.g., incidents or adverse weather;
- Adequate access to tools/procedures that remove disruptions quickly and/or supply additional, short-term capacity increases to compensate for capacity lost due to a disruption; and

- Adequate knowledge of which tools work most effectively for given disruptions and the ability to gain feedback on the performance of measures that are applied to improve travel reliability.

Planning and Programming-related benefits associated with an agency having a well-managed reliability-focused performance measurement system include:

- Improvement of information provided to decision-makers in support of strategic planning and programming, facilitating improvements in operations and planning,
- Assistance for agency executives in documenting accomplishments, providing a method for justifying the value of program investments and system improvements,
- Improved understanding of the value of one type of project/system improvement versus another, enabling cost/benefit analysis to be integrated into the agency's budgeting processes.

The key to understanding the needs of agencies in relation to reliability is to recognize that agencies and users (travelers) look at reliability statistics differently. Roadway agencies care about their roads, while customers care about their activities and trips. Although roadway agencies care about the customers' trips, their primary concerns are where, when, how often, and to what extent congestion occurs on their roadways. Each agency has financial obligations to deal with *their* roads, not others' roads; therefore, they must care more significantly about their own roads' performance.

Policy Makers

Policy makers are responsible for decisions typically related to funding for infrastructure capacity expansion, investment in operations management systems, and transportation system monitoring/information dissemination technologies. Broadly speaking, Metropolitan Planning Organizations (MPOs), DOT planning departments, and legislative bodies are all members of this group. Policy makers do not typically have a direct impact on the day-to-day procedures of monitoring travel time reliability; however, the decisions made at the policy level regarding the focus of improvements and spending (e.g., improved safety versus increased efficiency) have a trickle-down effect on an agency's effectiveness when it comes to providing reliable travel times to users. In general, policy makers and planning organizations are concerned with strategic and tactical plans, with a focus on recurrent traffic congestion. Few transportation agencies have adopted policies that mention managing their transportation systems for reliability; however, many agencies' transportation management objectives actually improve travel time reliability while working to improve roadway, capacity, efficiency and safety.

The current practice of "monitoring" travel time reliability amongst governmental and legislative bodies uses reliability performance measures (if the agency is tracking them) to determine how well the agency is performing ("output" measures). This is different than monitoring the effect those actions have on overall changes in travel time or delay experienced ("outcome" measures). Monitoring "output" measures is important for policy makers and allows them to respond to legislative and taxpayer concerns about whether their agency is "doing what they said they would" or what "we told them to do." Policy makers will need a specific set of charts or visual tools to make funding decisions about the transportation system.

Roadway System Managers

Transportation system managers are responsible for real-time and day-to-day operations of road networks, and include persons and entities such as Transportation Management Center (TMC) operators, state DOTs, and traffic information providers. Roadway system managers may make operational decisions and select and implement intelligent transportation systems. System managers are therefore directly responsible for protecting and improving reliability; they are the agency personnel most affected by issues limiting the effectiveness of an agency in providing reliable travel. Transportation Management Centers need and use travel time and reliability information to respond effectively to incidents and other events. Their needs are broadly characterized by surveillance, data processing, event response, and information dissemination to travelers and carriers. The L11 project identified some of the most commonly voiced concerns of roadway system managers:

- Lack of consistent, accurate (traffic/travel and reliability) data.
- Lack of budgetary resources to expand their data collection programs.
- Travel times affected by factors or circumstances out of their control (e.g., adverse weather).
- Modest or unnoticeable improvement in travel time reliability following an action.
- Resistance to the adoption of performance measures because of concerns about adding additional processing and workload for already overloaded employees.
- Lack of a current baseline against which to set goals.

Roadway system managers have three ways by which they can improve the reliability experienced by travelers:

- Improve the routine operation of roadways through infrastructure improvements.
- Reduce the number of disruptions that occur on the system and/or the duration of delay with the disruptions that do occur.
- Quality and timely delivery of information to their customers to put the user in the position of taking action to improve his or her overall travel experience.

For roadway system managers to provide a more reliable travel experience for their customers they need tools and resources that allow them to better manage and improve their existing transportation system. This allows them to maximize the performance of their system while minimizing the frequency and severity of events (factors) that cause disruption. The major factors causing disruption and impacting reliability are identified as incidents, weather, work zones, fluctuations in demand, special events, traffic control devices, and inadequate base capacity (FHWA 2005). Specifically, the factors to consider in evaluating reliability are recurrent and incident (non-recurrent) congestion. Recurrent congestion occurs predictably during peak hours, and at bottleneck locations, e.g., lane drops or weaving sections. Incident congestion is relatively unpredictable occurring during the peak or off-peak hours and at any location along roadway. In both cases, demand exceeds capacity and queues/delays are observed. However, in the case of incidents, the available capacity is further constricted by the occurrence of an event, e.g., crash, vehicle disablement, or debris on road. Typically, traffic incident occurrence is highly correlated with peak hours, complicating traffic operations. Fluctuation in demand created by the need of people to participate in daily activities at certain times and inadequate base capacity are principally responsible for the creation of recurrent congestion. Operational factors that further contribute to recurring congestion include adverse weather (when people drive at reduced speeds on slippery roads), special events, work zones, and traffic control devices. Factors that contribute to incident (non-recurrent) congestion can include adverse weather (e.g., flooding that reduces available capacity), traffic control devices (e.g., improperly timed traffic signals (see also Bo and Hiroaki 2008 for more discussion) and ramp meters), work zones, roadway geometry, and speeds, and driver/vehicle factors (e.g., driver distractions and equipment failure). Overall, there are several factors that contribute to traffic congestion. Therefore, in order to make operational decisions about the transportation system, system managers will need a different set of visual and analytical tools than the policy makers.

Reliability Experts

The L02 study team reached out to three groups as part of the interview effort. The first group (Group A) included individuals who work with travel time reliability monitoring systems for a highway or transit agency. The second group (Group B) included experts in the field of reliability and performance monitoring; primarily members of the L02 Technical Coordinating Committee (TCC). While efforts have yet to be made to specifically conduct interviews with individuals in Group A, several of the people interviewed under Group B could also be classified in Group A. The third group (Group C) included service providers in the area of travel time reliability monitoring. Given the surveys conducted by L11, the L02 project team did not conduct interviews with passenger travelers or freight movers. Rather, the results related to user needs from the L11 focus group discussions were used.

Group A - Individuals Who Work With Monitoring Systems

The L02 team identified 10 to 15 agencies for the Group A interviews based on existing relationships and recommendations from others in the profession. Table 2-9 lists these agencies as well as a general summary of the types of facilities and trips those agencies monitor. The following agencies were interviewed:

- Washington State DOT
- TriMet (Portland, Oregon)
- Virginia DOT
- Ontario Ministry of Transportation
- Port Authority of New York and New Jersey
- Kansas DOT (Kansas City Scout)
- Missouri DOT (Kansas City Scout)
- Jet Express, Inc.

Table 2-9: Extent of Travel Time Monitoring by Agency

Organization	Urban Highway Agency: Commuter Trips	Urban Highway Agency: Truck/ Delivery Trips	Urban Highway Agency: Transit Trips	Rural Highway Agency/ Resort Area: Recreational Trips	Rural Highway Agency/ Resort Area: Truck/ Delivery Trips	Transit Agency: Commuter Trips	Transit Agency: Recreational Trips
Florida DOT	X	X		X	X		
Utah DOT	X	X		X	X		
Washington State DOT	X	X		X	X		
Georgia Regional Transportation Authority	X	X	X				
San Antonio-Bexar City MPO	X	X					
Capital Area Metropolitan Planning Organization (Austin, TX)	X	X					
Puget Sound Regional Council (Seattle, WA)	X	X	X				
Metropolitan Transportation Commission (San Francisco, CA)	X	X	X				
Greater Cleveland Regional Transit Authority						X	X
TriMet (Portland, OR)						X	X
Chicago Transit Authority						X	X
Virginia DOT	X	X		X	X		
Ontario Ministry of Transportation	X	X		X	X		
Port Authority of New York and New Jersey	X	X					
Kansas DOT (Kansas City Scout)	X	X					
Missouri DOT (Kansas City Scout)	X	X		X	X		
Wisconsin DOT	X	X		X	X		
King County Metro						X	X
Mid-America Regional Council- Kansas City	X	X	X				
Southern California Association of Governments	X	X					
Jet Express, Inc.		X			X		
Bonneville County, Idaho Metropolitan Planning Organization	X	X					

The key for this effort was to ensure that agencies of different types, generally including urban highway agencies, rural highway agencies and resort areas, and transit agencies, would all be represented throughout the course of the interview process. In addition to conducting interviews with these three types of agencies, the L02 project team captured the different trip types, including recreational, commuter, truck/delivery, and transit trips.

Group B - Individuals Who Are Leaders in the Field of Reliability

The L02 project team invited all TCC members to participate in small group teleconference discussions. These teleconferences comprised the core of the interview process and allowed the L02 team to participate in more in-depth discussions regarding performance monitoring and travel time reliability.

The L02 project team developed a number of questions to help guide the discussions with Group B members. Each focus group interview was unique, and flexibility was built into the discussions to allow ideas to flow from the participants while still providing guidance to gather information on key aspects of travel time reliability. The primary goal of these questions was to answer the "who, what, where, when, and why" questions of monitoring travel time reliability. These questions were organized into five general categories and provided to the members of each interview prior to their scheduled interview time. For each category, Table 4-4 provides example questions, a summary of the key takeaway points, and the relative amount of information received.

As is shown in Table 2-10, we received medium to high amounts of feedback under four of the five general categories. Education (for both agency staff members and the traveling public) and outreach efforts related to reliability monitoring had limited amounts of feedback, with many individuals mentioning the on-going efforts of TRB, AASHTO, and FHWA, but not much discussion regarding specific educational efforts of staff within different agencies. In addition, there was very little discussion about public educational outreach programs on reliability.

Table 2-10: Interview Categories for Reliability Leaders and Information Received

Category	Key Questions	Key Takeaway Points	Relative Amount of Info. Rec'd
1. Data Collection Practices and Travel Time Measurement Tools	Do you currently use travel time reliability as a performance measure for your system? If so, how do you measure it? Where, when, and for what facilities, areas, corridors, or OD pairs do you measure it? What information is gathered to monitor travel time reliability? How do you obtain travel time information? Are the travel time reliability results archived and/or reported?	Many new and emerging data collection technologies exist, but agencies are still using inductive loops as the most common source for travel time and speed data. Quality control and management of tools and data is very time intensive and takes more resources than most agencies have available. Partnerships with other public and private agencies are vital when it comes to assembling the resources necessary to accurately record and archive data.	Medium
2. Communication to Users	What information is presented to the users of your system and how is it presented? In the future, what reliability information can you envision being delivered to system users and in what forms? Do you provide pre-trip information to users on system conditions? If so, what media is this communicated through?	A few agencies are experimenting with reporting reliability measures. Overall, most users seem more interested in knowing travel times rather than travel time variance. Reliability measures seem to be more useful when communicating pre-trip information. Users are demanding travel time information on alternate routes. Providing travel time or arrival information causes users to perceive the system as reliable. With all of the technology available, agencies need to better understand the most effective and efficient means by which to communicate reliability information.	High
3. Business Processes and Future Monitoring Plans	To what extent do you incorporate information about travel time reliability into day-to-day operations? Do you have quantitative or qualitative goals with respect to reliability? What are the challenges you face with setting reliability goals? Are there gaps in the travel time information you use that need filling? Are there other deficiencies that need improvement?	Using reliability measures is a goal of many agencies, but the way they are used varies. Examples are planning and programming, user cost assessments, performance assessments. Need to develop reliability initiatives at the national level and encourage partnerships at the local level to more easily reach goals established by initiatives. Few agencies monitor reliability on roadways other than freeways.	High
4. Performance Measures	What travel time reliability performance measures or indices do you monitor? Are these measures archived, tracked, or analyzed in any way? Under what system conditions do you monitor travel time reliability (relating to the seven factors influencing reliability)? What spatial and temporal levels of detail do you capture in your existing monitoring system and would you prefer more or less detail?	Understanding the "why" behind the variability is important for agencies to mitigate the problem behind the variability. Agencies would like more guidance on evaluating performance measures on a network level. Need to identify measures most clearly portrayed to the public and that are not facility or mode specific.	Medium
5. Education and Outreach	What resources do you most commonly use to educate your organization on travel time reliability monitoring practices? Does your organization provide public information programs to educate users on how to use travel time reliability monitoring resources? Do users generally feel the system is reliable and, if so, why? If not, what do you think could be implemented to change their perception?	The traveling public is intuitively aware of reliability concepts, but this intuition must be enhanced with educational tools that are marketable and easily accessible to the public. It is important to share information among agencies to advance the research and implementation of reliability programs. Effective outreach strategies must be centered on what users perceive and value and what they will listen to and comprehend in regard to reliability reporting. The guidebook ought to: 1) compile best practices, 2) provide specific examples, 3) provide guidance on reporting reliability for all user types, and 4) address integration with the private sector.	Low

Use Cases

Based on all of the interviews conducted and the prior L11 work, a set of 51 use cases were developed. These were intended to form the functional specifications for the TTRMS. The use cases were designed to fit the template shown in Table 2-11. Each one involves a definition of the type of person asking the question (user), the question being posed, the inputs needed to answer the question, the steps involved in answering the question, and the results expected.

Table 2-11: Use Case Template

User	The type of TTRMS user posing the question
Question	A description of the inquiry and why it would be posed.
Inputs	The data and information needed to answer the question. This description helps users understand the inputs required; and programmers understand the data inputs that must be assembled.
Steps	A list of the actions that have to be performed to answer the question.
Result	The TTRMS output at the completion of the use case.

Three additional notes about the use cases are important. First, even though people think about “on-time” as meaning “not missed”, there is no guarantee about being on time. Here “on-time” means arriving with a certain probability of not being late – or possibly early as is often the case for freight shipments. Second, anywhere the acronym TT-PDF is used (or TT-CDF, the cumulative density function), it refers to the travel time probability density function for individual vehicle travel times unless the text says otherwise. Third, fairly technical information is presented for the results – for example TT-PDFs for the routes that might be selected. This does not mean that such information is the only way to convey the results. Rather, it implies that such information is the basis for the answer; but the communication paradigm might be simpler, as in a single number (e.g., from L14).

The use cases are clustered around types of TTRMS users most likely to make the inquiry. They are also broken down into providers and consumers – i.e., the supply and demand sides of system use. The stakeholders – shown in Table 2-12 – come from four categories:

- *Policy and Planning Support*: Agency administrators and planners that have responsibility for and make capital investment and operational decisions about the system.
- *Overall Highway System*: Operators of the roadway system (supply), including its freeways, arterials, collectors, and local streets and drivers of private autos, trucks, and transit vehicles (demand).
- *Transit Sub-system*: Operators of transit systems that operate on the highway network, primarily buses and light rail (supply) and riders (demand).
- *Freight Sub-system*: Freight service suppliers (supply) and shippers and receivers that make use of those services (demand).

Table 2-12: User Types and their Classification

System User Type	Service Provider (Supply)	Users (Demand)
Policy and Planning Support	Administrators and planners	None (N/A)
Overall Highway System	Highway system operators (public or private)	Privately owned vehicle (POV) drivers, taxi drivers, limousine drivers, etc.
Transit Sub-System	Transit operators, transit vehicle operators	Transit passengers
Freight Sub-System	Carriers, freight movers, truck drivers	Freight customers (including both shippers and receivers)

The use cases are listed in Table 2-13. They are categorized consistent with Table 2-12 into those that pertain to agency administrators and planners, system operators and users, transit passengers, schedulers or operators, freight customers or operators.

Table 2-13: Use Cases for the Travel Time Reliability Monitoring System

Category	Subgroup	Use Cases
System Administrators and Planners	Administrators	AE1: See What Factors Affect Reliability AE2: Assess the Contributions of the Factors AE3: View the Travel Time Reliability for a Subarea AE4: Assist Planning and Programming Decisions AE5: Document Agency Accomplishments AE6: Assess Progress Toward Long-Term Reliability Goals AE7: Assess the Reliability Impact of a Specific Investment
	Planners	AP1: Find the Facilities with Highest Variability AP2: Assess the Reliability Trends over Time for a Route AP3: Assess Changes in the Hours of Unreliability for a Route AP4: Assess the Sources of Unreliability for a Route AP5: Determine When a Route is Unreliable AP6: Assist Rural Freight Operations Decisions
Roadway Network Managers and Users	Managers	MM1: View Historical Reliability Impacts of Adverse Conditions MM2: Be Alerted When the System is Struggling with Reliability MM3: Compare a Recent Adverse Condition with Prior Ones MM4: Gauge the Impacts of New Arterial Management Strategies MM5: Gauge the Impacts of New Freeway Management Strategies MM6: Determine Pricing Levels Using Reliability Data
	Drivers – Constrained Trips	MC1: Understand Departure Times and Routes for a Trip MC2: Determine a Departure Time and Route Just Before a Trip MC3: Understand the Extra Time Needed for a Trip MC4: Decide How to Compensate for an Adverse Condition MC5: Decide En-Route Whether to Change Routes
	Drivers – Unconstrained Trips	MU1: Determine the Best Time of Day to Make Trip MU2: Determine How Much Extra Time is Needed

Category	Subgroup	Use Cases
Transit System	Transit Planners	TP1: Determine Routes with the Least Travel Time Variability TP2: Compare Exclusive Bus Lanes with Mixed Traffic Operations
	Transit Schedulers	TS1: Acquire Reliability Data for Building Schedules TS2: Choose Departure Times to Minimize Arrival Uncertainty
	Transit Operators	TO1: Identify Routes with the Poorest Reliability TO2: Review Reliability for a Route TO3: Examine the Potential Impacts of Bus Priority on a Route TO4: Assess a Mitigating Action for an Adverse Condition
	Transit Passengers	TC1:Determine the On-Time Performance of a Trip TC2: Determine an Arrival Time Just Before a Trip TC3: Determine a Friend's Arrival Time TC4: Understand a Trip with a Transfer
Freight System	Freight Service Providers	FP1: Identify the Most Reliable Delivery Time FP2: Estimate a Delivery Window FP3: Identify how to Maximize the Probability of an On-Time Delivery FP4: Assess the On-Time Probability for a Scheduled Shipment FP5: Assess the Impacts of Adverse Highway Conditions FP6: Determine the Start Time for a Delivery Route FP7: Find the Departure Time and Routing for a Set of Deliveries FP8: Solve the Multiple Vehicle Routing Problem under Uncertainty FP9: Alter Delivery Schedules in Real-Time
	Freight Customers	FC1: Minimize Shipping Costs due to Unreliability FC2: Determine Storage Space for Just-in-Time Deliveries FC3: Find the Lowest Cost Reliable Origin FC4: Find the Warehouse Site with the Best Distribution Reliability

Only one of the use case analysis procedures is described here in detail. Table 2-14 shows AE1 that focuses on the contributions to the reliability of a segment or route.

Table 2-14: Assess the Contributions of Unreliability Sources (AE1)

User	Agency Administrator
Question	What Factors Affect Reliability?
Steps	1. Select the system of interest (e.g., a region or set of facilities). 2. Select the timeframe for the analysis: the date range as well as the days of the week and times of day. 3. Assemble travel time (travel rate) observations for the system for the timeframe of interest. 4. Label each observation in terms of the regime that was operative at the time the observation was made, that is each combination of nominal congestion and non-recurring event (including none). 5. Prepare TR-PDFs for each regime identified. 6. Analyze the contributions of the various factors so that the differences in impacts can be assessed.
Inputs	Travel times and rates for the system and date range of interest plus information about the nominal system loading that would have been expected and any non-recurring events.
Result	A set of TR-PDFs that portray the impacts of various factors on travel time reliability.

Summary

The traffic content business is a complex, growing field. The range of data sources available is growing constantly. Public agencies and private firms are using a wide array of technologies to assemble the data upon which their travel time assessments are based. Overall, however, with a few exceptions, travel time reliability information is seldom made available to potential users in a format that can help them make informed travel decisions. There is substantial variation in the format and sources by which reliability information that is currently disseminated by agencies.

The array of individuals and firms that want to make use of travel time reliability information is rich and expansive. In general, agency administrators and planners typically want summary information about system performance. They want to know how various factors affect reliability, like growing demand, or inclement weather, so they can make investment decisions or formulate policies that help to ensure system reliability will be acceptable. System operators, transit operators, and freight service providers think in terms of service provided: whether trips take longer or shorter than they ought to or they promised they would. These inquirers want technical, quantitative information, both (near) real-time data for operations management and archived historical trend data for strategic and investment planning. Drivers, transit riders, and shippers want qualitative, anecdotal and objective, quantitative information about reliability. They think in terms of: 1) deviations in trip time relative to the total trip time, or 2) how often they are able to arrive within a particular time window (or their shipments). What they experience affects departure times, mode choice, route choice, and even destination and location choices. Moreover, they make location decisions based on expected network reliability. Factors that affect reliability are clearly of interest to all system users. Some factors are internal to the system such as its operational control (e.g., signal timing), base capacity, and maintenance (e.g., work zones); others relate to the users, like incidents, unusually high demand, and special events; and still others are related to exogenous factors like weather and the performance of complementary and competing modes.

CHAPTER 3: FUNCTIONAL SPECIFICATIONS

The Travel Time Reliability Monitoring System (TTRMS) is intended to be an add-on to existing traffic management systems. Its structure is shown in Figure 3-1. Inside the main box are the three major modules: the data manager, the computational engine, and the report generator. The data manager assembles incoming data from traffic sensors and other systems, such as weather data feeds and incident reporting systems, and places them in a database that is ready for analysis as “cleaned data”. The computational engine works off the cleaned data to prepare “pictures” of the system’s reliability: when it is reliable, when it is not, to what extent, under what conditions, etc. In the exhibit this is illustrated by “regime TT-PDFs”, probability density functions showing the distribution of travel times under various conditions (regimes). The report generator responds to inquiries from users—system managers or travelers—and uses the computation engine to analyze the data and provide information that can then be presented back to the inquirer or decision maker.

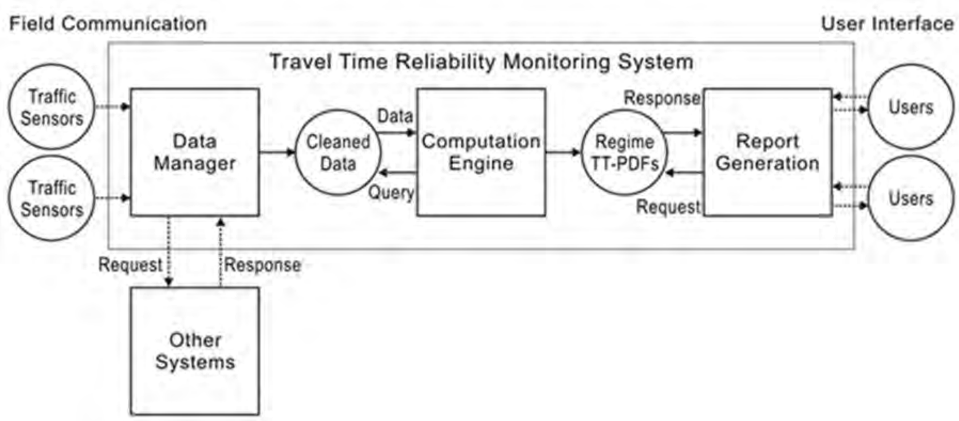


Figure 3-1: The Travel Time Reliability Monitoring System

Analytical Process

The TTRMS uses four key steps as illustrated in the conceptual diagram of information flow shown in Figure 3-2.

First, the TTRMS **measures** travel times. This is a complex technical topic due to the variability of traveler behavior and the plethora of different measurement sensors. Correctly *measuring* travel times along a given route requires a great deal of systems development effort and statistical knowledge. This report serves as a primer on how to measure travel times, effectively, using available technologies and statistical techniques. Measuring an individual travel time on a segment or route is the foundational unit of analysis for reliability monitoring.

Second, the TTRMS **characterizes** the reliability of a given system. This is the process of taking a set of measured travel times and assembling them into a statistical model of the behavior of a given segment or route. The statistical paradigm outlined in this report is that of using probability density functions to characterize the performance of a given segment or route, usually specific to a particular operating regime (a combination of congestion level and non-recurring events). This report gives specific advice on the statistical decisions required to effectively characterize the travel times. Characterizing the reliability of a segment or route is fundamental to making good decisions about what to do to improve the performance of that segment or route.

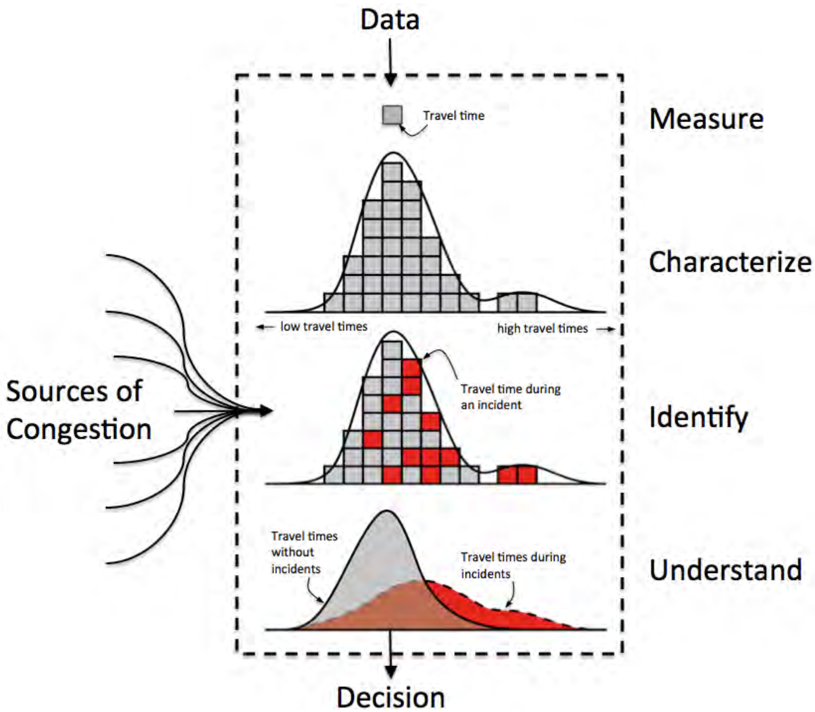


Figure 3-2: Information Flow in the TTRMS

Third, the TTRMS **identifies** the sources of unreliability. Once the reliability of a segment or route has been characterized, transportation managers need to understand the correlates of unreliability (and how to “fix” it). The report follows the list of factors that FHWA uses to describe why congestion arises, breaking these sources into the seven major influencing factors described previously—two internal and five external. It discusses how to organize data into time intervals when these influencing factors were at work and produce descriptions of travel time reliability (TT-PDFs) as associated with these various factors. Identifying the travel times impacted by these sources of congestion is required preparation for understanding system reliability.

Finally, the TTRMS helps operators **understand** the impact of these sources of unreliability on the system. For example, to mitigate the impact of incidents, service patrols, and changeable message signs that can reroute traffic may be considered. However, to mitigate work zone congestion, construction traffic mitigation and smart work zones may be considered. This final step in turning raw data into options and actionable decisions requires both quantitative and qualitative methodologies: operators need clear visualizations of data, as well as quantifications. This dual approach supports both data discovery and final decision-making about a given segment or route. Understanding reliability is the key to good decision-making about improving system reliability.

The TTRMS enables decision makers in a region to understand how much of the delay is due to unreliability, and prompts ideas about how to mitigate that delay. For example, it helps a freeway operator understand whether to deploy more service patrol vehicles (to clear incidents more quickly) or focus efforts on coordinating special event traffic (to reduce delay from stadium access)? A reliability monitoring system, as outlined in this report, can help an operator understand which of these activities is worth the investment, and what the payoff might be. Such systems add a new, powerful, practical traffic management tool to the arsenal of system operators. While, knowledge about the effectiveness of various mitigation actions can be scarce, service patrols in urban areas are known to be effective in ameliorating incident effects and reducing their durations; or changeable message signs can effectively divert travelers to alternate routes when displaying the right content and placed appropriately at decision points.

Key Features

This section describes the key features that the L02 study team believes need to be part of any TTRMS.

Monuments

The travel times should be based on travel times to and from *monuments*. A monument is a measurement point to and from which travel times are measured. As illustrated in Figure 3-3, the monuments should be at the midpoints of the physical links. This removes the travel time ambiguity that arises if intersections and interchanges are used.

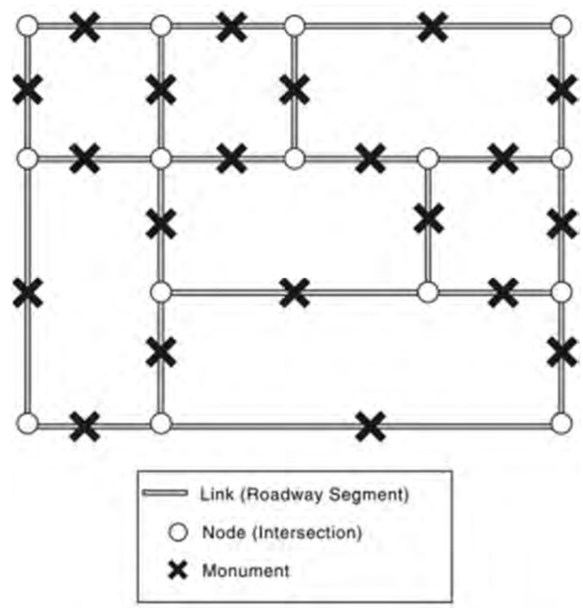


Figure 3-3: Travel Times based on Monuments

Vehicle trajectories between the monuments are all the same. They include the same delays associated with the turning movements. The correct turning movement delay is included in each monument-to-monument travel time. This is clearly important for arterials but it is also important for freeways. Ramp movements can have different travel times (e.g., direct ramp or loop ramp, as well as and any traffic control on the ramp—such as a signal—as is sometimes the case in Los Angeles).

The monuments also need to be locations that the traffic management center uses to monitor the system, as in the location of system detectors on both the freeway and arterial networks. This minimizes the database management tasks involved in keeping track of where the monuments are located. The can also be the location of toll tag readers and AVI sensors. They should not be placed at locations where standing queues occur.

Fundamental Units of Data

Every TTRMS will be based on some set of fundamental units of data. The L02 study team worked most often with 5-minute average speeds from system (loop) sensors and individual vehicle travel times (from AVI- or AVL-equipped vehicles). In the case studies, finer-grained system sensor data was often available (down to 30-second intervals), but it was not used. Aggregated values based on the individual vehicle travel times (e.g., averages) also could have been developed, but they were not. Hence, this final report and the guidebook documents most often refer to 5-minute system detector data and individual vehicle travel times.

An advantage to the system (loop) detectors is that they provide information that is based on all the vehicles in the traffic stream (Enam and Al-Deek 2006). The disadvantage is that no individual vehicle data are provided. The individual vehicle data (e.g., speed) are observed but not reported out by the monitoring station.

An advantage to the AVI- and AVL-data is that data for individual vehicles are reported (List *et al.* 2005a, List *et al.* 2005b, List *et al.* 2006, Demers *et al*. 2006a, Demers *et al*. 2006b, Byon *et al*. 2006, Dion and Rakha 2006, Feng *et al.* 2011, Fontaine and Smith 2005, Li *et al.* 2006,Hoeitner *et al*. 2012, Vanjakshi *et al.* 2009, Liu *et al.* 2010, Xialiang and Koustsopoulos 2008, Lin *et al.* 2003, Ma and Koustsopoulos 2010, Pan *et al.* 2007, Soriguera and Thorson 2007, Quiroga and Bullock 1998, Kaparias *et al.* 2008, Wasson *et al.* 2008, De Fabritiis *et al.* 2008, Ma *et al.* 2009, Liu *et al.* 2007, Wojtowicz *et al.* 2008Yamamoto *et al.* 2006, Yamazaki and Kurauchi 2012). This includes speeds, travel times and, in the case of AVL data, complete trajectories (Cetin *et al.* 2005, Yang *et al.* 2011, Ernst *et al.* 2012, Haghani *et al.* 2010). The disadvantage is that only some vehicles are observed, whether it is only the vehicles equipped with discoverable Bluetooth device or those equipped with tags. (See Kwon *et al.* 2007, Martchouk *et al.* 2011 for an interesting discussion on this topic.) Hence, there can be a bias in the observations vis-à-vis the overall traffic stream.

Investigators have also used buses, trucks and other vehicles as probes for collecting travel time data, but these information sources are not reviewed in detail here. Studies that have examined these sources include Hall and Nilesh 2000, Berkow *et al*. 2008, Bertini *et al*. 2005, and Chakroborty and Kikuchi 2004, Uno *et al.* 2009, Zhu *et al.* 2011.

Imputation to fill Data Voids

Described more fully in the next section, it is important to use imputation to fill voids caused by missing data (see also Wang *et al* 2008). To monitor travel time reliability, high-quality, real-time data must be available. Missing data interferes with this objective. Hence, within obvious limits, estimating values for voids is important. This pertains to data like spot speeds (spot rates) from system detectors as well as segment and route travel time data obtained from AVI- and AVL-based systems.

Real-Time Data for Non-Recurring Events

Information about non-recurring events needs to be collected in real-time from sources that provide such information. Some monitoring systems already collect incident data and make it available for current and future analysis. But weather data are often not collected; and the same is true for special events. The “problem” is that much of this data is perishable; and if it is not collected as events unfold, it can be lost. If that happens, then it becomes either very labor intensive or impossible to determine why specific travel times arose. For special situations or special analyses it may be possible to assemble this information ex-post-facto – the L02 study team did this a number of times during the case studies and use case analyses – but for operating agencies this is not a reasonable option.

This design feature has several implications. One is that the sources for this information have to be identified and real-time data feeds have to be established. Another is that data structures need to be created to store the data. A third is that fields have to be added to the travel time monitoring records so that linkages are created between the travel times and the non-recurring events. Finally, tools and techniques have to be developed that allow the monitoring system to “automatically” link the non-recurring events to the travel time observations. This is not trivial because the non-recurring events may be on adjacent facilities - upstream downstream, or even in the opposite direction - of the segment where the unusual travel times arose.

Regimes for Data Classification

The TTRMS needs to classify travel time observations on the basis of the regime (operating condition) that was operative at the time when the travel times were obtained. This avoids misinterpreting and misunderstanding the impacts of congestion and non-recurring events.

The recommendation of the study team is that these regimes be based on a combination of a nominal congestion condition (e.g., uncongested, low, moderate, and high congestion) and a non-recurring event condition (e.g., none, weather, incident, special event) as shown in Figure 3-4. Ultimately, it is for these conditions that the PDFs are developed and by which, through the PDFs, the reliability performance of the segment, route, or other facility is understood and analyzed; and for which actions are taken to improve reliability.

Reliability Regimes						
	Non-Recurring Condition					
		One of Several				
Congestion Condition	None	Weather	Incident	High Demand	Special Event	Work Zone
Uncongested						
Low						
Moderate						
High						

Figure 3-4: Classifying the Travel Time Observations by Operating Regime

For most practical applications, it appears sufficient to assess the congestion condition at a 5-minute granularity. One minute seems too short; and 15 minutes is too long. In 15 minutes, the operating conditions can change dramatically, especially during heavy congestion.

The non-recurring event categories should be consistent with the FHWA “sources of congestion.” The “insufficient base capacity” condition is captured by the congestion condition categories (i.e., situations where the D/C ratio is high enough that sustained queuing occurs). The high demand category is equivalent to “fluctuations in demand”.

Travel Rates in Addition to Travel Times

The TTRMS should focus on analyzing travel rates as well as travel times. The travel rate is obtained by dividing the travel time by the distance traveled. Travel rates make it possible to compare the performance of one segment with another; and one route with another (in terms of the distribution of the travel rates involved). Spot rates are also important to study. They are the inverses of the spot speeds. They are measured at a specific location by observing the speed and computing the inverse.

Probability Density Functions and Cumulative Density Functions

The TTRMS should focus on creating and analyzing travel time (and travel rate) PDFs. Through the case studies and use cases, it was found that the PDFs (and CDFs) were both necessary and sufficient to address the reliability issues involved or the questions posed. A corollary is that the TTRMS can certainly produce other metrics derived from the PDF, e.g., the travel time index or the buffer time index, and it does so by analyzing the PDF.

Figure 3-5 shows the kinds of CDFs that the TTRMS should produce. It plots the distribution of 5-minute average travel rates on Interstate-8 westbound in San Diego across a three-month period under various regimes.

Since the plots are CDFs, each point on each line shows how many 5-minute average travel times for that regime were equal to or less than the value shown on the *x*-axis. For example, when inclement weather occurs during heavy (recurrent) congestion, one half (50%) of the travel rates (seconds per mile) are up to 70 sec/mile. That is, 50% of the travel rates are this long or *shorter/smaller*. The 90th percentile travel rate is 110 seconds per mile. Or put another way, 9 out of every 10 vehicles is traveling at that rate or faster.

The value comes from comparing one CDF with another. For example, analysts can compare the distribution for high recurrent congestion *and* inclement weather with high recurrent congestion without inclement weather. Without inclement weather, 50% of the vehicles are traveling at 52 sec/mi instead of 70 sec/mi—considerably faster. And at the 90th percentile, the difference is even more dramatic: 58 sec/mi versus 110 sec/mi.

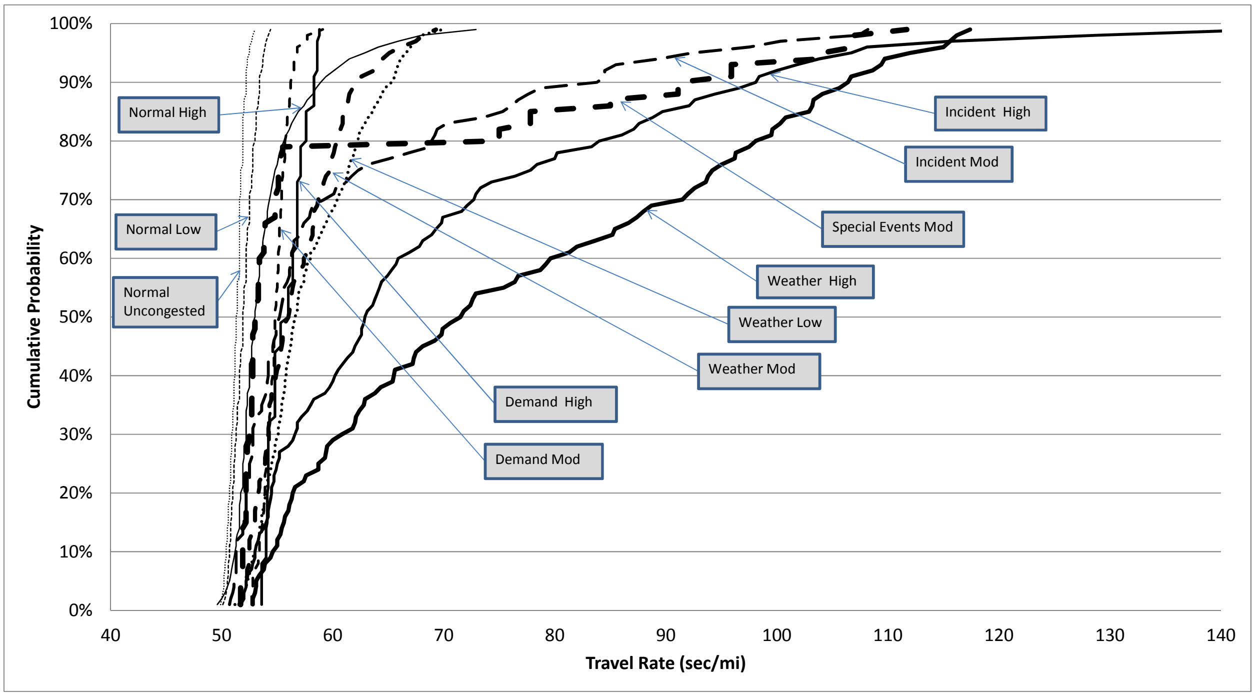


Figure 3-5: Information Revealed by the CDFs

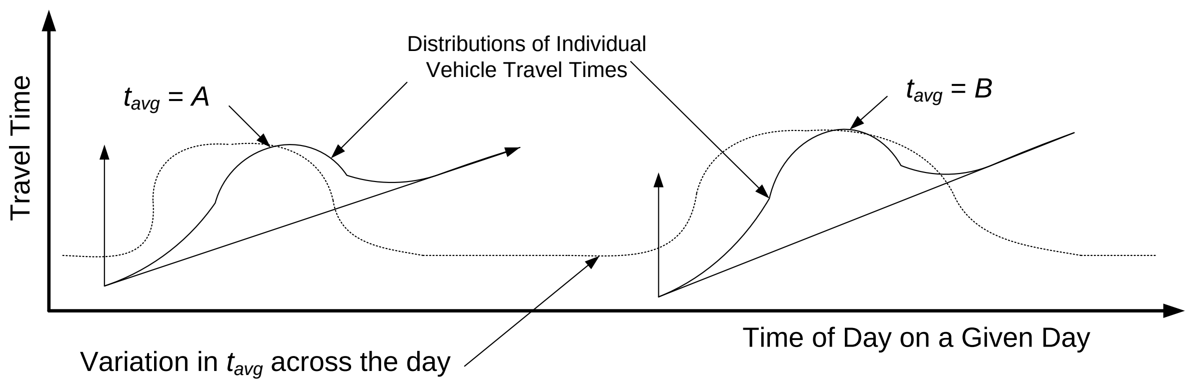
Not only does the exhibit indicate that the difference between the two conditions is large, but it also suggests that taking appropriate actions to mitigate these impacts would produce significant benefits in terms of improving reliability. The mitigating actions would be intended to cause the travel times (or travel rates) during incidents to get much closer to those when there are no incidents. Moreover, after the mitigating actions were taken, the TTRMS would be able to show how reliability improved.

Times for Individual Vehicles as Well as System Averages

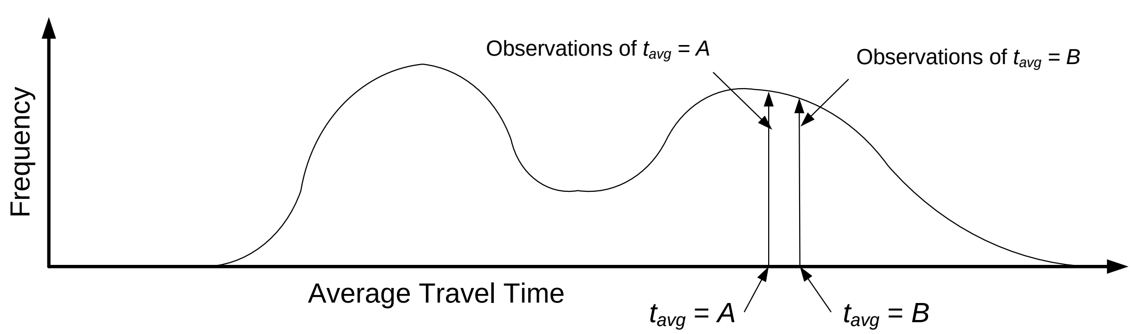
The TTRMS should be designed to collect and analyze individual vehicle travel times as well as averages from system detectors. While aggregated system average data are far more common today, the individual vehicle travel times address issues of system performance from the users' perspective.

Figure 3-6 illustrates how these measures are related. At any given point in time (e.g., during a given five-minute time period on a given day) vehicles traverse a given segment or route. They produce travel times that can be summarized by a distribution. Two examples are shown in part a) of the figure, one toward the beginning of the day; and another toward the end. System detectors (e.g., loops and cameras) observe spot speeds (spot rates)for all of the vehicles but only at specific locations. Bluetooth sensors, toll tag readers, and similar devices, observe travel times for some of the vehicles.

Across an extended timeframe, say a year, a distribution of the average travel times can be created as shown in part b) of the figure. This distribution can be based on the same 5-minute time period each day – which analysts often do – or some collection of five-minute time periods (such as the morning peak) that represents a given operating condition. It is these distributions of average travel times that system operators use today to monitor the performance of their networks and make assessments of where and when corrective actions should be taken to reduce the variability in travel times (i.e., improve reliability).



a) Genesis of the travel time data



b) Resulting distribution of t_{avg}

Figure 3-6: System Average and Individual Vehicle Travel Times

The distributions of individual vehicle travel times can also be developed and studied if the data are available so that the system performance received by (given to) the individual users can also be assessed. It is uncommon for system managers to examine these distributions today, but as vehicle monitoring technologies become more prevalent, it is likely that such information will be used for decision-making purposes.

Segment-Level Travel Times

Segment-level travel times are the fundamental building blocks in terms of measure of time for a highway network. A segment is a path between two monuments. In the case of system level detectors, segments are often defined as being sections of freeways (or arterials) immediately upstream and downstream of a system detector as illustrated in Figure 6-2 (See Kwon *et al*. 2000). For AVI-based systems, segments are often links (one-way arcs) between AVI monitoring stations. For AVL-based systems, segments can be defined to and from whatever locations seem most useful or appropriate still in keeping with the notion of where to locate monuments.

Non-Parametric Analysis Techniques

Another key feature is that the TTRMS analyzes the PDFs using non-parametric techniques – or more simply by just focusing on the entire density function itself (Rosenblatt 1956 and Silverman 1986). As described in Section 2, the density functions are frequently multi-modal and the details of each mode are critical in understanding what is or has happened from a reliability perspective. It seems that no parametrically based distribution – or even multi-modal parametrically based distribution – can serve adequately as a building block upon which the TTRMS can be based. Figure 3-7 illustrates this point in the context of travel times between South Lake Tahoe and Placerville, CA along US-50. Notice the extraordinarily rich diversity in the shapes of the CDFs.

Route PDFs from Segment PDFs Using Correlation

Since the data for specific routes is likely to be too thin to estimate route-level PDFs and CDFs directly, such information has to be synthesized by combining segment-level data. The TTRMS has to be able to do this. Chapter 6 describes ways to do this, but the main stipulation is that the correlation in travel times (travel rates) from one segment to the next has to be taken into account. The travel time observations are inherently correlated because the driver populations overlap between adjacent segments, and drivers are at least somewhat consistent in their speed management.

Several methods for combining segment PDFs have been developed. They are described comprehensively in the Guidebook and its supplements; and portrayed briefly in Chapter 6. Other types of modeling efforts include Dong and Mahmassani 2011, Sun and Gao 2012, Ishak *et al.* 2007, Feng *et al.* 2012, Van Hinbergen and Van Lint 2008, Ramezani and Geroliminis 2012, Rice and VanZwet 2004, Susiwati *et al.* 2011, van Lint and van Zuylen 2005, van Lint *et al.* 2008, and Jintanakul *et al*. 2009.

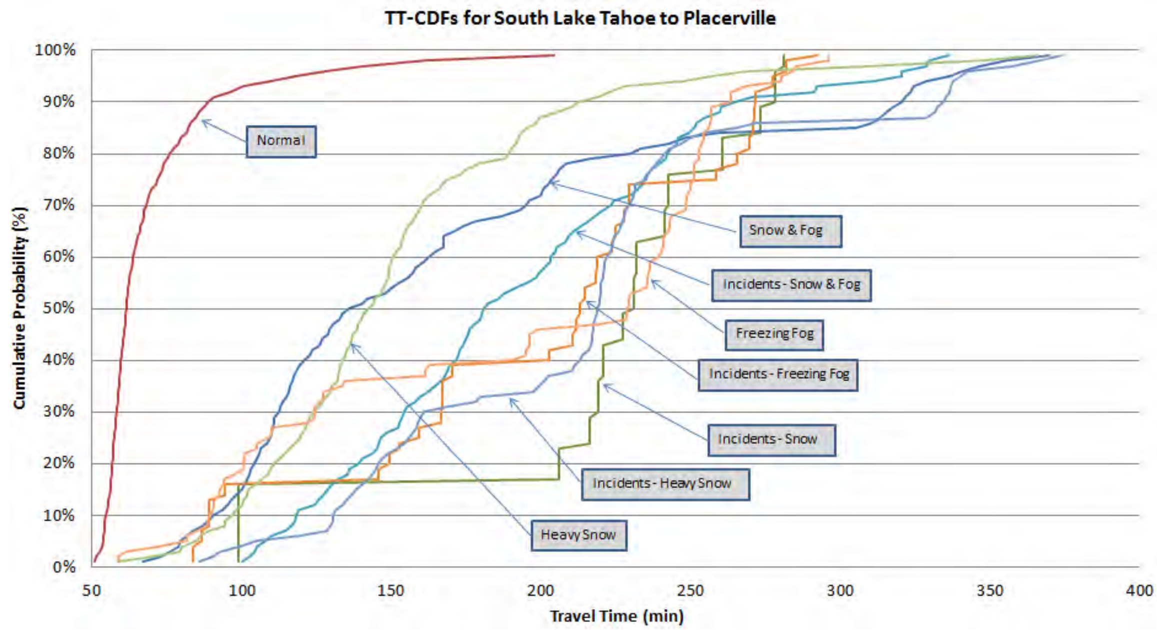


Figure 3-7: Variations in the TT-CDFs for trips from South Lake Tahoe to Placerville

PDFs as the Basis for Archiving

Many options exist about what data to archive for use in reliability analyses. Some experts suggest that "everything" should be kept. Because these people tend to be thinking about keeping the observations of average speeds for the system detectors (loop detectors) at an interval of every 30 seconds (or every minute) or so, this is not unreasonable. Data storage is becoming cheap; and by keeping everything, the "raw" data are then available for future analysis. Of course, they are not keeping the actual observations of individual vehicle detection events, or speeds. They are keeping summaries (averages) based on those data.

Whether it is wise to keep everything in the context of AVI- or AVL-related data is not so clear. For AVI systems this would mean keeping every timestamp for every vehicle observed at every AVI- location. For AVL-based systems, this would mean keeping every GPS ping. Most likely, these options are not reasonable. Moreover, liability issues associated are associated with such information.

For system detector data, like loops, it does seem logical to keep "everything". This means: keep the average speeds, volume counts, occupancies, etc. that are collected every 30-seconds or every minute from every detector in the system. A five minute level of granularity is probably the upper bound on the interval between archived observations that is still useful for reliability analyses. Fifteen minutes is too coarse. In 15 minutes, a lot can happen during the peak hours. It probably also makes sense to add fields that indicate the regime was extant when the data were collected: either the region identifier itself or a combination of two fields: one that indicates the nominal congestion level that would have been present under normal conditions and a second that indicates the non-recurring event (including none) that was occurring (including "none") during the 5-minute time period.

For AVI- or AVL-based data, it seems valuable to record segment-level CDFs on a periodic basis. Even though some researchers are experimenting with parametrically-based procedures (Guo *et al*. 2012, Hesham *et al.* 2006), this is more useful than storing the parameters for a pre-selected density function. The study team could not identify a parametrically-based density function that worked well.

The study team used two mechanisms to create these CDFs. In the first, the 51 most recent AVI- or AVL-based travel time observations were recorded on a periodic basis. The number of observations was chosen so that a data point would be recorded for every 2nd percentile up to and including the 100th. Every 5 minutes was the most common frequency with which this was done although every 15 minutes seems like a plausible answer for archiving purposes as well. The vehicle IDs were not kept – and for liability reasons they probably should not be – although keeping them makes it possible to track individual vehicles across successive segments. In the second instance, the 51 AVI- or AVL-based observations were recorded every time 25 new observations were obtained; which means half of the samples overlap from one set of stored values to the next. Of course, other variations are possible, like having only 10 of the values overlap, or none.

The other piece of information that seems logical to include along with the 51 observations is the timespan covered by those observations – the difference between the time of the newest and the oldest observation. The timespan gives an indication of how closely the 51 observations correspond to the time period to which they were assigned (e.g., the 5 minute time period in the case of the first mechanism; and the timestamp of the last observation in the second.) Given the penetration rates that exist today and the locations where the Bluetooth data were recorded – this timespan tended to be about an hour at night and only 10-15 minutes during the peak hours. (It is helpful that there is more traffic during the peak hours – when these CDFs are most important and change most significantly.

Of course, for special studies or situations where detailed analysis is desired, keeping "everything" still makes sense.

Summary

The TTRMS is intended to be an add-on to an existing traffic management system. It is broken down into three major modules: a data manager, a computational engine, and a report generator. The data manager assembles incoming information from traffic sensors and other systems, such as weather data feeds and incident reporting systems, and places them in a database that is ready for analysis as "cleaned data". The computational engine works off the cleaned data to prepare "pictures" of the system's reliability: when it is reliable, when it is not, to what extent, under what conditions, etc. In the exhibit this is illustrated by "regime TT-PDFs". The report generator responds to inquiries from users—system managers or travelers—and uses the computation engine to analyze the data and provide information that can then be presented back to the inquirer or decision maker.

The value of the TTRMS comes from helping agencies understand the reliability performance of their systems and monitor how reliability improves over time. It equips them to answer questions like:

- What is the distribution of travel times in the system?
- How is the distribution of travel times (or rates) affected by recurrent congestion and non-recurring events?
- How are freeways and arterials performing relative to reliability performance targets set by the agency?
- Are capacity investments and other operational actions helping improve the reliability of the travel times?
- Are operational improvement actions and capacity investments helping to improve the travel times and their reliability?

CHAPTER 4: DATA COLLECTION, ASSEMBLY, AND CLEANING

As the project unfolded, it became increasingly apparent that clean and complete data was critically important if meaningful travel time reliability information was to be obtained. (Karr *et al*. 2006 provide a valuable examination of data quality issues.) This proved to be one of the main insights derived from a team decision to focus on using field data rather than simulation to develop and test the TTRMS.

Data Quality

Two main data quality issues emerged during the project. The first related to AVI sensor data. The second pertained to AVL-based timestamp and location observations.

Passage Times for AVI Sensors

For AVI-based sensors there is an issue of attributing passage times – deciding when a given vehicle passes by the sensor. For toll tag readers (which are also AVI sensors) this is not a significant problem: the timestamp corresponds to when communication with the tag takes place. But for other AVI sensors where no specific transaction occurs, the in-vehicle device is likely to be within range of the sensor for an extended period of time, and sometime within that window is the best choice for the time stamp.

The reason this is important is measurement error. It is important to avoid creating noise in the travel time values by being imprecise about when a specific vehicle passes by a specific location. If the travel times between sensors are about 60 to 120 seconds, and the timestamps have a variation of ±10 seconds on when the sensor was actually passed, then the travel times can be as much as 20 seconds shorter than up to 20 seconds longer than the actual travel time that occurred. This is an error of ±33% if the travel time is actually 60 seconds!

This “problem” surfaced for the study team when the Bluetooth data along US 50 between Sacramento and South Lake Tahoe was being studied. Figure 4-1 shows the MacID responses from a Bluetooth device that was detected by one of the Bluetooth readers that was along US 50.

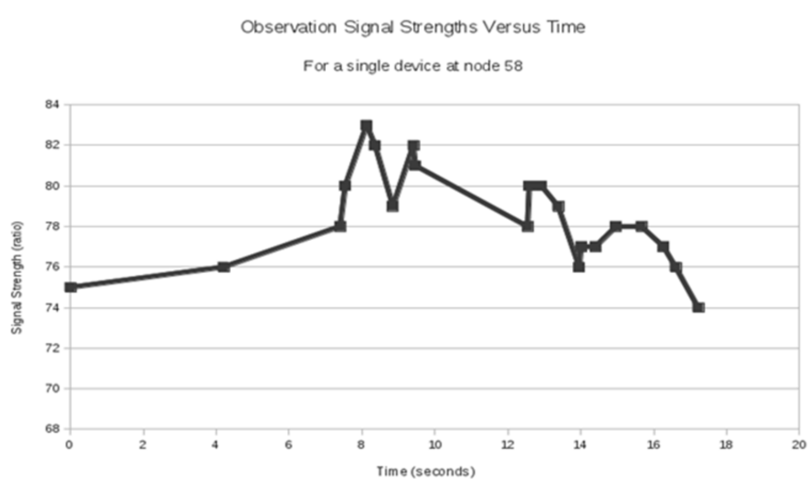


Figure 4-1: MacID responses for a Vehicle

In this instance, the device is observable only for 20 seconds and the signal strength peaks at between 7 and 10 seconds. Hence, the assignment of a passage time in this instance is clear. It should be at about 9 seconds.

However, the team's understanding of most Bluetooth readers is that they do not monitor signal strength to determine a passage time. Rather they use the average of the first and last time the device was observed. In the case of the vehicle whose detection is shown in Figure 4-1, this is not likely to be a problem. It was first observed at "0" seconds and last observed at "17", so the average would be 8.5, which is also when the strongest signal strength was observed.

But the use of this average time can be problematic. Figure 4-2 shows another vehicle that was within range of the sensor for about 700 seconds (almost 12 minutes). Plotted again in this instance is the signal strength of the device's response versus time. It seems likely that the device was closest to the sensor about 15-20 seconds after coming into range. It could be that 15 seconds is the "best" passage time to use.

But maybe two values are better than one. If two values were used, the first would then be used to compute the travel time "to" this sensor; and the second, to compute the travel time "from" this sensor to the next. Measurement error would be minimized. On the other hand, if the 15 second value was used, this would add about 10 minutes for the travel time "from" this sensor to the next one visited – time that was actually spent near the sensor – not traveling to the next one. Unless the distance to the next sensor was more than 100 minutes away (almost 2 hours), use of the 15 second value would introduce a measurement error of more than 10 percent.

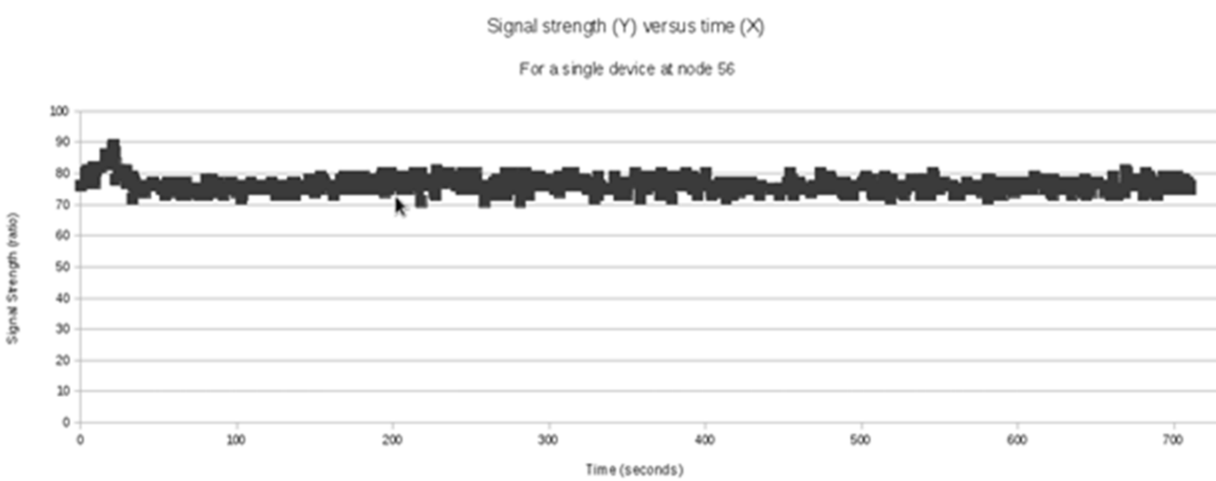


Figure 4-2: MacID responses for a Second Vehicle

It might be best to use a data processing rule that says: if the difference between the first and last timestamp is short (say less than 20 seconds) then use the timestamp from the strongest signal response. Otherwise, use two values, one of which corresponds to the earliest observed time and the other, the last.

Times and Locations for AVL-Equipped Vehicles

Automated Vehicle Location (AVL) technologies track vehicles as they travel. Hence, entire trips can be observed, including the path employed. Moreover, actual travel (and not trip times) can be computed for segments and routes by differencing the timestamps for when the vehicles pass specific locations in the network. Trips that involve stops can be removed so that their trip times do not bias the travel times *or* the times associated with the stops and other side-trips can be removed so that actual travel times are obtained (see Hellinga and Fu [2002] for an example of how to remove biases).

An important detail is that the AVL data are not intrinsically tied to the underlying highway network. As illustrated by Figure 4-3, the latitudes and longitudes reported are based on the information at the disposal of the GPS device, not the physical location of the highway segment being traversed.

Hence, the AVL data need to be matched to specific segments for their data to be used in estimating travel times. One way to do this is through map matching algorithms. The data received from the vehicle-based sensors (longitude, latitude, point speed, bearing, and timestamps) are snapped to segments in the study network. Map matching is one of the core data processing algorithms for associating AVL-based travel time measurements with a route. A typical GPS map-matching algorithm uses latitude, longitude, and bearing of a probe vehicle to search nearby roads. It then determines which route the vehicle is traveling on and the resulting segment and route travel times.

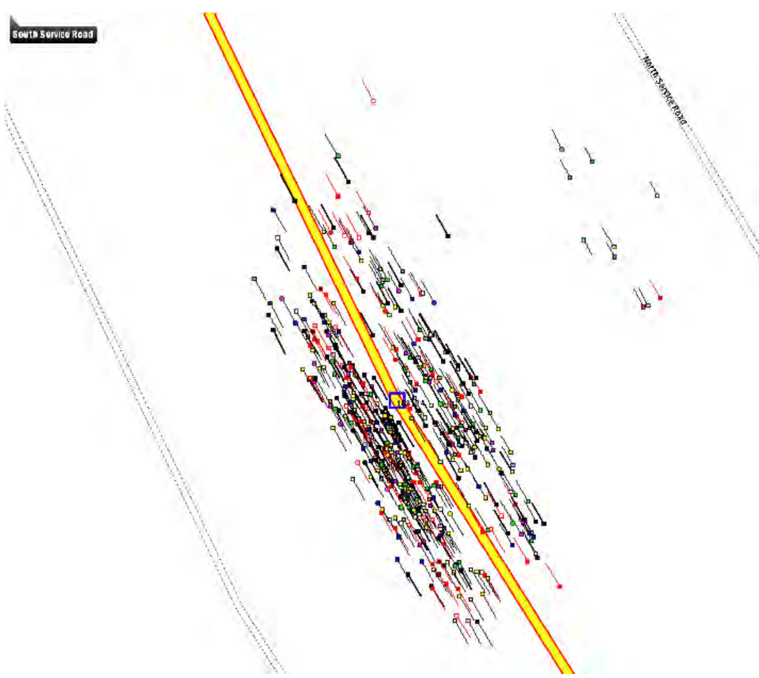


Figure 4-3: Locations and Headings Reported by AVL-Equipped Vehicles Trips
Source: ALK Technologies

In many cases, as shown in Figure 4-4, there can be multiple answers to the map-matching problem.

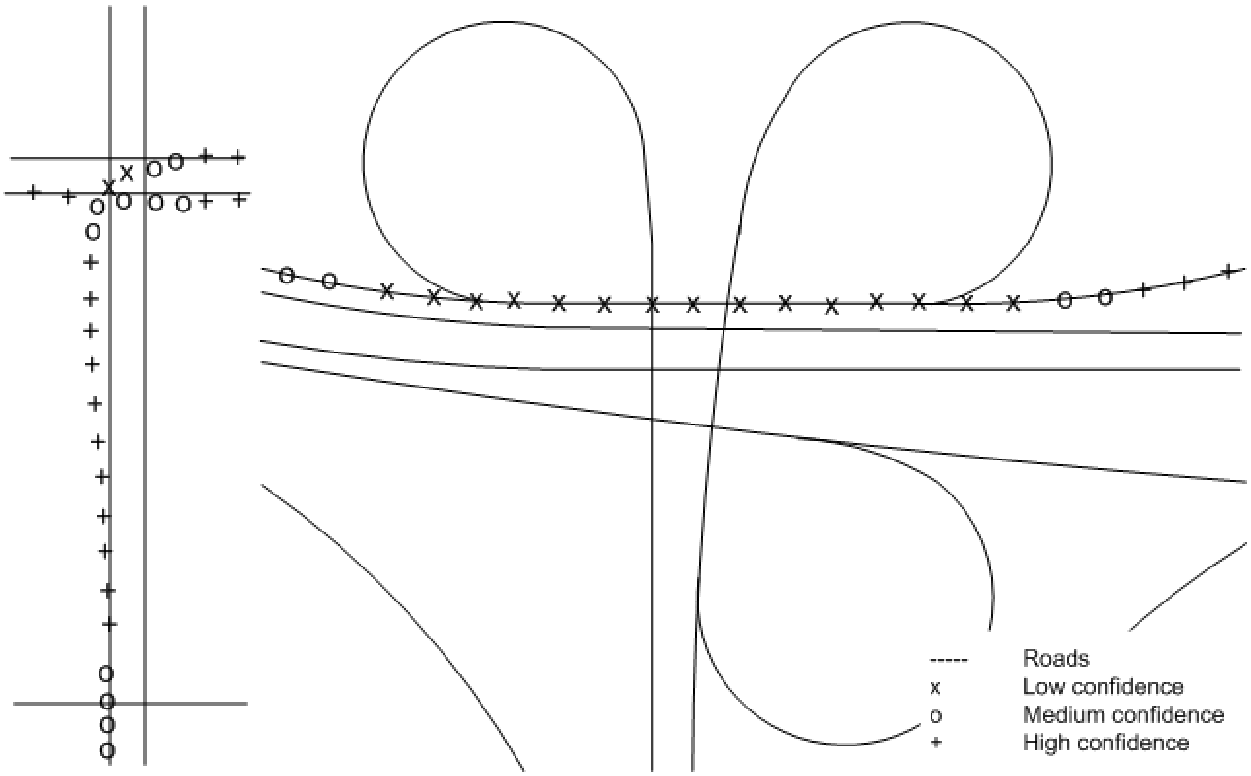


Figure 4-4: Example of Map Matching Challenges for AVL Data

Thus, a number of GPS data mining methods have been developed to find the closest or most probable match. Map-matching algorithms for transportation applications can be divided into four categories: (1) geometric; (2) topological; (3) probabilistic; and (4) advanced. Geometric algorithms use only the geometry of the link, while topological algorithms also use the connectivity of the network. In probabilistic approaches, an error region is first used to determine matches, and then the topology is used when multiple links or link segments lie within the created error region. Advanced algorithms include Kalman Filtering, Bayesian Inference, Belief Theory, and Fuzzy Logic.

Most AVL-based systems use monuments of some kind to compute segment and route travel times. One technique for establishing the timestamps associated with monuments involves filtering the pings to select the one that is closest to the monument. This was the technique employed in selecting the pings displayed in Figure 4-3. There is no control over when the pings are issued (every few seconds) and the expectation is that a ping will be issued at some point in time when the vehicle is near each monument. A second technique involves having the vehicles generate their own "monument-to-monument" travel times. When the AVL-equipped vehicle passes each monument it creates a message packet indicating the monument it just passed, the associated timestamp, the previous monument passed, the timestamp associated with that previous monument passage, and the next monument in the path. In this case, data records akin to the AVI detector-to-detector records are created and can be used to create segment and route-specific travel times. Moreover, in some systems the path followed is also included in the data packet, so the route followed is known as well.

Imputation

Imputation is the process whereby voids in the data are filled by estimation based on data from nearby or similar detectors. The details about data collection, assembly, and cleaning are addressed in the Guidebook or Supplement A. Figure 4-5 illustrates the idea. Concurrent data from nearby sensors are used to estimate a value for the missing data item.

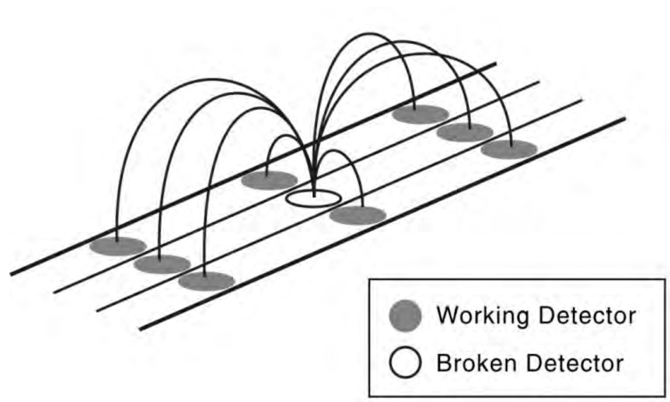


Figure 4-5: Imputation of Traffic Data

The imputed value is computed based on one or more formulas and the input data. Then the value is marked as being synthesized, and when possible, a confidence in the value is saved as well. (See for example Chen *et al.* 2003.)

One of several options involves using occupancies and volumes from the detectors in adjacent locations. Infrastructure-based detectors can be considered neighbors if they are in the same location in different lanes or if they are in adjacent locations upstream or downstream to the bad detector. In this approach, an offline regression analysis is used to continuously determine the relationship between each pair of neighbors in the system. The dependent variable is the flow or occupancy at a detector (when the detector was good) and the independent variables are the flow or occupancy at adjacent detectors. Then, when a detector is broken, its flow and occupancy values can be determined by using the estimated regression parameters. The regression equations can take the form given in Equations 3-1 and 3-2 as follows:

$$q_i(t) = \alpha_0(i, j) + \alpha_1(i, j) \bullet q_j(t) \qquad \text{Equation 3-1}$$

$$k_i(t) = \beta_0(i, j) + \beta_1(i, j) \bullet k_j(t) \qquad \text{Equation 3-2}$$

where:

- *(i,j)* is a pair of detectors,
- q is flow,
- k is occupancy,
- t is a specified time period (for example, 5 minutes), and
- $\alpha_0, \alpha_1, \beta_0, \beta_1$ are parameters estimated between each pair of loops using five days of historical data. These parameters can be determined for any pair of loops that report data to a historical database.

Notably, there are some limitations associated with using linear regression because the observations used for estimation are not independent and the values of flow and occupancy have to be positive.

An imputation need that also surfaced during the project pertains to filling in missing segment travel times for AVI- or AVL-based data. The use of "super segments" seems to be the best way to impute travel times (and travel time distributions) for segments whose endpoint detector is malfunctioning. Figure 4-6 illustrates this idea. If AVI detector B is broken, the super segment A–C provides a way to impute vehicle travel times for both segments A–B and B–C.

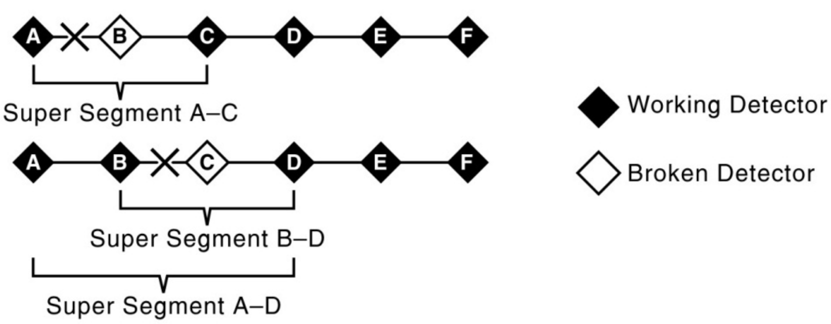


Figure 4-6: Super Segment Examples

When all three detectors are working properly, regression equations can be developed that predict the travel times for A–B and B–C based on the travel time for A–C. Then, when detector B is malfunctioning, these equations can be used to impute individual vehicle travel times (or the average or some other percentile value such as the median) based on the travel time observations between A and C. The same idea applies to the segments B–C and C–D if detector C is malfunctioning, only there are two super-segments that could be used to impute the missing values (i.e., super segments A–D and B–D). The super segment that is the best predictor of the travel times on the subject segment (i.e., which might be either B–C or C–D) should then be used to impute the missing travel times.

Where infrastructure-based detection is present, one can use the point speeds (spot rates) from those sensors to adjust and/or cross-check the imputed distribution of travel times. Also, equations (e.g., regression) can also be developed to use the system detector data directly in doing this. (Of course, the infrastructure-based point speeds can be used directly to estimate average travel times for the subject segments.)

A temporal median approach, equivalent to the one described for option 3 for infrastructure-based imputation, can also be utilized. A temporal median is the median of the historical, non-imputed route travel time values reported for that segment for the same day of week and time of day over the most recent several weeks. Finally, it should be noted that imputing data when there are too many non-functioning sensors can reduce the value of the imputation and the results.

Non-Recurring Event Data

It is important to collect non-recurring event data in real-time, rather than waiting until after-the-fact. These data tend to be perishable and consequently hard to find after the event is over.

The primary non-recurring events that affect reliability are incidents, weather, construction, and special events. The ability of agencies to collect data on these events, and the types of data they can collect, will vary between locations.

Transportation Incidents

There are many viable sources for collecting incident data. Most state (and some local) emergency response agencies use Computer Aided Dispatch (CAD) systems to respond to incidents; these systems have feeds to connect with. The benefit of this data source is that it is in real-time, but the drawback is that the data has not been cleaned (for example, incident locations may not be clearly specified and durations may be inaccurate). Many State Departments of Transportation have databases with cleaned up incident records for state highways (for example, the Caltrans Accident Surveillance and Analysis System), for the purpose of performing detailed analyses. These sources can also be leveraged for reliability monitoring. Another potential source for incident data is the local Transportation Management Center (TMC), where operators usually enter incident information into their management software. Finally, private sources such as Traffic.com often collect incident data at a high level of specificity from various sources, including video, mobile (patrol) units, and emergency communication frequencies. While many potential sources for incident data exist, it should be noted that these data are often incomplete, many times lacking severity indicators, clearance times, and exact incident locations.

The following variables can be used to relate traffic incidents with travel time variability: location, date, type, starting time and duration, full time to clearance, severity, and lanes impacted. In addition, transit incidents, such as bus collisions or disablements can disrupt the operations of a transit system and cause major delays. Such incidents are increasingly being detected by the AVL systems used by transit agencies.

Weather

One source for weather data is existing weather stations operated by various governmental organizations or research bodies. For example, the most accurate sources of weather information are the Automated Surface Observing System (ASOS) and Automated Weather Observing System (AWOS) stations maintained and used for real-time airport operations by the Federal Aviation Administration (FAA). Another good source is an online interface from the National

Climatic Data Center (NCDC) of the National Oceanic and Atmospheric Administration (NOAA), which provides hourly, daily, and monthly weather summaries for 1,600 U.S. locations. For mountainous rural areas, the major sources of weather-related delay are closures and chain control stations. These data are frequently available from rural traffic management centers, although collecting feeds of such data is rare and problematic. One of the richer sources of this data may be Highway Advisory Radio (HAR) networks, which broadcast closure and chain control locations and are frequently available via statewide feed. Any weather data obtained from sources not directly on a monitored route will have to be associated with nearby routes in the system.

Another option for collecting weather data is to directly install Environmental Sensor Stations (ESS) at key roadway locations. Many states have used these to build Road Weather Information Systems, which archive weather data and use it in roadway-related decision making.

The following variables can be used to relate weather with travel time variability: air temperature, type of precipitation, amount of precipitation, visibility, wind speed, pavement temperature, and surface condition.

Transit agencies can use similar methods to monitor weather conditions and develop operational plans that can help them deal with potential disruptions in service and variability in travel times during a variety of adverse weather events.

Work Zones

There are a few different sources for construction-related lane closures. Many states have lane closure systems that serve as a communication interface between the contractors and state agencies to facilitate lane closure management; this data source can be obtained in real-time. Private sources are another option; for example, Traffic.com reports both scheduled and unscheduled construction events. Another option is to manually obtain construction-related information from changeable message sign logs or feeds.

The following variables can be used to relate work zones with travel time variability (see also Haseman *et al.* 2010): start time and duration, start and end locations, and lanes impacted.

Special Events

One option for special events is to manually review calendars for major event venues near a route. Another option is to obtain event data from TMCs, many of which collect event logs to know when and where to activate event-based signal timing plans.

The following variables can be used to relate special events with travel time variability: location, routes affected, duration, type of event, and attendance.

Data Storage

The data storage regime for the non-recurring events is dependent on exactly which variables are collected, and at what granularity. The spatial and temporal resolution of non-recurring events data is an important consideration that impacts the strength of the relationships developed with travel time variability. Data on non-recurring events, to some degree, needs to be aggregated to the same temporal and spatial resolution, in that it all needs to be spatially collected by route and temporally collected for each day in the analysis period. Collecting data on some of the sources at higher spatial and temporal resolutions would lead to more accurate analysis.

The data on non-recurrent events does not need to be stored in the same tables as the route travel times, since the analysis to link travel time variability with its causes is typically a manual exercise. As such, the database for non-recurring events can be uniquely designed to store the data that each agency is able to collect.

Summary

It cannot be over-stressed that high quality data need to be available for a travel time reliability monitoring system to be effective and useful. While it is possible to do some degree of reflective, ex post facto analyses of system performance on the basis of weak data, real-time decision making by system operators and users cannot be done if the data are weak.

This chapter has addressed the issue of collecting and managing the data feeds needed to assess and manage travel time reliability. Two main data feeds are reviewed: 1) the travel time data collected from system detectors and/or AVI- and AVL-equipped vehicles and 2) non-recurring event data. Both are critical to properly analyze and manage system performance. The first provides evidence of the traffic load on the system as well as the travel times being provided. The second indicates whether there were extenuating circumstances under which the system was functioning at the time when the travel times were observed.

CHAPTER 5: SENSOR SPACING AND SAMPLING FOR TRAVEL TIME RELIABILITY MONITORING

Introduction

Operating agencies have historically created monitoring systems that use sensors placed at strategic locations along their freeway networks. Figure 5-1 shows a section of freeway in California where there are 10 sensors in 5 miles, or a sensor about every 0.5 miles. This is a bit dense, but typical. A spacing of a mile or more is common. Of course, putting them at an equal spacing has no particular value; the sensors need to be installed either at locations where congestion rarely occurs – so flow rates can be monitored, like the first, fifth, ninth, and tenth sensor or at places where bottlenecks arise, like the second, third, fourth, fifth, and eighth sensor, so that queuing can be detected.

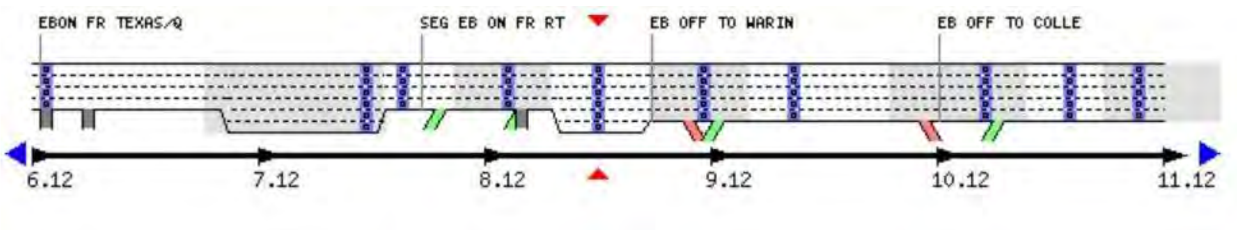


Figure 5-1: Typical Sensor Spacing on a Freeway

The advent of vehicle-based sensing technologies, including those that provide speeds for short TMC segments, are revolutionizing these ideas because sensor placement becomes less of an issue: nothing has to be installed in the roadway surface. Moreover, actual travel times can be observed if the vehicles are re-identified, for example using their MAC-IDs or tag numbers.

In addition, and different from sensing the general health or status of the network, as is the purpose for the sensor deployments shown above, monitoring travel time reliability has a different objective. One needs to sense the status of the system (in time or in space) in a way that produces a defensible image of the travel times that are occurring, as well as their changes in time and space.

For example, Figure 5-2 shows the temporal pattern of AVI-based travel time observations on I-5 in Sacramento, south of US 50 for 2/18/2011 when there was an incident immediately preceding the PM peak. Notice that the rise and fall in travel times during the incident is dramatic: growing from 5 minutes to 35 minutes in the span of 20 minutes and then dropping back to about 7 minutes in another 30 minutes. Also, the travel times in the PM peak, which are typical for this location, rise from 5 minutes (without the incident) to 12 minutes in an hour and a half and then fall back to 5 minutes in another hour and half.

To adequately observe such transients, especially the first, from the incident, one would have to be sampling the travel times every 1 to 2 minutes so that the rapid rise could be observed as well as the subsequent fall. The PM peak that follows could adequately be monitored with samples at every 5 to 10 minutes.

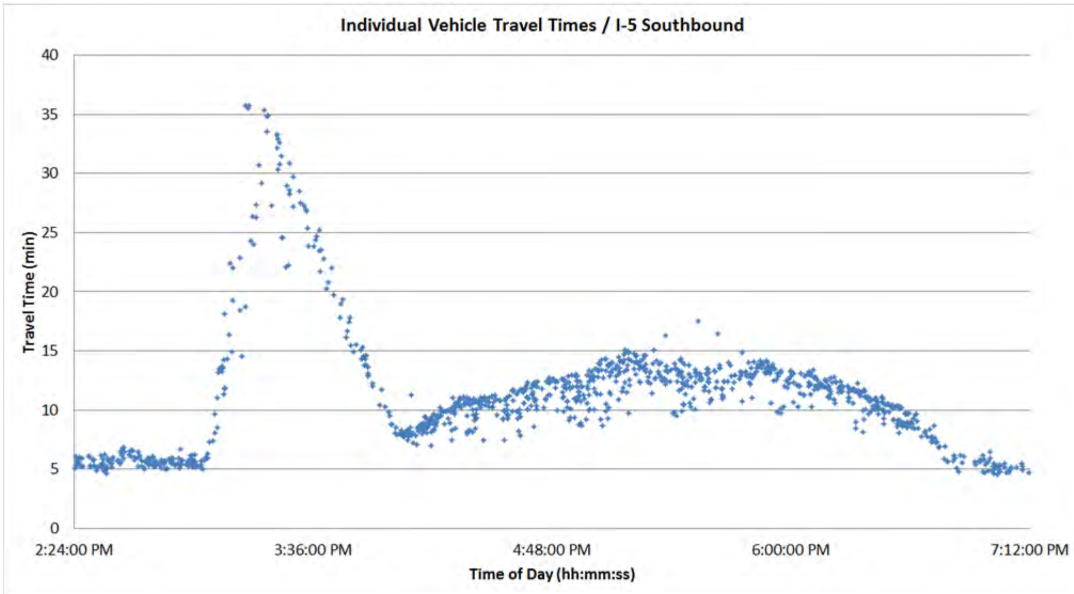


Figure 5-2: An example of two travel time transients – an incident followed by a PM peak

Of course, a difference exists between how many samples are needed *ex post facto* to reproduce an observed waveform, like the ones discussed above, compared with monitoring the travel times that unfold in real time. Not only are the rates of change unknown, but latency (how long will it be after the event occurs) becomes an issue. In the examples above, a monitoring rate of every 15 minutes would be too slow to spot the incident in any meaningful way; and it would be adequate but not ideal to observe the PM peak. On the other hand, an interval of a minute would be adequate for both. In between, a sampling rate of 5 minutes would detect both, but provide a less responsive and less accurate representation of the incident-related transient. These data tend to suggest that a sampling rate of 5 minutes or shorter is likely to be adequate.

In the spatial domain it is more difficult to understand what is adequate. The challenges are twofold. The first is to observe the vehicle trajectories in a suitable manner—in space, not in time—to create defensible travel times. The second is to identify a spacing that allows one to pinpoint the places of reliability trouble, in terms of queuing and momentary slow-ups, so that corrective actions can be taken. Fortunately, the objective is *not* to reproduce the exact vehicle trajectories. As can be seen in Figure 5-3, to do that would require a sample to be taken every 10 or so feet because the transient slow-downs or speed-ups span only 30 to 50 feet, and adequately representing them would require 5 or so observations.

Two concepts are helpful in bounding the lower end of the spatial sampling interval: the spatial geometry of highway design, and expectations about how long it should take before an incident can be identified. In the context of the first, ramp lengths and weaving sections are rarely shorter than 300 to 500 feet, so detector spacing shorter than this would be difficult to implement. Second, and in a separate dimension, shockwaves travel at rates in the range of 10 to 30 mph (15 to 45 ft/sec), so sensors placed 500 feet apart would be able to detect growing queues 10 to 30 seconds after their formation; at 1,000 feet, it would be 20 to 60 seconds.

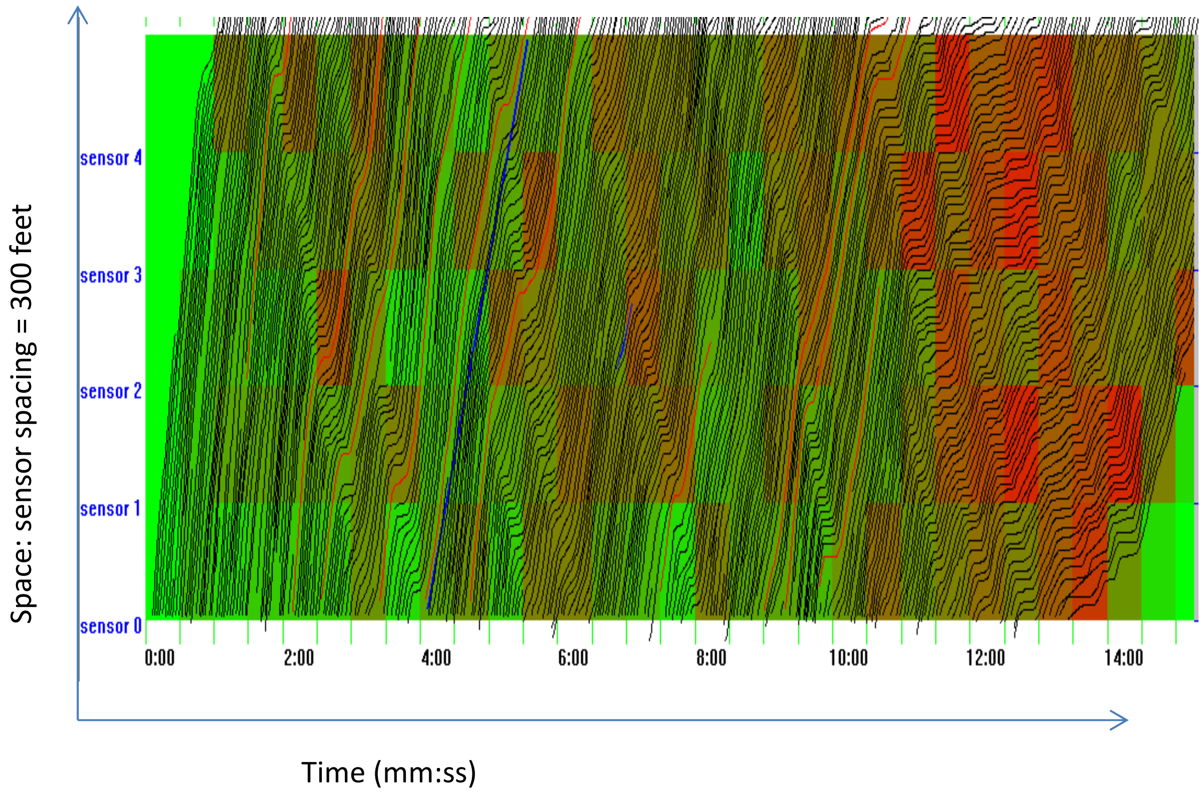


Figure 5-3: Vehicle trajectories in space and time

A Formal Technique

To treat the topic more formally, a procedure focused on the information contained in the sampled data and the ability of the sampled data to reproduce the actual, underlying waveform can be used to gain a sense of how closely the detectors need to be spaced.

The questions that need to be addressed are three-fold:

1. What criteria should be used to determine the sampling rates?
2. What methodology can be used to approximate continuous time series from discrete data samples?
3. How should minimum and practically acceptable temporal and spatial sampling rates be selected?

Quantifying Information Gains

A fundamental question is how to select a measure or a set of criteria that can quantify information gain or accuracy improvement at various locations. For link travel time estimation applications, "link traffic flow volume", "O-D flow coverage" and "link travel time estimation errors" have been widely used as criteria for determining the priority of point detector locations.

In comparison, the essential goal of traffic sensor network design for travel time reliability monitoring applications covers not only reducing average estimation errors for link travel times (see Park *et al.* 2007, Lyman and Bertini 2008), but also capturing the day-to-day and within-day dynamics under both recurring and non-recurring conditions. If the day-to-day or within-day travel time distributions are expressed in terms of probability density functions (PDFs) or cumulative density functions (CDFs), then the criteria of minimizing the average link travel time estimation errors might not adequately emphasize, and possibly just ignore, many non-recurring and important random sources such as incidents. In this study, a Kolmogorov–Smirnov test (K–S test), a nonparametric test for the equality of continuous, one-dimensional probability distributions, can be used to see if the CDFs constructed from sample sequences significantly differ from the ground truth CDFs of travel times under different sensor spacing and reporting configuration scenarios.

Approximating Temporal Patterns from Discrete Samples

If traffic measurements (from a continuous traffic process) are available at some time interval (say, 30 s) and spatial spacing (say, every 0.1 mi), one can strive to select a pattern smoothing method that will identify statistically significant system-wide trends (due to incidents/weather conditions or special events) while filtering out the noise associated with driving behavior or measurement errors. A wide range of time series-based methods exist for traffic state estimation, including autoregressive moving average models and Bayesian learning models, as well as Kalman filtering. Overall, the above methods predominantly operate in the time domain and are suitable for estimating time-dependent dynamics. However, these methods face modeling difficulties in identifying the underlying system process (signals) variability, which is compounded by multiple components, such as day-to-day trends, within-day variability and non-recurring events.

An innovative technique adapts a digital signal processing (DSP) method to process the raw travel time measurements, and further use a spectrum analysis framework to transform travel times (analogous to signals in a DSP model) from the time-series domain to the frequency domain, where a large data set will be decomposed into components of different frequencies. Mathematically, the following generalized model is used to fit the travel time series x_t.

$$x_t = a_0 + \Sigma\, [a_k * cos(\lambda_k * t) + b_k * sin(\lambda_k * t)] \text{ (for } k = 1 \text{ to } q)$$

where t is the sampling interval, and the length of the sampling interval $|t|$ can be 1 minute (along time dimension) or 1 foot (along space dimension). (Note that $\frac{1}{|t|}$ is the sampling frequency.) In addition, x_t is the travel time sampled at t, k is a specific wavelength, and a_k and b_k are the magnitudes of the sine and cosine waves for wavelength k. (It is useful to note that sine waves of wavelength L can be identified by using a sampling rate of about $L/8$ or higher. This provides 4 samples in every half cycle.)

Example. The above modeling approach can be applied using standard Fast Fourier Transformation (FFT) techniques. This first example focuses on the time domain. Seven weekdays of travel volume data, represented as the time series in Figure 5-4 are mapped to the frequency domain representation in Figure 5-5 using a standard Fast Fourier Transformation.

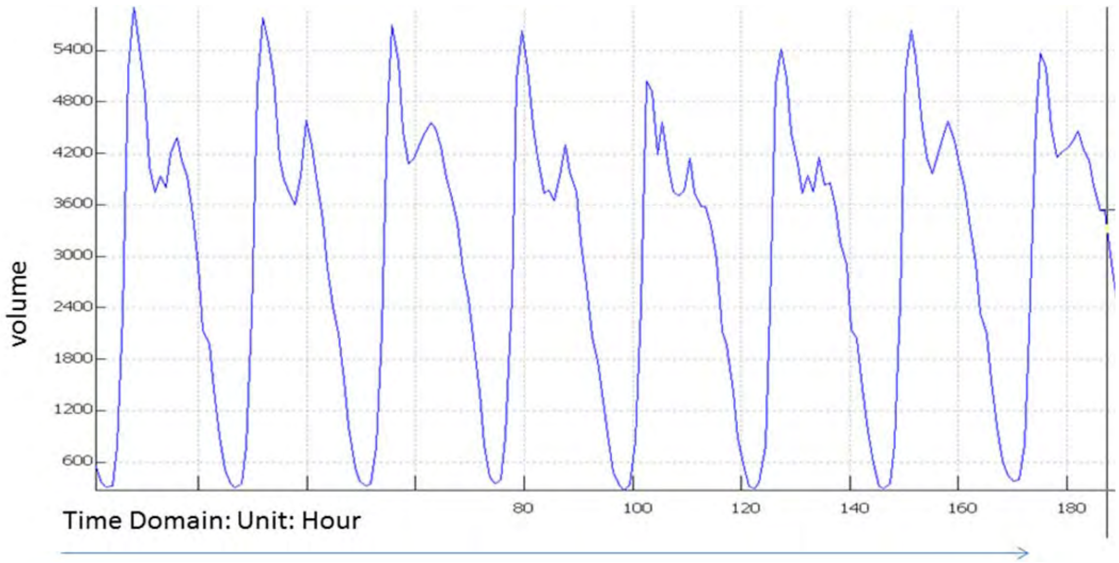


Figure 5-4: Time series of observed weekday volumes, PeMS, 2/1/2006 to 2/10/2006

The spectrum analysis in Figure 5-5 clearly indicates that there are at least 7 to 10 major waves/ harmonics present in the observed data, each one representing a frequency component with a different cycle length. For example, the first wave has a frequency of 1/0.04 per hour, which corresponds to a daily 24-hour cycle.

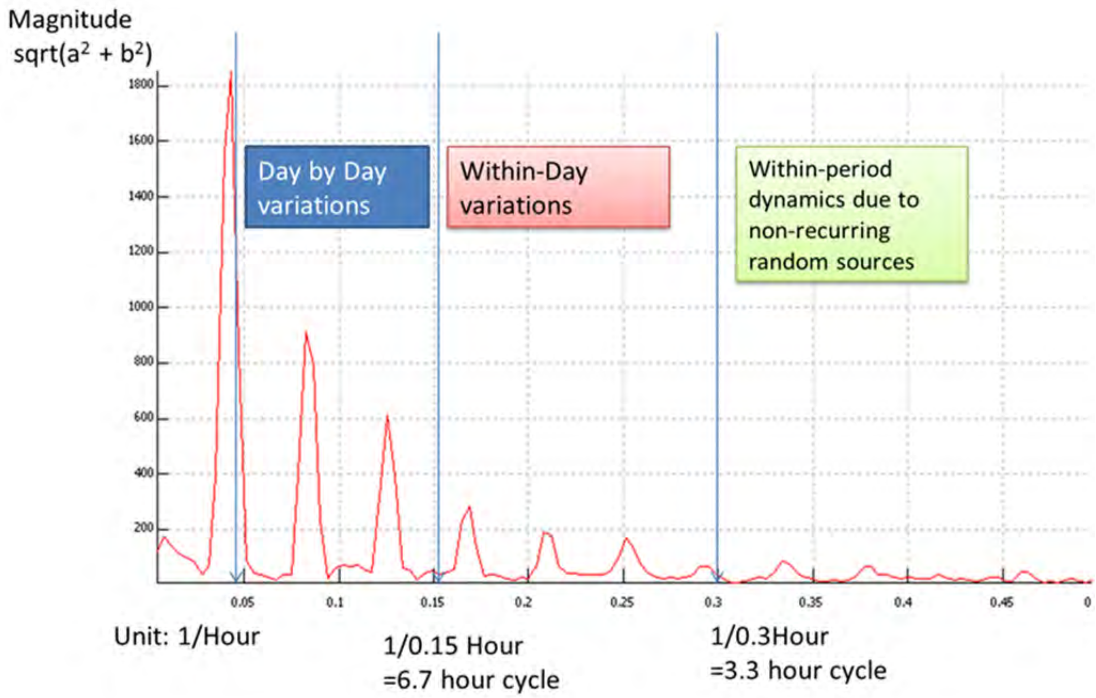


Figure 5-5: Frequency domain representation for travel flow data along time dimension

Using the first three frequency components (up to a 6.7 hour wavelength – a frequency of 1/0.15 cycles per hour), it is possible to capture the day-by-day trends as can be seen in Figure 5-6. (Using the 8 samples per cycle thumb-rule, a wavelength of 6.7 hours can be sensed by taking samples every 50 minutes.)

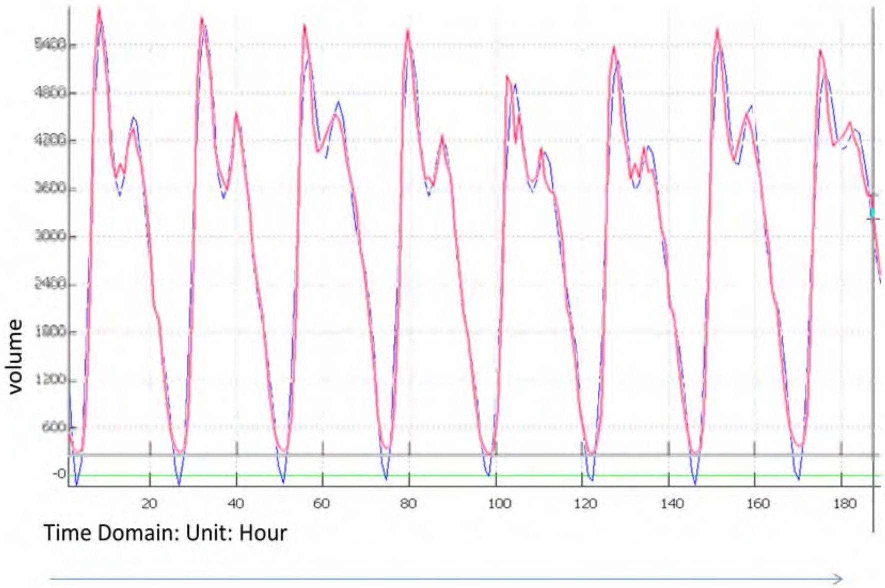


Figure 5-6: Reconstructed time series data that captures day-to-day trends, restored by using a cut-off frequency=0.15, 3 harmonics (Blue = reconstructed time-series, Red = original time-series)

If shorter wavelengths are included, for example down to 3.33 hours, the within-day dynamics can be captured at a finer resolution, as illustrated in Figure 5-7. (To obtain 8 samples of a 3.33 hour wavelength, sampling every 25 minutes would be needed.)

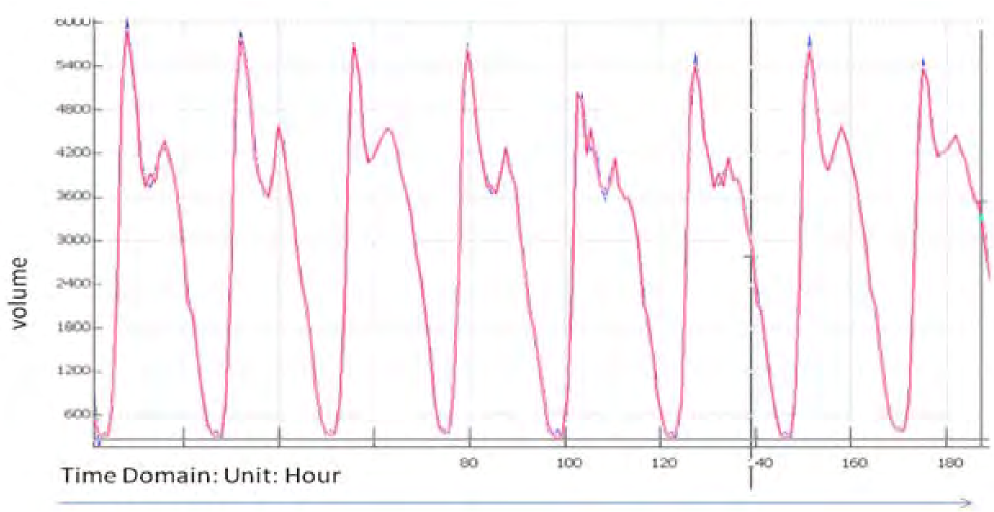


Figure 5-7:Reconstructed time series data that captures within-day dynamics, restored by using a cut-off frequency=0.3, 7 harmonics (Blue = reconstructed time-series, Red = original time-series)

Temporal Sampling Rates

After identifying the distribution of wavelengths within the sampled data (e.g., the PeMS data), one can use the classical Nyquist–Shannon sampling theorem to determine the theoretical minimum sampling rate. That is, if a function x_t contains no frequencies higher than B hertz, it is completely determined from a series of sample points spaced $\frac{1}{2B}$ seconds apart. Moreover, in practical DSP applications, a practically acceptable sampling rate is about $\frac{1}{8B}$, which filters out possible measurement errors and other random factors (e.g., heterogeneous driving behavior in our application).

In the above specific example of traffic flow estimation, Figure 5-7 suggests a system frequency of B = 0.45, that is, a minimum temporal sampling rate of about $\frac{1}{2B}$= 2.2 hours is required to fully capture the within-day variation and a sampling rate of $\frac{1}{8B}$ = 16 min satisfies the practical considerations. Interestingly, the latter coincides with the common practice of 15-30 minute time intervals for sampling traffic flows.

Approximating Spatial Patterns from Discrete Samples

To evaluate the travel time or traffic speed frequency distribution along the space dimension, one can again apply a Fast Fourier Transformation to a sequence of GPS traces, and accordingly identify trends of spatial variations. The notion of spatial variations is somewhat more difficult to comprehend, but once understood, its application becomes sensible and obvious. The following examples illustrate the concept. For instance, if a car is moving in a recursive stop and go pattern every 0.5 miles on a freeway, then its speed frequency profile should include a wavelength of 0.5 miles (from one stop to the next). If a car periodically stops at a sequence of intersections with a spacing of 0.3 miles, then the spectrum analysis should find a spatial wavelength of about 0.3 miles (again stop to stop). It should be remarked that, due to the complex geometric roadway features and traffic dynamics, the spatial frequency distributions might be much more difficult to identify compared to the travel speed frequency distribution on a single location.

Examples. In the following numerical example, 5 GPS traces are used from vehicle trajectories, which cover multiple freeway segments a length of 35,520 feet (6.7 miles), to find acceptable spatial sampling rates. The second-by-second location data are converted to a spatial resolution of 20 feet, leading to a total of 1776 samples. Figure 5-8 gives the spatial frequency analysis results. As expected, the spatial-dimension spectrum pattern is less clear compared to the above time- dimension spectrum pattern in Figure 5-5, although the magnitude of waves decreases as the frequency increases in general.

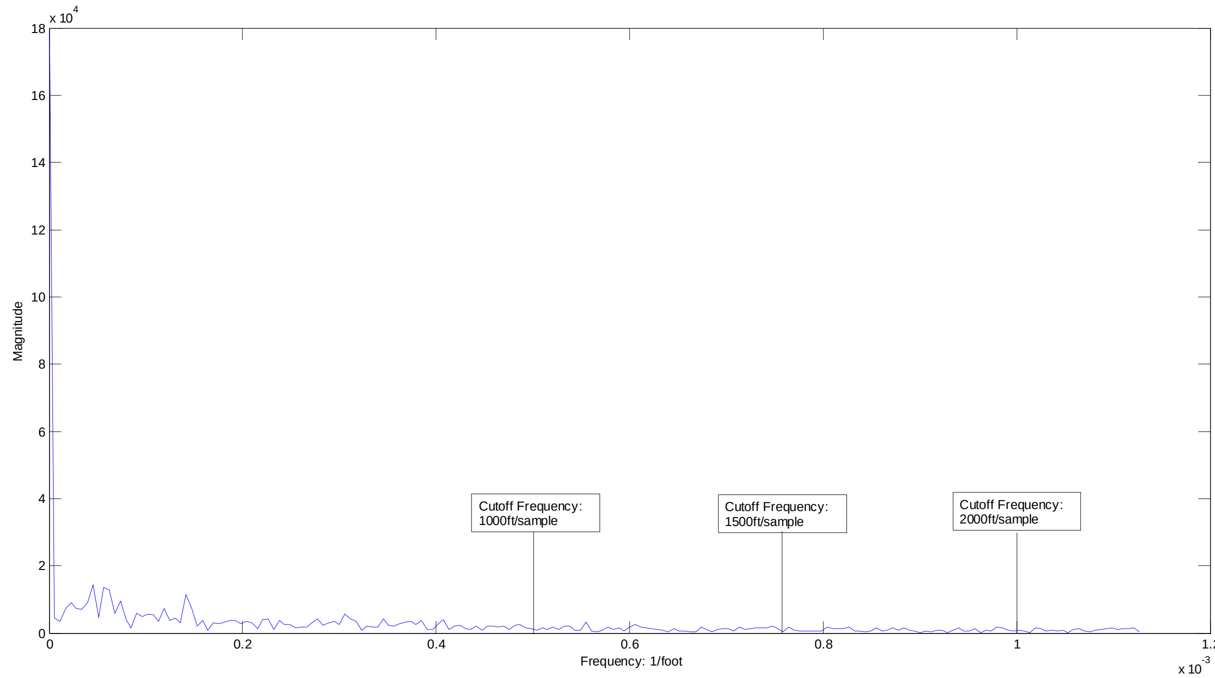


Figure 5-8: Frequency domain representation for GPS location-based speed data along space dimension

As it is difficult to determine the cut-off frequency from the spectrum analysis results, the reconstructed time series curves and KS statistics are compared for three levels of spacing: 1,000 feet, 1,500 feet and 2,000 feet. The analysis results are shown in Figures 5-9 to 5-12.

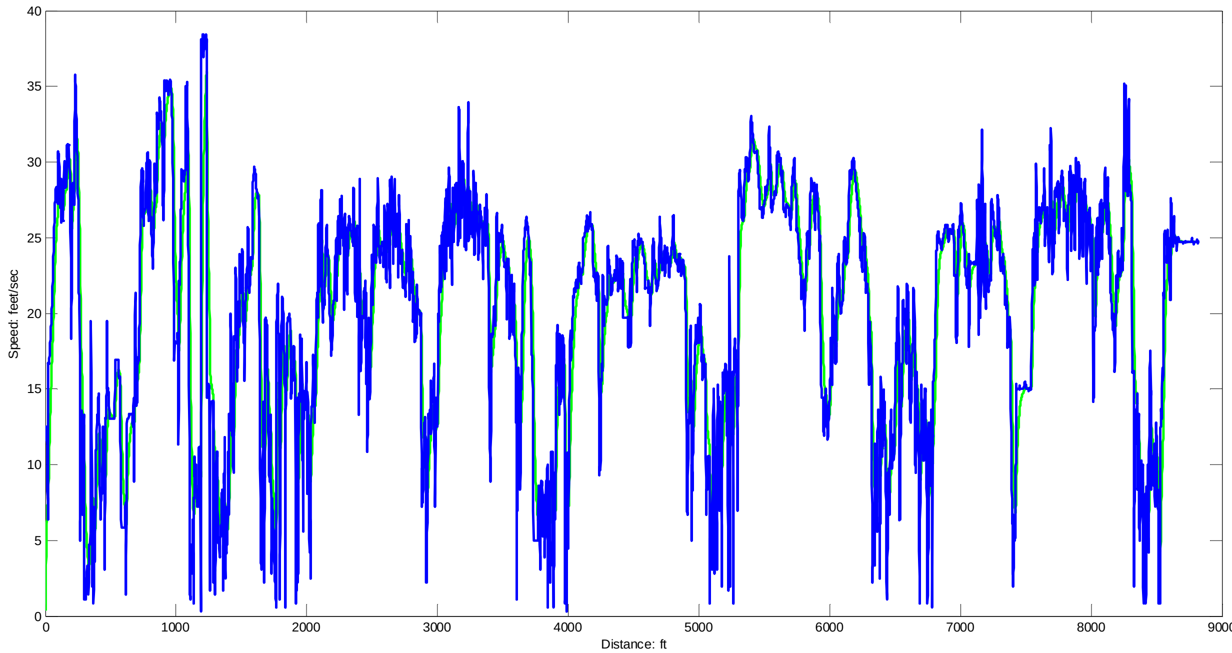


Figure 5-9a: Reconstructed speed time series for GPS traces under cutoff frequency of 1,000 ft/sample

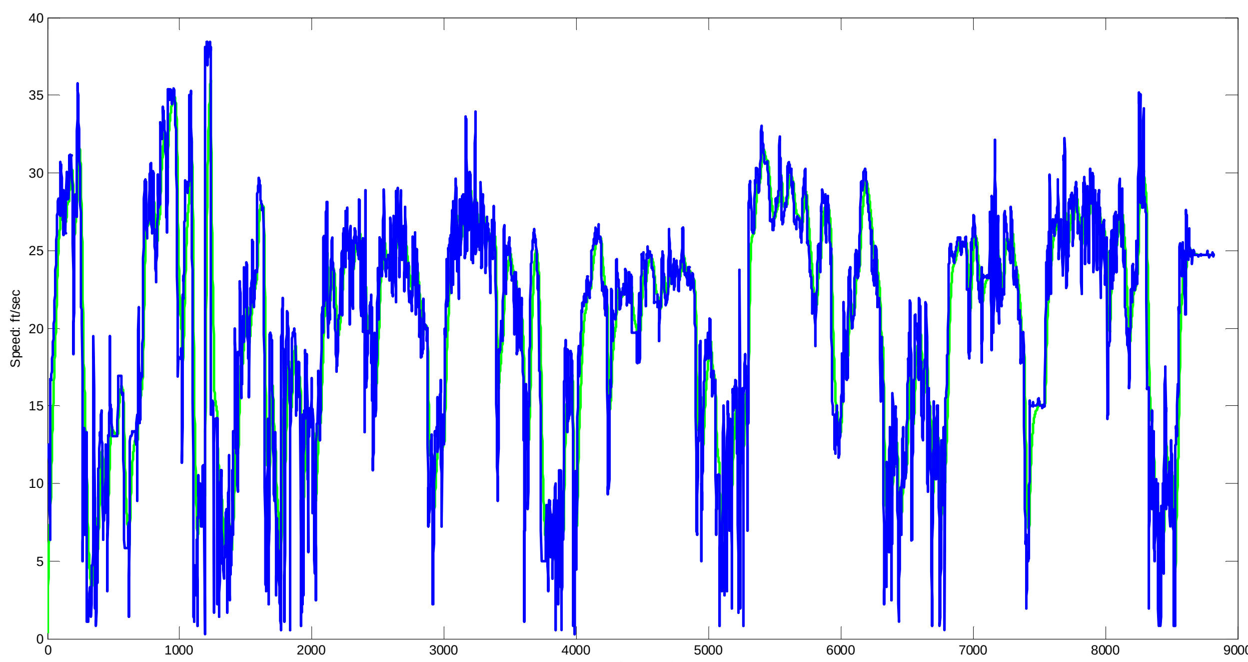


Figure 5-9b: Reconstructed speed time series for GPS traces under cutoff frequency of 1,500 ft/sample

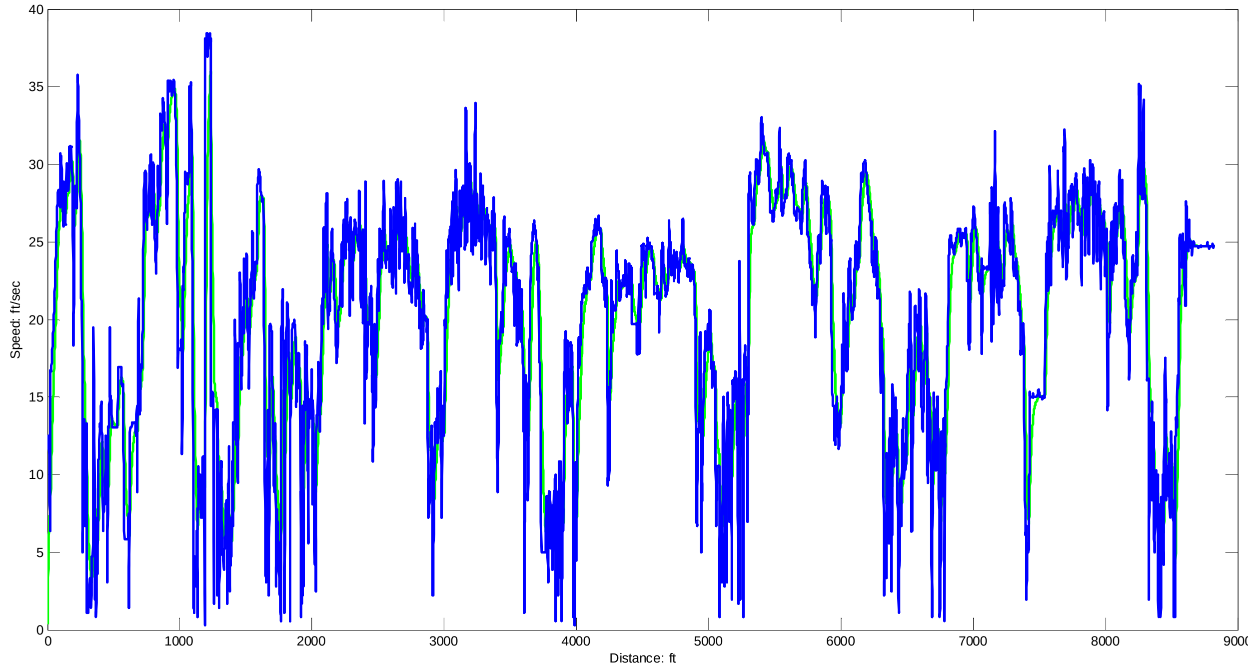


Figure 5-9c: Reconstructed speed time series for GPS traces under cutoff frequency of 2,000 ft/sample

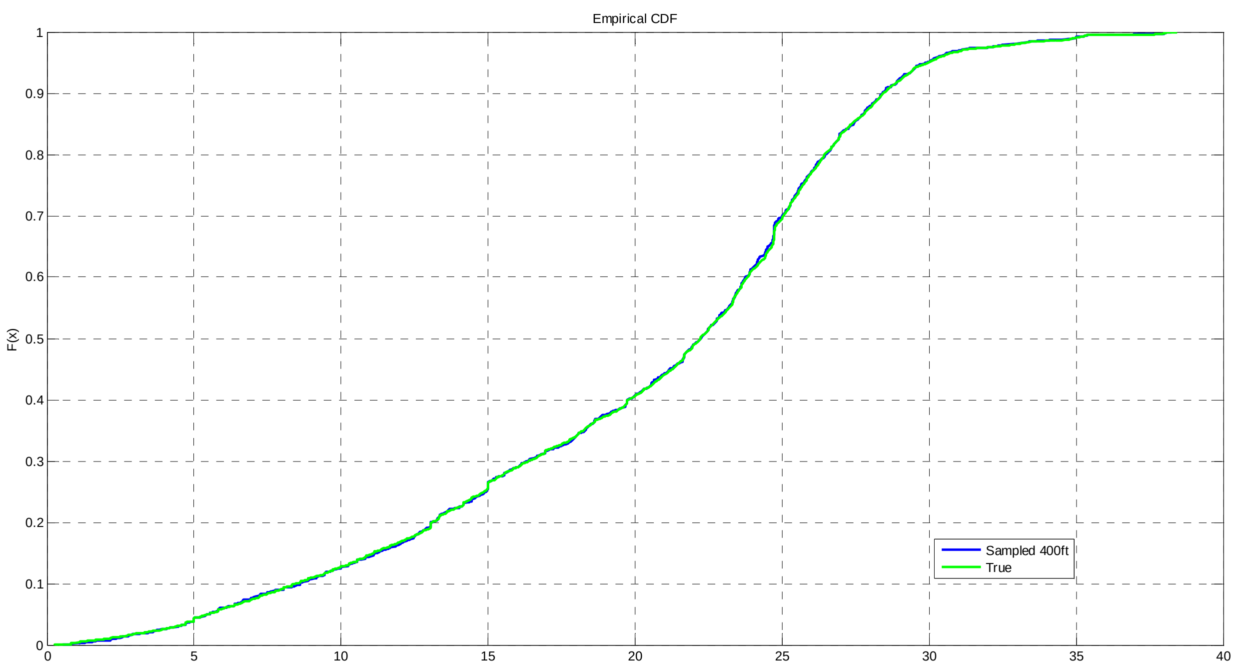


Figure 5-10: Reconstructed and ground truth traffic speed CDFs under sampling spacing of 1,000 ft/sample

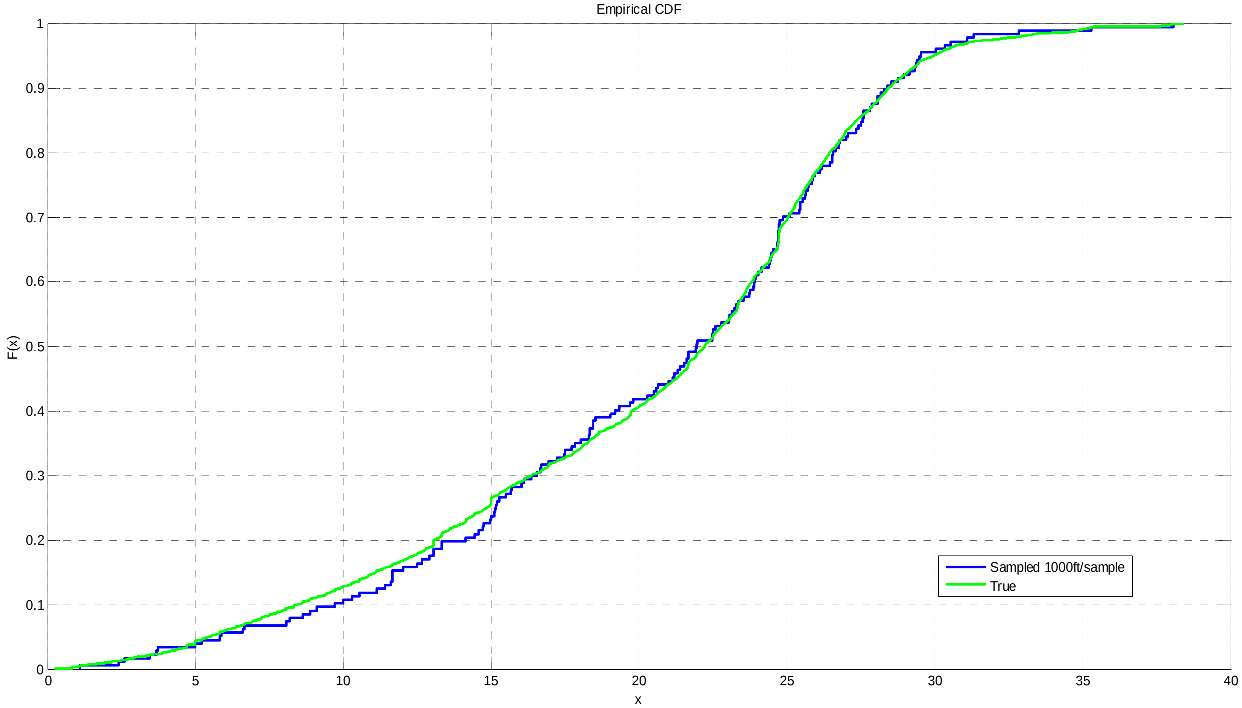


Figure 5-11: Reconstructed and ground truth traffic speed CDFs under sampling spacing of 1,500 ft/sample

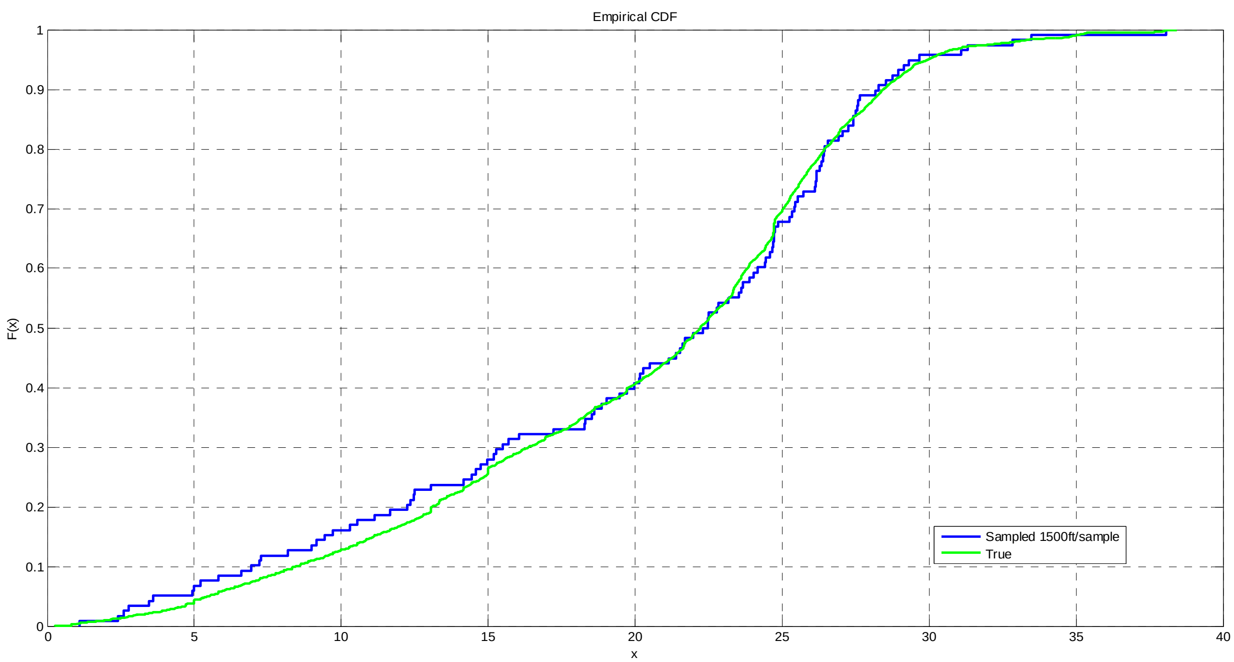


Figure 5-12: Reconstructed and ground truth traffic speed CDFs under sampling spacing of 2,000 ft/sample

For distances of 100 feet to 1500 feet, Table 5-1 lists the absolute, percentage differences and K-S statistics for each of the sampled CDF functions. The table suggests that a cut-off frequency of 1500 feet can deliver statistically sound approximations to the final travel speed CDF curves.

Table 5-1: Percentage Travel Speed Differences and Absolute Differences for K-S Value for Different Detector Spacings

Detector Spacing (ft)	Percentage difference of travel speed CDF	Max Absolute Difference in terms of K-S statistics
100	2.05%	0.01
200	3.05%	0.01
400	4.94%	0.02
500	4.12%	0.02
1000	6.38%	0.04
1500	13.71%	0.05

In the second experiment, the point speed data from the GPS traces are converted to travel rates (1/speed). The spectrum pattern still lacks a clear indication of what the cut-off frequency should be. In general, identifying the cut-off frequency is difficult in its own right and may require a large data set to uncover the inherent patterns. The above analysis results show $\frac{1}{B}$ = 1,500 feet is a reasonable estimation of system frequency, which leads to a suggested minimum spacing of $\frac{1}{2B}$ = 750 feet or a slightly impractical spacing of $\frac{1}{8B}$ = 200 to 300 feet for better approximation results.

Using the NGSIM vehicle trajectory data from I-80 in Oakland, California, it is possible to further identify the vehicle-by-vehicle travel time frequency distribution in Figure 5-13, which indicates 1/B = 0.02 Hz (=50 seconds) as being a logical cutoff frequency. At the minimum sampling rate of $\frac{1}{2B}$, this indicates a need to maintain a sampling interval of 25-30 seconds to obtain high-quality travel time variability distributions.

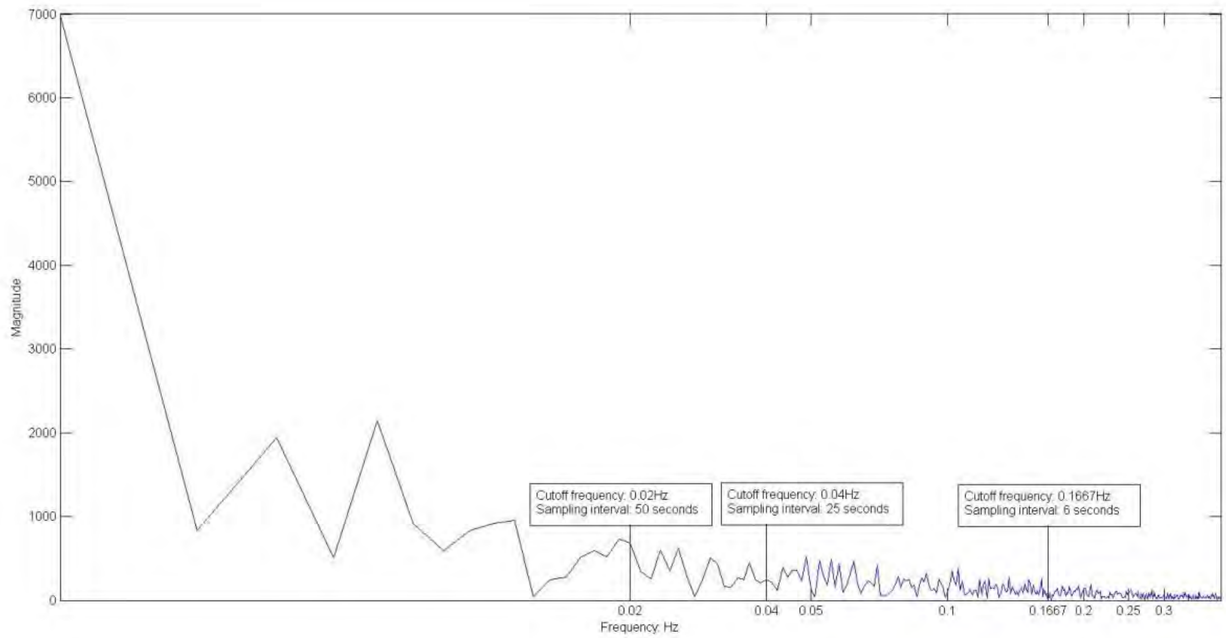


Figure 5-13: Frequency distribution for end-to-end travel time data along a freeway segment, based on the NGSIM data set

By using the sampling rates of 30 seconds and 300 feet, we obtain the aggregated cell-based traffic state representation illustrated in Figure 5-14. Compared to the background vehicle traffic trajectories, which contain significant stop and go shockwaves, the recommended space-time sampling interval appears to reasonably capture the traffic dynamics under this severe congestion condition.

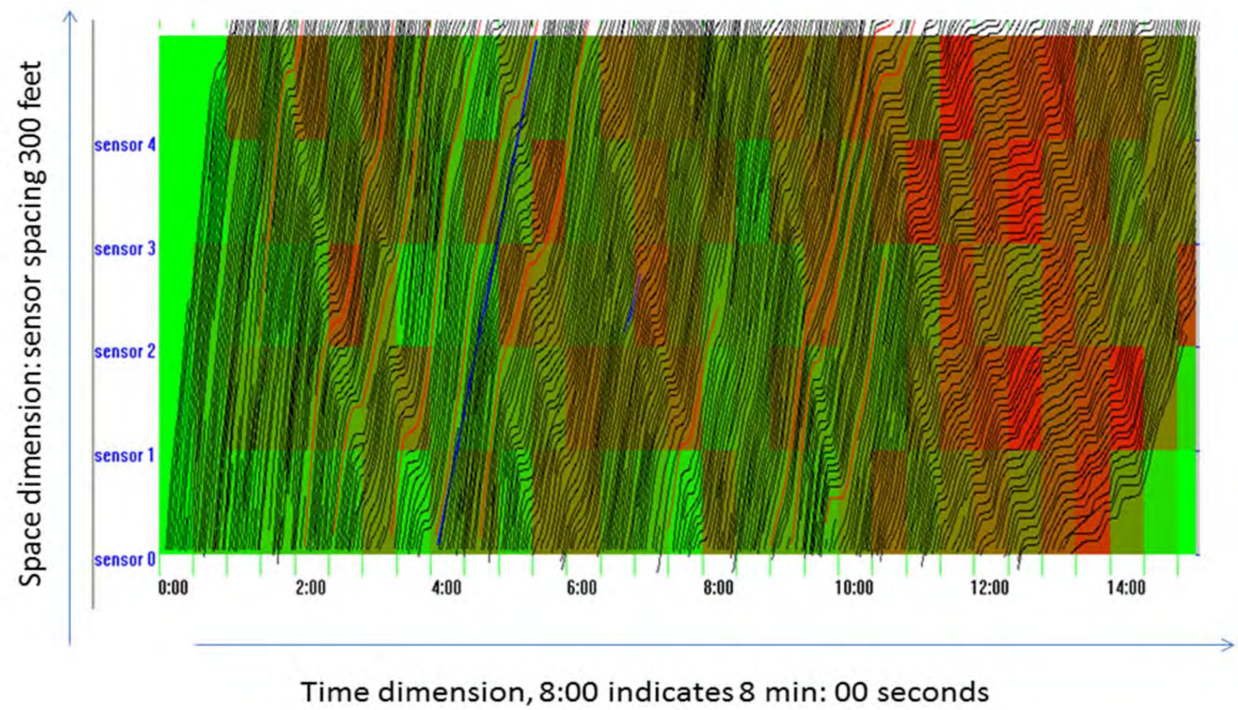


Figure 5-14: Space-time vehicle trajectory and aggregated density representation with a sampling rate of 30 seconds and 300 feet, color coded from green to red representing the aggregated density from low to high

Summary

This chapter has examined the issue of sampling rates in both time and distance to capture acceptable "pictures" of the trends in travel time reliability that are occurring, especially on freeway facilities. The conclusions drawn are that:

- temporal sampling intervals in the range of 1 to 5 minutes should be adequate for most situations where both recurring and non-recurring events occur, although 30 seconds is somewhat better;
- longer sampling intervals can be used where transients are not expected (e.g., off-peak) or where separate means exist for detecting incidents;
- spatial sampling intervals in the range of 750 to 1,500 feet are desirable in locations where queuing transients are expected; and
- longer spatial sampling intervals can be used where queuing is not expected or a separate means exists for detecting incidents.

CHAPTER 6: DATA PROCESSING AND ANALYSIS

Data processing and analysis lies at the heart of the TTRMS. This section provides an overview of the data processing and analysis that are part of the TTRMS. More elaborate descriptions can be found in Chapter 3 of the Guidebook and Supplement B.

Processing Steps

The processing steps employed by the TTRMS are provided by Figure 6-1. The cascading steps transform the raw data into information about travel times and travel time reliability.

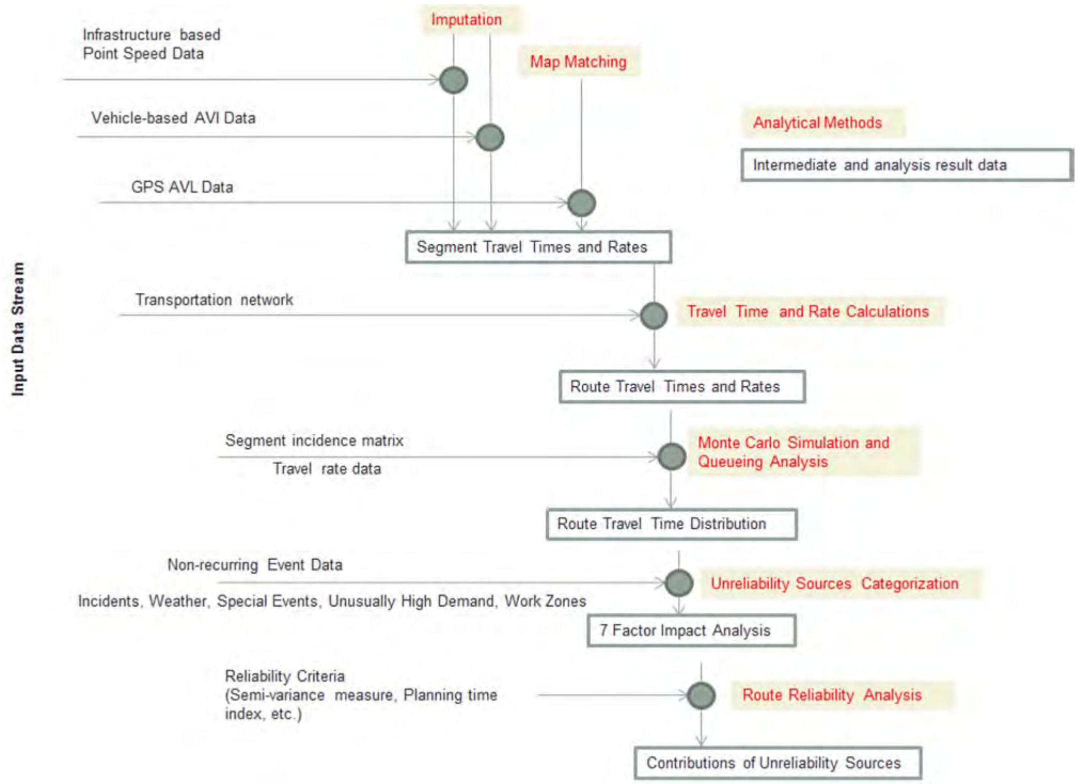


Figure 6-1: Steps in the Reliability Analysis Process

As can be seen, the process starts with the definition of the monuments (monitoring points – real or virtual) – being the locations to and from which travel times will be measured and monitored. As explained in Section 2, they should be located between (and not at) the network junctions so that turning movement delays do not confound the reliability analysis. Undoubtedly, there are logical locations for these monitoring points: lane additions and drops, the location of toll tag readers, of AVI monitors, etc. It seems that most TMC systems already assign segments to system detectors as illustrated by Figure 6-2. More discussion about monuments can be found in Chapter 2.

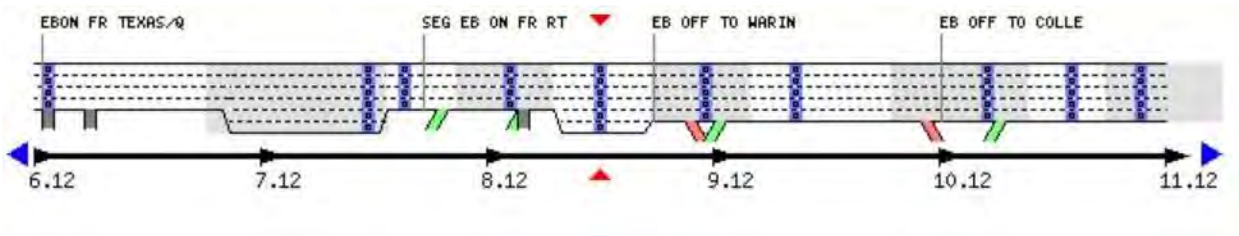


Figure 6-2: Freeway Segments and Segment Boundaries from Legacy Systems

Once the monuments have been established, the incoming data can be processed to prepare segment-level travel times – which are the basis for the reliability analysis and assessment.

The data from infrastructure-based sensors must be enhanced to provide segment-level travel times. Imputation is used to fill voids where data are missing, and then augmented with average speed information where it was not collected directly. (see Wosyka and Pribyl 2012, Van Zwet *et al.* 2003, Zou *et al.* 2008, Zou *et al.* 2009, Shen and Hadi 2012 for a useful discussion on inferring speeds for single loop detectors.) Further inference transforms these spot speeds into average segment level travel times; and those average travel times can then be extended further to develop synthetic distributions of individual vehicle travel times where and when needed.

The data from AVI- and AVL-based systems needs to be processed as well, but in a different way. One has to be sure that the AVI- and AVL-based observations actually pertain to the segments of interest. In the case of AVI data, the sensors are typically located above or adjacent to the roadway, so it is highly likely that the observations pertain to the facility of interest. For AVL-based systems, map matching is required to determine which facilities the observations pertain to. The GPS coordinates are often not sufficiently precise to make this linkage clear. Once suitable observations have been identified, the data can be summarized directly to create segment-level PDFs of the individual vehicle travel times as well as averages (for comparison to and use with the system detector-based data).

The segment travel times and rates are then combined to develop route-level travel times and rates. The combination process is not trivial because strong correlations exist among the times observed on adjacent segments, but it is possible to generate these multi-segment density functions. (Of course, if the AVL or AVI data are sufficiently numerous that direct observations of route level travel times exist, then the travel times and travel time distributions can be observed directly.)

Non-recurring event data are collected from external sources so that the operative conditions in the network can be correctly characterized for any given point in time and location. Variable values based on these data are added to the segment and route-level travel time data so that the effects of these conditions can be ascertained (and the effects of mitigating actions assessed). Congestion-level information also needs to be added so that its impacts can be seen and assessed. Combinations of congestion level and non-recurring events form regimes, the principal categories of system operation for which the reliability performance is differentiated.

A view of the processing steps is found in Figure 6-3. This portrayal ties together the four types of data feeds. It also shows how those feeds have to be processed to generate segment and route-level PDFs.

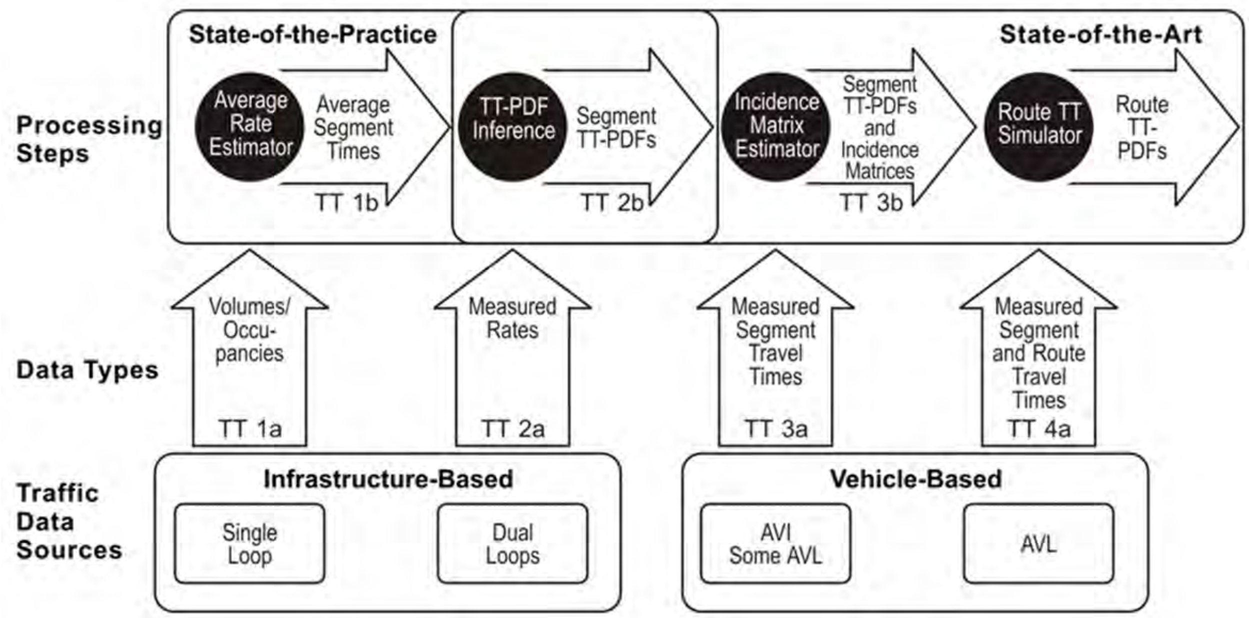


Figure 6-3: Data Processing and Integration to Yield Segment and Route PDFs

Yet another perspective is provided by Table 6-1. The narrative in the table indicates how various types of information can be obtained from the various data feeds typically available.

Table 6-1: Creating Reliability Information from Various Data Feeds

Generating PDFs and Measures of Interest				
Enhancement or Metric	Data Type			
	Type #1	Type #2	Type #3	Type #4
	Single Loops	Double Loops	AVI	AVL
Passage Times	Not applicable	Not applicable	Use signal strength or bounce-back time	Use passage times for Lat/Lon locations
Average Spot Rates	Use occupancy, flow, and assumed vehicle length	Directly computed by the sensor	Not needed	Not needed
Spot Rates for Individual Vehicles	Cannot be obtained	Could be obtained	Use signal strength or bounce-back times	Use GPS speeds at Lat/Lon locations
Average Times or Rates for Segments	Combine adjacent sensor spot rates	Combine adjacent sensor spot rates	Determine from adjusted IV-PDFs	Determine from adjusted IV-PDFs
Segment IV-PDFs	Use average times or rates and IV-PDF typical of the traffic conditions	Use average times or rates and IV-PDF typical of the traffic conditions	Adjust the observed IV-PDFs to account for unequipped vehicles	Adjust the observed IV-PDFs to account for unequipped vehicles
Incidence Matrices	Base on field studies or similar segment-to-segment flow conditions elsewhere	Base on field studies or similar segment-to-segment flow conditions elsewhere	Use equipped vehicles on adjacent segments	Use equipped vehicles on adjacent segments
AVG-PDFs for Segments or Routes	Add estimated segment or route times or rates	Add estimated segment or route times or rates	Compute from segment or route IV-PDFs	Compute from segment or route IV-PDFs
IV-PDFs for Routes	Simulation based on IV-PDFs and Coincidence Matrices	Simulation based on IV-PDFs and Coincidence Matrices	Use equipped vehicles or simulation based on IV-PDFs and Coincidence Matrices	Use equipped vehicles or simulation based on IV-PDFs and Coincidence Matrices

Segment Travel Time Calculations

Segment travel times and their PDFs lie at the heart of the TTRMS. It is via these times and their distributions that reliability performance is assessed and improvements over time are monitored. It is also via these data that route-level travel times and PDFs are developed; as well as area- and sub-area-wide aggregate assessments. Hence, it is critical that high-quality segment travel times and their PDFs be developed from whatever data sources are available.

As has been stated before, today most agencies base their segment travel times on speeds obtained from system detectors (loops) and/or third-party sources (e.g., from INRIX). To restate what is probably obvious, in the first case, average spot rates (actually speeds) are being collected at specific locations. In the second, average speeds are related to TMC segments.

Also, as stated before, an advantage to the system (loop) detectors is that they base the speeds on all the vehicles in the traffic stream, not just equipped vehicles. Two disadvantages are that no individual vehicle spot speeds are reported out (although they are observed) and the data do not actually reflect segment travel times.

A growing number of agencies are obtaining data from third party sources. AVI- and/or AVL-equipped vehicles are the ultimate source of the TMC segment data these companies provide, but the reported data do not typically indicate how many vehicles were observed for the values reported or the speeds for the individual vehicles.

In a limited number of instances, agencies are installing their own Bluetooth sensors or tag sensors to obtain individual vehicle travel times. These data truly are segment travel times, but only for the equipped vehicles. Hence, a disadvantage is that only some vehicles are observed, whether it is only the vehicles equipped with discoverable Bluetooth device or those equipped with tags. Hence, there can be a bias in the observations vis-à-vis the overall traffic stream.

This discussion assumes the TTRMS can work with data from both system detectors and individual vehicle monitoring systems. Hence, there are discussions about developing distributions of individual vehicle travel times as well as *average* travel times. As was shown in Figure 3-6, the two are related. Moreover, assuming the individual vehicle travel times are not biased (which they might be), the means from the individual vehicle travel times should match the mean travel times (from the spot rates) reported by the system (loop) sensors. (The study team checked this correspondence on I-5 in Sacramento and found that the two did match closely, but the system detector average travel times tended to lag the averages from the individual vehicle observations and miss some of the variation that occurred.)

The discussion below also assumes that the travel times are "tagged" by additional information that indicates the operating condition (regime) that pertains to each observation. This means the data, however selected from the overall dataset, can be categorized for further analysis based on the regimes represented in the selection. This means the influence of associative (causal) factors can be studied.

Individual Vehicle Travel Time PDFs from AVI or AVL Data

In this instance, the development of PDFs for individual vehicle travel times is straightforward. The one stipulation is that the observation points have to be at both ends of the segment (or some form of interpolation has to be used). For AVL systems, interpolation is almost always required because the vehicles may not report their status exactly at the segment end points unless the vehicles have been told to do that. (In some AVL systems, they can be instructed to do so.)

It is possible that the observations can be biased vis-à-vis the overall traffic stream if the equipped vehicles can traverse the segment in some manner that the unequipped vehicles cannot or do not. This might be the case if there was a toll booth in the middle of the segment and the equipped vehicles could pass through the toll booth without stopping while the unequipped vehicles could not.

There are also minor issues about whether the vehicle data to employ should be based on time of entry into the segment, time of exit, or some other rule. Of course, when the averages are being computed in real time, the vehicle travel times are not observed until the vehicles exit the segment. Most people seem to use the time of entry as the criterion for selection. The analyses conducted in this study used that rule.

Several types of individual vehicle PDFs can be developed from these data. The first is the PDF for a specific timespan during the day (e.g., a 5-minute time period) based on some period of time (e.g., an entire year). The years-worth of observations makes it possible to examine the extent to which the distribution of travel times (rates) varies, the impacts of congestion when no non-recurring event exists, the impact of non-recurring events, and the consistency that does or does not exist within observations for the same operating condition (regime). The second is the distribution of individual vehicle travel times for some timespan during the day (e.g., 7:00 a.m. to 9:00 a.m. on workdays) as well as a period of time (e.g., a year). Embedded in such data is a mix of both non-recurring event conditions as well as congestion conditions. That is, the data will represent a mix of regimes. A third is the distribution of average travel times for some timespan (e.g., a 5-minute time period) where all of the observations have the same operating condition (regime). An example would be (uncongested with no non-recurring event). A fourth is all the observations for an entire year. This data set would clearly represent a wide range of regimes, which implies it would likely be multimodal. This data set would also be ideal for studying the differences that exist among regimes in terms of the distribution of the travel times.

Individual Vehicle Travel Time PDFs from System Sensor (Loop) Data

Tracking individual vehicle travel time PDFs from detectors is challenging, but it can be done. The task would be simpler if the system sensors reported individual vehicle spot rates (which are observed), but they do not report such data, at least presently. Perhaps in the future they will be able to report such data. The detector would not have to pass back each vehicle speed observation. Rather, it could report the sum of the squares of the vehicle speeds and the sum of the cubes of the vehicle speeds. These two additional data items would be sufficient.

Assuming that only the average speeds and the number of vehicles observed were available, a procedure that can be used is as follows. It requires the conduct of field studies for a limited number of locations and regimes to establish distributions of individual vehicle spot rates (or spot speeds) for typical operating conditions (combinations of congestion levels and non-recurring events). Assuming these field studies have been conducted, then the observed average spot rates (spot speed) and occupancies can be used to find a regime that best matches the current conditions. On the basis of this result, one of the distributions of spot rates can be chosen. The average in the selected distribution can be adjusted up or down so that it matches the average spot rate that has been observed. The resulting distribution of spot rates can then be multiplied by the segment length to estimate the distribution of individual vehicle travel times.

Examples of the spot rate distributions can be seen in data from the Berkeley Highway Lab (a section of I-80 located adjacent to Berkeley, California) where individual vehicle travel times were recorded on selective days in January 2011. Hundreds of observations were recorded for each of several regimes as can be seen from Table 6-2.

Table 6-2: Bluetooth Observations from the Berkeley Highway Lab

Condition	Observation Counts by Day and Condition				
	13-Jan	20-Jan	22-Jan	24-Jan	Total
Free Flow	1183	1446	1727	1566	5922
Transition into Peak	121	328	160	126	735
Transition from Peak	84	310	80	149	623
Peak(Congested)	1099	639	594	552	2884
Total	2487	2723	2561	2393	10164

Figure 6-4 shows the TR-PDFs (travel rates) for the free flow regime overall and for each day. Notice that the distributions are all very similar and the variances are relatively small. The minimum is about 50 sec/mi (about 72 mph), the 50th percentile is at about 70 sec/mi (51 mph) and the 95th percentile is at about 86 sec/mi (about 42 mph). The coefficient of variation is about 0.15. In this instance, this travel rate PDF could be used to estimate off-peak PDFs for individual vehicles for all the times during the year when the facility was lightly loaded.

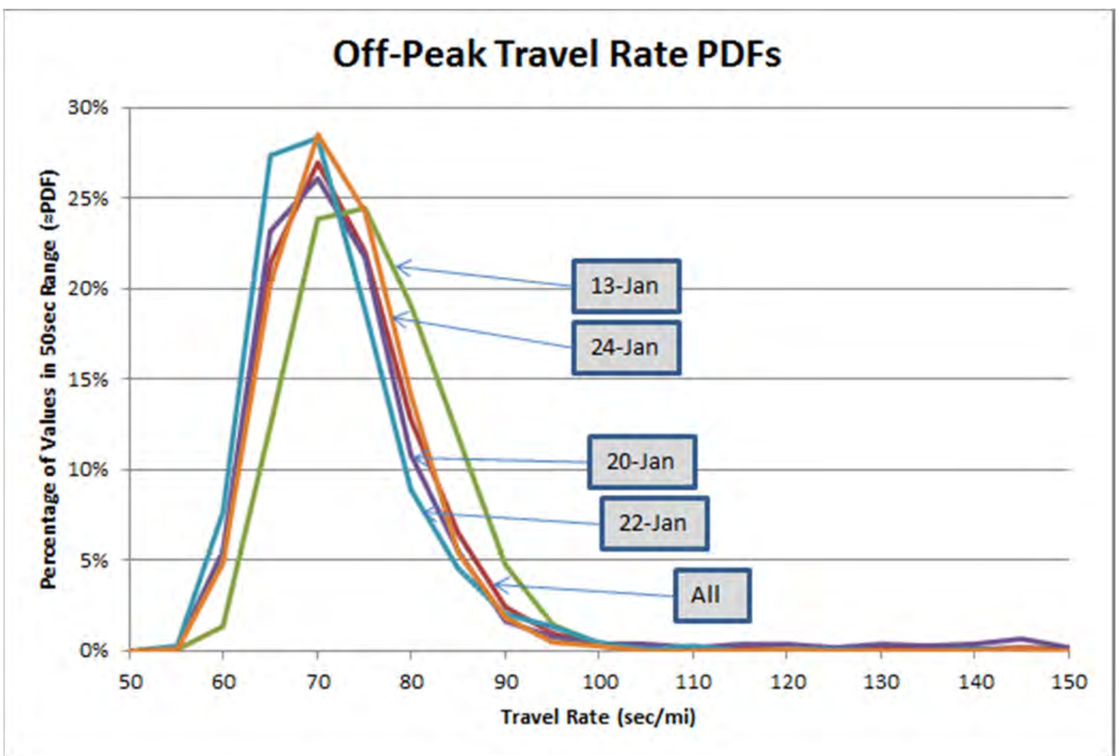


Figure 6-4: Off Peak Travel Rates measured by Bluetooth Sensors for the Berkeley Highway Lab (segment length of about 4500 feet)

In contrast, the travel rate PDFs during the peak period congestion involve significantly larger travel rates, a wider distribution, and much more variability day-to-day as shown in Figure 6-5. The minimum travel rate is about 60 sec/mi (60 mph), the 50th percentile ranges from 150 to 190 sec/mi (19 to 24 mph) and the 95th percentile ranges from 180 to 360 sec/mi (10 to 20 mph). Two reasonable options are 1) to use the overall PDF for all the days and adjust it to the median travel rate (see also Arezoumandi and Bham 2011) being observed at a given point in time or 2) select the PDF whose median travel rate most closely matches the extant travel rate and then adjust that PDF to the extant travel rate.

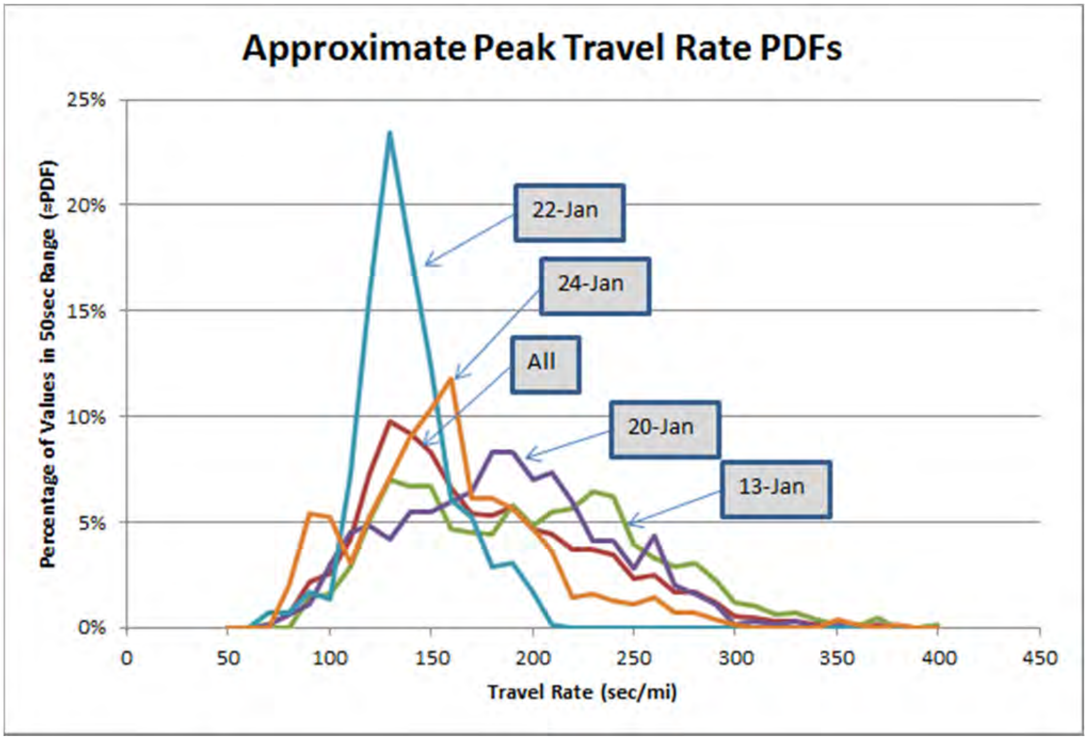


Figure 6-5: Peak Condition Travel Rates measured by Bluetooth Sensors for the Berkeley Highway Lab (segment length of about 4500 feet)

The transitions to and from the peak flow conditions are more challenging, but the data still provide good guidance. Figure 6-6 shows the travel rate PDFs for the transition to peak flow conditions observed on the four days. Evidence of both off-peak and peak conditions can be seen. The density functions appear to be multimodal (bimodal). The minimums are about 60 sec/mi (60 mph), the median ranges from 90 to 130 sec/mi (30 to 40 mph), and the 95th percentile ranges from 160 to 400 sec/mi (10 to 20 mph).

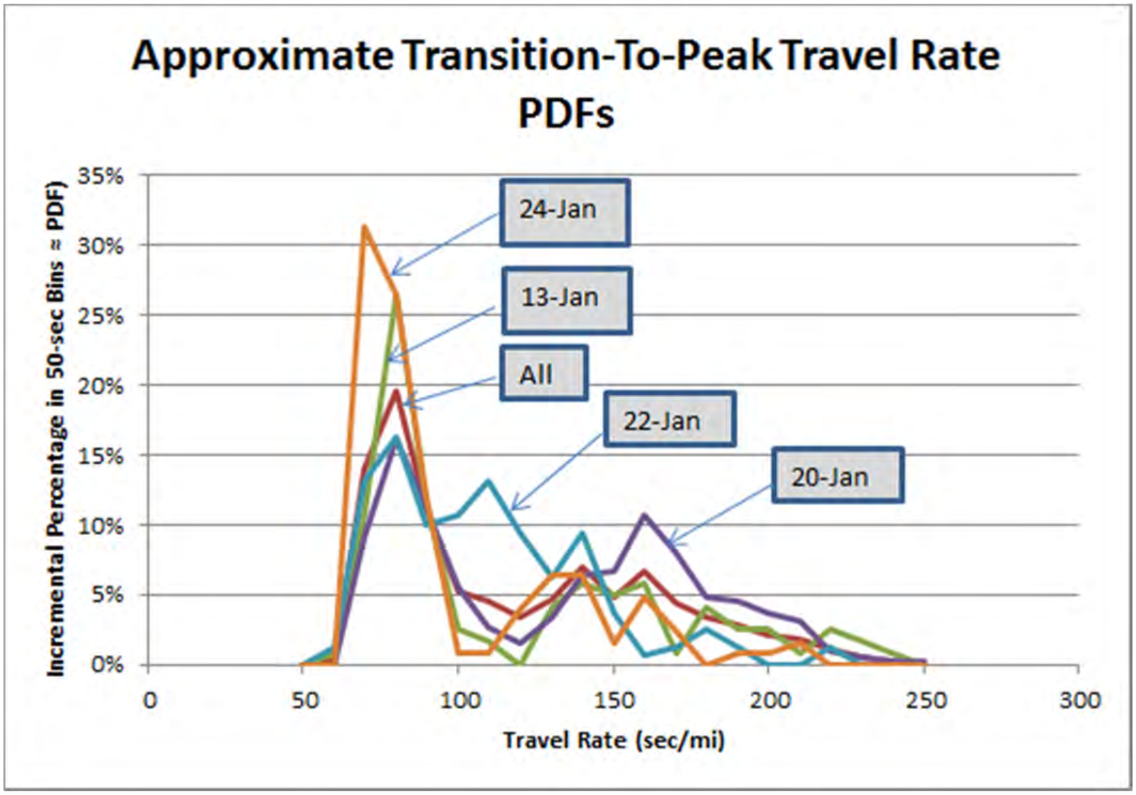


Figure 6-6: Transition to Peak Condition Travel Rates measured by Bluetooth Sensors for the Berkeley Highway Lab (segment length of about 4500 feet)

The same types of summary distributions can be developed from these synthesized individual vehicle travel times as was the case for the actual observations. The one caveat is that the synthesized observations are inherently tied to the underlying system detector observations and the frequency with which those observations exist. For example, if the system detector observations are only available every five minutes, then the synthesized individual vehicle travel time observations are available only every five minutes as well unless additional inference is used to synthesize individual vehicle travel times for intervening points in time.

Average Segment Travel Times from AVI or AVL Data

For AVI and AVL data, average segment travel times can be computed by averaging the individual vehicle travel times. The same types of PDFs identified in the previous two discussions pertain as well to the averages derived from the AVI and AVL data. The one difference is that there may be a bias in the results obtained if the AVI- and/or AVL-equipped vehicles have driving attributes that are different from the unequipped vehicles.

Average Segment Travel Time PDFs from System (Loop) Sensor Data

From system (loop) sensor data, the development of PDFs for average segment travel times is both simple and complex. The reason it is simple is that the data reported back by the system (loop) sensors are average spot speeds at specific locations. For third party data feeds, they are also average speeds for specific short highway segments (TMC segments) the fundamental observations are average speeds based on observations from equipped vehicles.

The complexity arises from the fact that the observations are only spot rates. They are not observations of travel times across the segment. In most cases, the agency associates specific sections of the highway network with each system (loop) detector. This was illustrated by Figure 6-2. Moreover, they assume that the *average* spot rates (inverses of the average speeds) observed by the system sensors are the travel rates for the entire segment. Field studies can be conducted to establish adjustment coefficients if the match is not exact under certain conditions.

Route Travel Time Calculations

Route travel time PDFs are also of great interest in monitoring the performance of a given system. The routes can be short, such as a sequence of segments across a few miles, or long, such as from a significant traffic origin to a significant destination.

Route travel time PDFs are clearly of interest to the system users. It is these travel times that they will actually experience and to which they will relate – rather than the segment travel times. Moreover, route travel time PDFs are very useful when the agency wants to portray to various stakeholders information about the reliability of the system and how it has improved over time.

The challenge with route travel times is that they are difficult to observe. For any specific origin-destination pair, the data tend to be too sparse to allow the estimation of route travel time PDFs directly. Segment PDFs have to be combined to obtain the result. Hence, the question is, how should the segment-level PDFs be combined to produce credible route-level PDFs.

This section describes three procedures for developing route-level PDFs from the segment-level data. The first procedure is based on Monte Carlo simulation of the traffic flow behavior on successive segments combined with incidence matrices for tying together those results. The second procedure uses a lane-by-lane Monte Carlo simulation of a cascading sequence of bottleneck locations to estimate the overall travel time distributions. The third procedure adds together travel times for identical percentiles across the segments to obtain the route-level PDF. All three procedures have value.

The Importance of Correlation

It is clear that correlation exists among segment travel times especially when the segments are short. This phenomenon affects the manner in which one needs to combine segment-level TT-PDFs to form route level TT-PDFs. One cannot add the variances by assuming that the TT-PDFs are uncorrelated.

To illustrate, Figure 6-7 shows scatterplots for individual vehicle travel times on subsequent segments along a six-mile section of freeway in Sacramento, California.

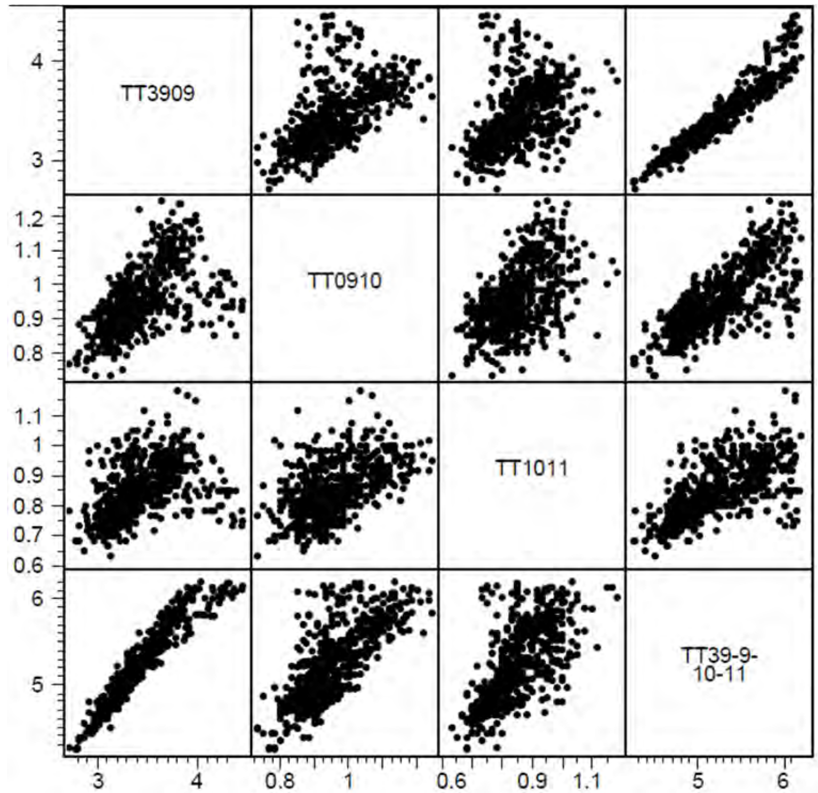


Figure 6-7: Correlations among Individual Vehicle Travel Times for a Sequence of three Segments along I-5 in Sacramento, California

The sequence of AVI monitoring stations is 39, 9, 10, and 11. Travel times are TT3909, TT0910, and TT1011. The scales in minutes for the travel times are shown along the left-hand and bottom borders. The scatterplots above the diagonal show the correspondence between travel times on adjacent segments (TT3909 versus TT0910 and TT0910 versus TT1011) and then two away (3909 versus 1011); and then each one is plotted against the overall travel time (TT3909 versus TT39-9-10-11; TT0910 versus TT39-9-10-11; and TT1011 versus TT39-9-10-11. The scatterplots are symmetric about the diagonal.

Most significantly, not only do the travel times on adjacent segments show a significant degree of correlation but the travel times on each segment are correlated with the overall travel time. The scatter does not increase dramatically as the segments become further separated as would be the case if the travel times were uncorrelated. In fact, the correlation between the travel times is strong as can be seen in the top right-hand scatterplot which shows the correspondence between travel times on the first segment (TT3909) and the overall travel times (TT39-9-10-11). The only way these scatterplots can look like this is if the travel times are tightly correlated.

Monte Carlo Model with Incidence Matrices

This first way to estimate route-level PDFs involves the use of segment-level travel time PDFs and incidence matrices that indicate the correspondence (correlation) between rates on adjacent segments. The method is based on Hu (2011) who studied this idea using a VISSIM model of the Berkeley Highway Lab section of I-80 East in San Francisco. The facility is five-lanes wide and experiences significant congestion during the afternoon peak. Using the model, AVI-like data were generated for two adjacent segments "AB" and "BC" and then those distributions were combined to produce the PDF for "AC".

The incidence matrix was developed using the following procedure, which could also be used by operating agencies conducting short field studies:

1. Observe vehicles traversing AB, BC, and AC and note their travel times (and rates) for AB, BC, and AC.
2. Create a small number (say, 10) of travel rate bins for both AB and BC.
3. Create an incidence matrix that shows the frequency with which specific bin-to-bin combinations of the travel rates arise (e.g., a travel rate on AB in bin X and a travel rate on BC in bin Y.
4. Use the following procedure to generate a PDF for the travel rate on AC:
   a. Select a first random variable x_1.
   b. Select a travel rate τ_{AB} based on x_1.
   c. Identify the *AB* travel rate bin in which τ_{AB} belongs.
   d. Use τ_{AB} and the length of segment *AB* to determine when the vehicle will arrive at the beginning of segment *BC*.
   e. Select a second random variable x_2.
   f. Identify the *BC* travel rate bin from which τ_{BC} should be obtained based on x_2.
   g. Select a third random variable x_3.
   h. Select the *BC* travel rate τ_{BC} on the basis of the lower and upper bounds for the *BC* travel rate bin and the value of x_3.
   i. Compute the travel rate τ_{AC} using the following expression:
      $$\tau_{AC} = (\tau_{AB} * d_{AB} + \tau_{BC} * d_{BC})/ d_{AC}$$

The process needs to be repeated for every successive combination of segments in the route. A sufficiently large number of realizations generated in this manner will result in creating a defensible TT-PDF for the route.

An example of an incidence matrix can be seen in Table 6-3. The left-hand column shows the ranges of travel rates experienced by the vehicles as they traversed the upstream segment *ab*. The top row shows travel rates pertaining to the vehicles as they traversed the downstream segment *bc*. The values in the matrix show the percentages of vehicles that experienced specific combinations of upstream and downstream rates. For example, 24% of the vehicles experienced an upstream rate between 80 and 100 sec/mile *and* a downstream rate between 70 and 80 sec/mi. Interpreted a different way, the matrix also shows that, 37% (8% + 24% + 5%) of the vehicles experienced travel rates between 80 and 100 sec/mi. Of those vehicles, 21% (8/37) experienced travel rates between 60 and 70 sec/mi on the downstream segment *bc*, 65% (24/37) had travel rates between 70 and 80 sec/mi and 14% (5/37) had travel rates between 80 and 90 sec/mi.

Table 6-3: Example of an Incidence Matrix

		τ_{bc} (sec/mi)						
		50	60	70	80	90	100	110
	60							
	80							
	100			8%	24%	5%		
	120			6%	21%	7%	1%	
	140			1%	7%	2%		
	160			1%	3%	2%		
τ_{ab} (sec/mi)	180			1%	2%	1%		
	200				2%	1%		
	220				1%	1%		
	240				1%	1%		
	260							
	280							
	300							
	>300							

The value of using this incidence matrix can be seen in Figure 6-8. It shows the close correspondence between the distribution of route travel rates estimated by the Monte Carlo procedure and the distribution that pertained to the actual vehicle travel rates.

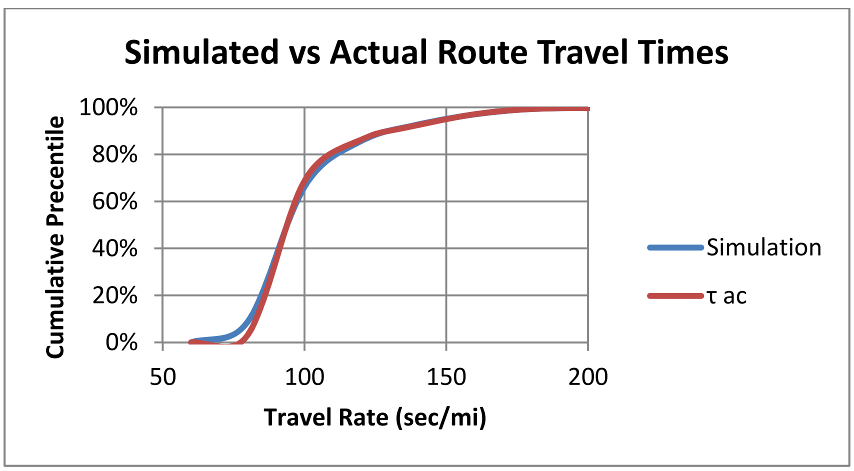


Figure 6-8: Simulated versus Actual Travel Rates for a Route

Point-Queue Based Model

In this second procedure, Monte Carlo simulation is again used, but within a different paradigm. A probe- and point-queue based end-to-end travel time prediction model is used to estimate the route-level travel time distribution. Vehicles pass through the network in specific lanes and their overall travel times are recorded.

The procedure captures the important traffic-related factors that affect end-to-end travel times: the prevailing congestion level, queue discharge rates at the bottlenecks, and flow rates associated with merges and diverges. Based on multiple random scenarios and a vector of arrival times, the experienced delay at each bottleneck along the corridor is recursively estimated to produce end-to-end travel time distributions. The model incorporates stochastic variations of bottleneck capacity and demand, to explain the travel time correlation between sequential links.

Figure 6-9 provides an illustration of a system that has been studied using this model.

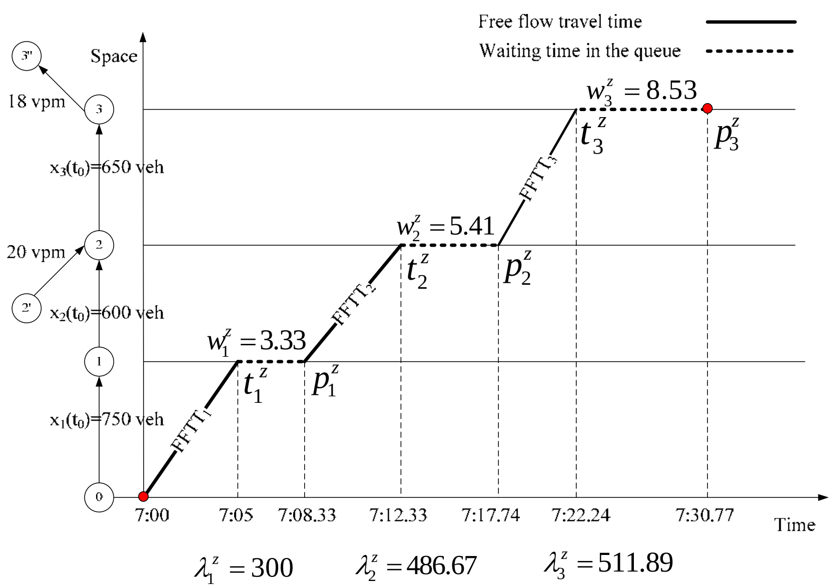


Figure 6-9: Example of a Network Simulated using the Point-Queue based Model

In each Monte Carlo simulation a probe vehicle is assumed to enter the network at a prescribed time (e.g., 7:00 a.m.). It proceeds at free flow speed to the first downstream bottleneck, assumes a position in queue (based on the estimated number of vehicles ahead of it, waits to be discharged, and then, when discharged proceeds downstream to the next bottleneck location. A set of analytical equations are developed to calculate the number of queued vehicles ahead of the probe vehicle as it proceeds through the network. Ultimately, its arrival time at the downstream location is noted and its travel time (and travel rate) recorded. Assembling these simulation run results into a dataset of travel times allows the distribution of travel times and rates to be reported.

An illustration of the results obtained is presented in Figure 6-10. One can immediately see how the model captures the richness in the distribution of travel times that actually arise for vehicles as they proceed through the network and the simulation model's ability to mimic that distribution.

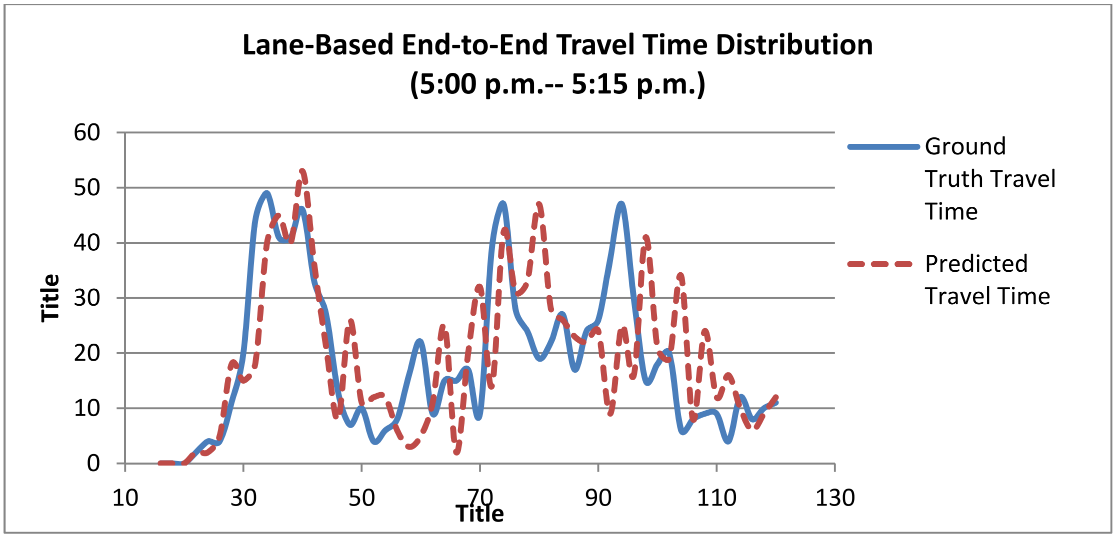


Figure 6-10: Actual Travel Times versus those from the Point-Queue based Model

Co-Monotonicity-Based Model

The co-monotonicity-based procedure is based on the idea that one can add travel times for identical percentiles across the segments to obtain the route-level PDF. When it is possible to do this, the system exhibits co-monotonicity. Co-monotonicity implies that individual percentile values from each of a set of random variables can be added together to obtain the percentile values for the distribution of the sum (Dhane *et al.* 2002a, Dhane *et al.* 2002b).

In a traffic sense, the hypothesis is defensible if drivers are consistent in the speeds they want to achieve and the manner in which they drive. In other words, if a specific driver travels through the network on two (or more) separate days, under similar network conditions, there will be minimal variation in his or her driving behavior.

The possibility of using this technique was tested using Bluetooth data collected on I-5 Sacramento, California. First, the hypothesis of driver consistency was tested. Every individual MAC ID that appeared more than once for a given regime conditions was tracked, and its corresponding average travel time ($\tilde{\tau}_n^i$), standard deviation of travel times ($\tilde{\sigma}_n^i$), and coefficient of variation ($C_v^i = \frac{\tilde{\tau}_i^{n})}{\tilde{\sigma}_i^{n})}$) were computed where $\tilde{\tau}_n^i$ is the average of 'n' observed travel times for a specific MAC ID 'i' , $\tilde{\sigma}_n^i$ is the standard deviation of 'n' observed travel times for the corresponding MAC ID 'i', and C_v^i is the coefficient of variation for the corresponding MAC ID 'i'. Each 'dot' in the Figure 6-11 represents a specific MAC ID, its *x*-value represents average travel rate in seconds per mile, and the corresponding y-value represents coefficient of variation.

One can see that the variation in individual driver travel times under the normal-uncongested regime is almost negligible. The same is true for the normal-low congestion regime. The variation in travel times grows as the network operating conditions become more congested.

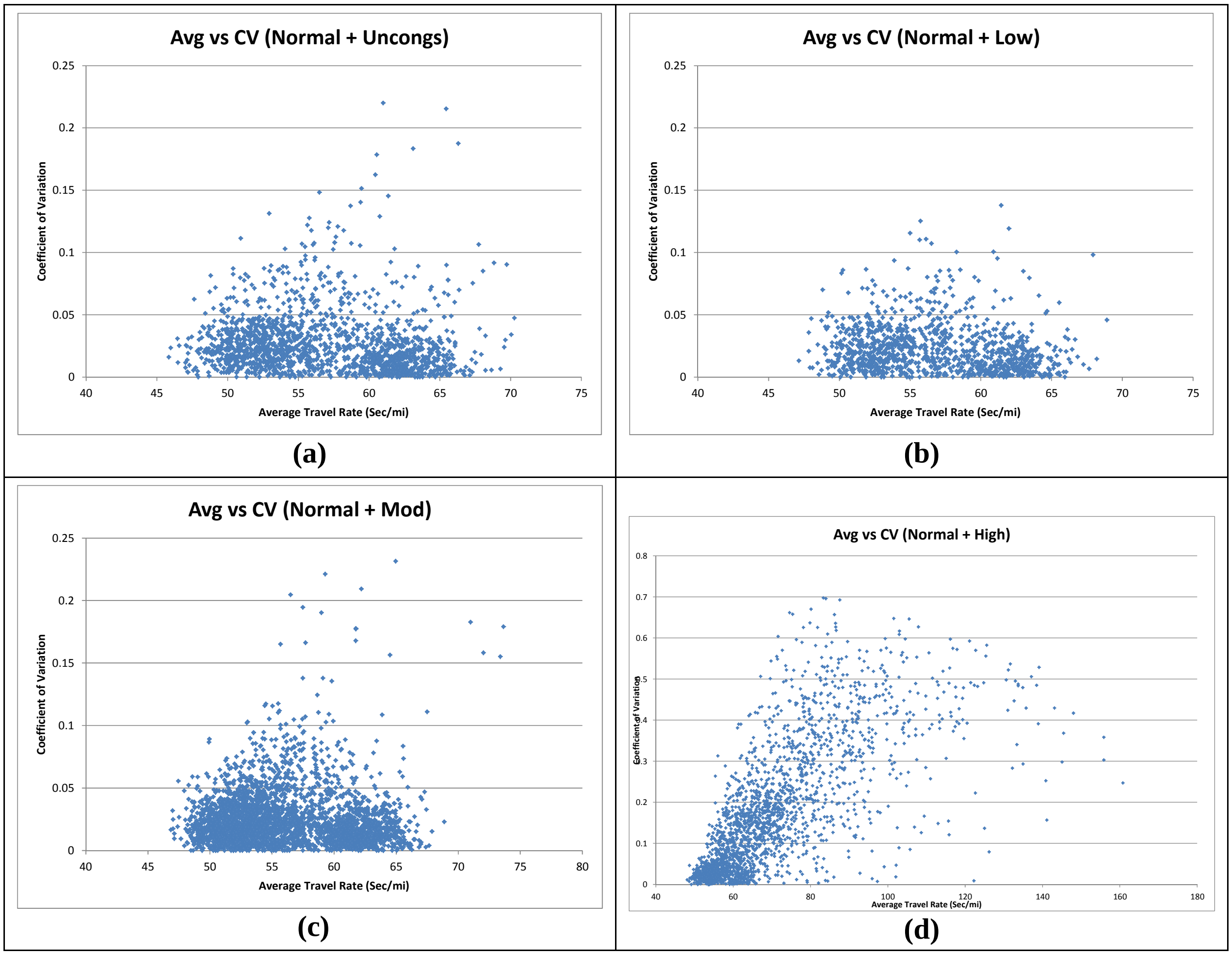


Figure 6-11: Average versus Coefficient of Variation Plot for Different MAC ID's Under Various Regimes

Hence, the applicability of co-monotonicity was tested based on the I-5 data. Table 6-4 provides a comparison of the southbound travel times (for four regime conditions) on I-5 predicted by summing the segment travel times against the overall route travel time. For example, the second, third and fourth columns show the percentile travel times for segments 39-9, 9-10, and 10-11 based on the travel times for those individual segments. The fifth column shows the travel times obtained if these percentile values are simply summed. That is, the values in this fifth column do not represent the percentiles of any underlying distribution. They simply are the algebraic sums of the percentile-based travel times shown to their left. On the other hand, the sixth column shows the percentile travel times that are obtained when the travel times for the overall route are used as the basis for developing the percentile-related travel times. The last column shows that the differences between the naïve sums and the empirically derived results are nearly identical for Uncongested, Low, and Moderate congestion conditions. When congestion on the facility is high, one could notice the differences in almost every percentile are more than 1%, and all percentiles greater than 70% have the difference varying between 2% to 6%.

Table 6-4: A Comparison of actual percentile travel times for a given route against values obtained by summing the travel times for the same percentile on the individual segments

	Segment Travel Times (Uncongs)					
Percentile	Seg 39-9	Seg 9-10	Seg 10--11	Sum	Route 39-11	% Diff
5%	0.810	0.757	0.778	4.450	4.500	1.11%
10%	0.829	0.773	0.814	4.566	4.600	0.74%
15%	0.843	0.788	0.833	4.650	4.667	0.36%
20%	0.856	0.803	0.852	4.733	4.733	0.00%
25%	0.866	0.818	0.861	4.792	4.800	0.17%
30%	0.875	0.834	0.870	4.850	4.867	0.35%
35%	0.884	0.848	0.889	4.916	4.917	0.02%
40%	0.896	0.864	0.908	4.992	4.983	0.18%
45%	0.903	0.871	0.926	5.041	5.050	0.18%
50%	0.917	0.886	0.944	5.125	5.117	0.16%
55%	0.926	0.902	0.972	5.200	5.200	0.00%
60%	0.940	0.916	0.991	5.283	5.300	0.32%
65%	0.958	0.939	1.000	5.383	5.433	0.92%
70%	0.977	0.955	1.019	5.484	5.550	1.19%
75%	0.995	0.977	1.037	5.591	5.625	0.60%
80%	1.012	1.000	1.056	5.692	5.700	0.14%
85%	1.028	1.015	1.074	5.784	5.783	0.02%
90%	1.046	1.038	1.092	5.892	5.875	0.29%
95%	1.074	1.075	1.130	6.067	6.017	0.83%

(a)

	Segment Travel Times (Low)					
Percentile	Seg 39-9	Seg 9-10	Seg 10--11	Sum	Route 39-11	% Diff
5%	0.806	0.743	0.778	4.417	4.467	1.12%
10%	0.824	0.773	0.797	4.534	4.567	0.72%
15%	0.838	0.788	0.814	4.617	4.650	0.71%
20%	0.849	0.803	0.833	4.691	4.717	0.55%
25%	0.861	0.818	0.852	4.767	4.767	0.00%
30%	0.870	0.818	0.861	4.808	4.833	0.52%
35%	0.880	0.834	0.870	4.867	4.883	0.33%
40%	0.889	0.848	0.889	4.933	4.942	0.18%
45%	0.898	0.864	0.908	5.000	5.000	0.00%
50%	0.908	0.879	0.926	5.067	5.067	0.00%
55%	0.921	0.894	0.944	5.150	5.150	0.00%
60%	0.933	0.909	0.972	5.233	5.233	0.00%
65%	0.949	0.925	0.991	5.326	5.350	0.45%
70%	0.968	0.939	1.009	5.424	5.467	0.79%
75%	0.986	0.962	1.019	5.525	5.567	0.75%
80%	1.002	0.985	1.037	5.624	5.650	0.46%
85%	1.019	1.000	1.064	5.725	5.733	0.14%
90%	1.037	1.030	1.092	5.849	5.817	0.55%
95%	1.060	1.061	1.130	6.001	5.950	0.86%

(b)

	Segment Travel Times (Mod)					
Percentile	Seg 39-9	Seg 9-10	Seg 10--11	Sum	Route 39-11	% Diff
5%	0.810	0.743	0.778	4.434	4.500	1.47%
10%	0.829	0.773	0.797	4.550	4.583	0.72%
15%	0.838	0.788	0.814	4.617	4.650	0.71%
20%	0.847	0.795	0.833	4.675	4.700	0.53%
25%	0.858	0.803	0.842	4.729	4.767	0.81%
30%	0.866	0.818	0.852	4.784	4.817	0.69%
35%	0.875	0.834	0.870	4.850	4.867	0.35%
40%	0.884	0.841	0.889	4.908	4.917	0.18%
45%	0.894	0.848	0.908	4.967	4.967	0.00%
50%	0.903	0.864	0.926	5.033	5.033	0.00%
55%	0.917	0.879	0.944	5.117	5.117	0.00%
60%	0.928	0.894	0.963	5.192	5.200	0.15%
65%	0.942	0.916	0.981	5.283	5.317	0.64%
70%	0.961	0.939	1.000	5.391	5.450	1.08%
75%	0.981	0.955	1.019	5.500	5.550	0.90%
80%	0.999	0.970	1.037	5.595	5.633	0.67%
85%	1.014	1.000	1.074	5.717	5.700	0.30%
90%	1.033	1.015	1.092	5.817	5.800	0.29%
95%	1.060	1.061	1.139	6.009	5.950	0.99%

(c)

	Segment Travel Times (High)					
Percentile	Seg 39-9	Seg 9-10	Seg 10--11	Sum	Route 39-11	% Diff
5%	0.861	0.773	0.814	4.683	4.767	1.76%
10%	0.884	0.803	0.833	4.816	4.867	1.05%
15%	0.903	0.818	0.852	4.917	4.967	1.01%
20%	0.919	0.818	0.870	4.991	5.050	1.17%
25%	0.938	0.834	0.889	5.092	5.150	1.13%
30%	0.958	0.848	0.908	5.200	5.250	0.95%
35%	0.977	0.856	0.926	5.292	5.367	1.40%
40%	1.000	0.864	0.926	5.383	5.483	1.82%
45%	1.026	0.879	0.944	5.509	5.583	1.33%
50%	1.051	0.886	0.963	5.625	5.700	1.32%
55%	1.086	0.894	0.981	5.774	5.783	0.16%
60%	1.120	0.909	1.000	5.933	5.883	0.85%
65%	1.153	0.925	1.019	6.084	6.000	1.40%
70%	1.185	0.939	1.037	6.233	6.117	1.90%
75%	1.218	0.955	1.074	6.400	6.233	2.68%
80%	1.259	0.970	1.092	6.583	6.378	3.21%
85%	1.301	0.993	1.130	6.792	6.550	3.69%
90%	1.352	1.023	1.176	7.049	6.750	4.43%
95%	1.431	1.075	1.259	7.466	7.100	5.15%

(d)

Moreover, not only does comonotonicity seem to hold, but it does so in spite of the fact that the density functions are multi-modal. Figure 6-12 shows both the density function for four operating regimes. In all four instances, the density functions are bi-modal. Except for highly congested conditions, the match is strong between the density function obtained by adding the percentiles and that which was actually observed.

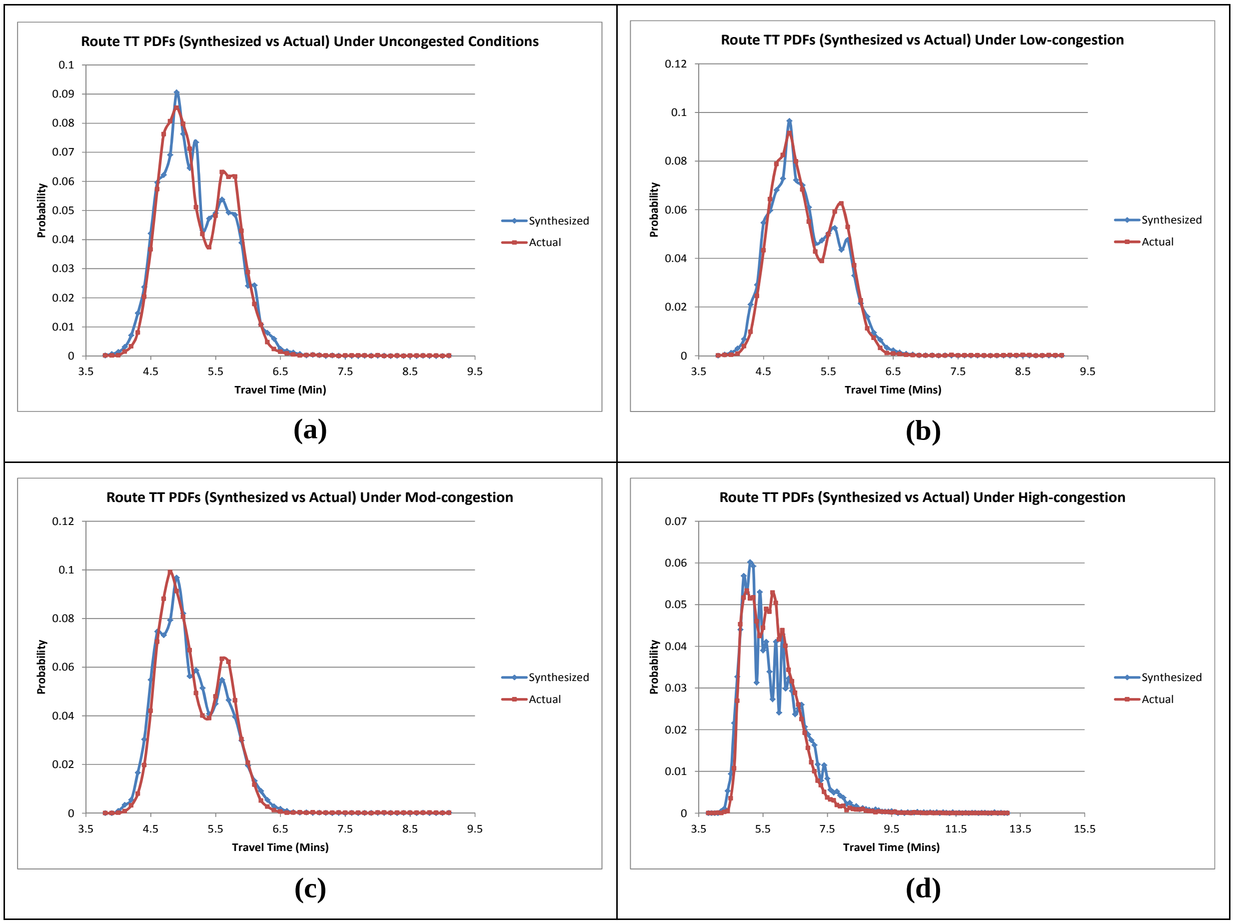


Figure 6-12: Comparisons between the travel time density functions synthesized from individual segment percentiles (squares) and the observed values (dots).

PDFs for Route-Level Average Travel Times or Rates

A common procedure for computing average travel times from infrastructure-based sensor speeds involves the following steps:

1) Calculate the average travel time for the first route segment using the average travel time at a specific point in time;
2) Get the average speed for the next segment at the time the vehicle is expected to arrive at that segment, as estimated by the calculated average travel time for the first route section; and
3) Repeat step 2 until the average travel time for the entire route has been computed.

Put a slightly different way, this procedure involves "walking" the time-space matrix for the detectors. That is, the travel time employed for the n^{th} sub-segment is the value in the time-space matrix that pertains at the time that sub-segment is reached given that the initial start time at the initial sub-segment is at the beginning (or middle) of the initial 5-minute time period.

A slightly more sophisticated approach was developed by Hu (2011). It combines the travel *rates* for the applicable time periods (based on the time-space matrix) in a way that ensures the best possible contribution to the travel rate *between* adjacent sensors is obtained. The intent is to capture the effects of variations in congestion levels between the segments (e.g., due to merges, diverges, and lane drops).

The approach computes the arithmetic average of the two spot rates is computed and then adjusts that result by a factor γ:

$$\tau_s = \gamma\left[\frac{\tau_1 + \tau_2}{2}\right] \qquad \text{Equation 6-1}$$

In Equation 6-1, τ_s is the travel rate for the segment and τ_1 and τ_2 are the travel rates for the upstream and downstream detectors respectively. The value of γ is dependent on the traffic flow conditions on the segment, i.e., the regime that is extant at the time for which τ_s is desired (e.g., the level of congestion present). The appropriate value of γ can be obtained from a look-up table once the values have been calibrated for the regimes.

An alternative equation uses two parameters α and β to combine the spot rates:

$$\tau_s = \alpha\tau_1 + \beta\tau_2 \qquad \text{Equation 6-2}$$

Again, the values of α and β are dependent on the traffic flow conditions on the segment, i.e., the regime that is extant at the time for which τ_s is desired (e.g., the level of congestion present). They can be obtained from a look-up table once the values have been calibrated for the regimes.

Influencing Factor Analysis

A major purpose of the TTRMS is to empower agencies to improve the reliability of their systems. The objective is to guide agencies toward actions that can be taken to improve reliability. For example, if the agency's facilities are experiencing unreliable travel times largely due to incidents, the agency might choose to increase spending on incident management systems or on roadway safety improvements (see also Tsubota *et al.* 2011). This analysis can also help agency administrators set benchmark goals against which they can test future improvements.

The process for conducting these analyses includes the following steps:

1. Select the region or facilities of interest
2. Select a timeframe of interest
3. Assemble travel rate data for each facility
4. Generate TR-PDFs for each facility
5. Understand variations in reliability due to congestion
6. Develop TR-CDFs for each combination of recurring congestion level and non-recurring event(s)

The aim is to create separate TR-PDFs for each combination of 1) type of non-recurring event, including “normal” (i.e., no non-recurring event) and 2) recurring congestion level (i.e., low, moderate, high).

The technique for doing this involves two sub-steps:

- Identify types of non-recurring events in the data
- Identify the reliability impacts of congestion

The first sub-step is to identify types of non-recurring events in the data. The data for each route are plotted against time of day and VMT/hour to identify outliers. Starting with the most extreme (largest) outliers first, web-based databases should be queried to see if an explanatory non-recurring event can be identified for the date and time when the unusual travel rate occurred. For an operational TTRMS, this process should be automated and be conducted in real-time because event information tends to be perishable data. Categories of non-recurring events may include incident, weather, special event, and demand. Data points not falling into any one of these categories should be classified as being Normal.

When identifying categories of non-recurring events, Demand should always be the last category considered, after explanations related to Weather, Special Events, or Incidents are identified. Moreover, the latter three categories always trump the Demand designation. Values in the Demand category are extracted from those remaining in the Normal category after those explained by Weather, Special Events, or Incidents have been removed. This removal process should be iterative; there is nothing permanent about the Demand designation, unlike the other three categories.

When identifying data points in the Demand category, the VMT/hour value for a given 5-minute observation should be compared against the average for that 5-minute time period. If the value is more than two standard deviations above the mean, the data point should be given a Demand designation. A second analysis should also be conducted because this technique does not work during the highly congested time periods when VMT/hour is constrained by capacity (because the VMT/hour cannot be higher).

The second analysis seeks sequences of 5-minute time periods when the VMT/hour is high *and* the travel rate is high. This analysis identifies conditions when the demand-to-capacity ratio is higher than the volume-to-capacity ratio, implying there are queues in the system.

The second sub-step (in the fourth step) involves labeling each observation based upon the nominal loading of the system expected for each observation. This is done by analyzing the observations that remain once the non-recurring events have been removed.

The purpose of the congestion level designations is to differentiate the observations based on the reliability performance to be expected based on system loading, such as congestion. Many metrics could be used to assess this impact, such as the buffer time index, the planning time index, or the travel time index. However, the authors of this report used the semi-variance measure because the semi-variance is sensitive to how the data are distributed above the minimum value. As explained in the methodological description, the semi-variance σ_r^2 is a one-sided variance metric that uses a reference value r instead of the mean as the basis for the calculation and only observations x_i that are greater than (or less than) that reference value are used:

$$\sigma_r^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - r\right)^2 \quad and \quad \sigma_r = \sqrt{\sigma_r^2} \quad \exists\, x_i \geq r$$

Based on this analysis, the Normal data can be broken down into different recurring congestion-related categories.

The fifth step involves looking at the semi-variance trends so that the variations in reliability due to congestion can be understood. Low semi-variance values indicate high reliability on a route. The comparison of semi-variance values throughout the day can be used to identify peak time periods and how reliability changes throughout the day.

The sixth step involves developing TR-CDFs for each combination of recurring congestion that would normally occur (from the analysis above) and non-recurring event (from the first categorical analysis). These combinations are the “regimes” in which the facility operates per the definition of that terminology presented earlier. The TR-CDFs are created by appropriately binning the 5-minute travel time observations.

An example application of this procedure is contained in the use cases document. The system that was studied comprises three freeway routes from A to B in San Diego, as shown in Figure 6-13: I-5, I-805/ CA-15/I-5, and I-805/CA-163/I-5.

Source: https://maps.google.com/ accessed 9/7/2012

Figure 6-13: Three Routes Examined in Use Case AE-1

In subsequent text, these three routes are identified more succinctly as I-5, CA-15, and CA-163. The timeframe of interest was 2011, all weekdays, and all 24 hours during those days. The data were average travel rates from A to B for each route based on system detector data obtained by walking the time-space matrix for hypothetical trips that start every five minutes during the day on all three routes.

The travel rates are displayed in Figure 6-14 plotted against time of day and Figure 6-15 plotted against VMT/hour. Since the data for the entire year are shown, there are 72,000 values for each route. Hence, the total number of data points in the combined graphs is 216,000. Travel rates are needed because normalizing by the distance makes it possible to compare the performance of one route with the others without having the differences in length confound the analysis.

Step 4 involves labeling each observation—all 216,000 in this case—in terms of the regime that was operative for each observation. Since regime labels were added ex-post-facto, the process involved three sub-steps. The first was to add a non-recurring event designation. The data for each route was plotted against time of day and VMT/hour (system loading) as shown in Figures 6-14 and 6-15 respectively. Hourly VMT data (effectively VMT/hour) were obtained from PeMS. The actual hourly values were assigned to the 6th five-minute observation in each hour (25 minutes) and the other 5-minute values were generated by interpolating between these values. Starting with the most extreme (largest) outliers first, web-based databases are queried to see if explanatory non-recurring events can be identified for the dates/ times when the unusual travel rates occurred. For this particular system, the types of non-recurring events were: Incidents, Weather, Special Events, and Demand. An Incident was an accident or some other disruptive traffic event – recorded in the PeMS database or some other source; Weather was an inclement weather event; Special Event was an unusual event – often sports-related; and Demand was a condition when the VMT (implicitly, the traffic flows) was higher than normal for the time-of-day at which the high travel rate arose. Data points not falling into any one of these categories remained in a “Normal” category. (A weakness of this approach is that non-recurring events that do not create outliers might be missed.)

The “Demand” designation was always the last one added. That is, explanations were sought related to Weather, Special Events, or Incidents before using “Demand” as the explanation. Moreover the former three categories always superseded the “Demand” designation. Hence, values in the “Demand” category were extracted from those remaining in the “Normal” category after those explained by Weather, Special Events, Incidents, or other non-recurring events (e.g., work zones) are removed. Moreover, this removal process was *iterative*; there was nothing permanent about the “Demand” designation, unlike the other three. Second, the identification of the “Demand” category data points had two facets. The first involved comparing the VMT/hr value for a given 5-minute observation with the average for that 5-minute time period. If the value was more than two standard deviations above the mean, it was given a “Demand” designation. Then, because this technique did not work during the highly congested time periods when VMT/hr was constrained by capacity – because the VMT/hr cannot be higher – a second analysis was conducted.

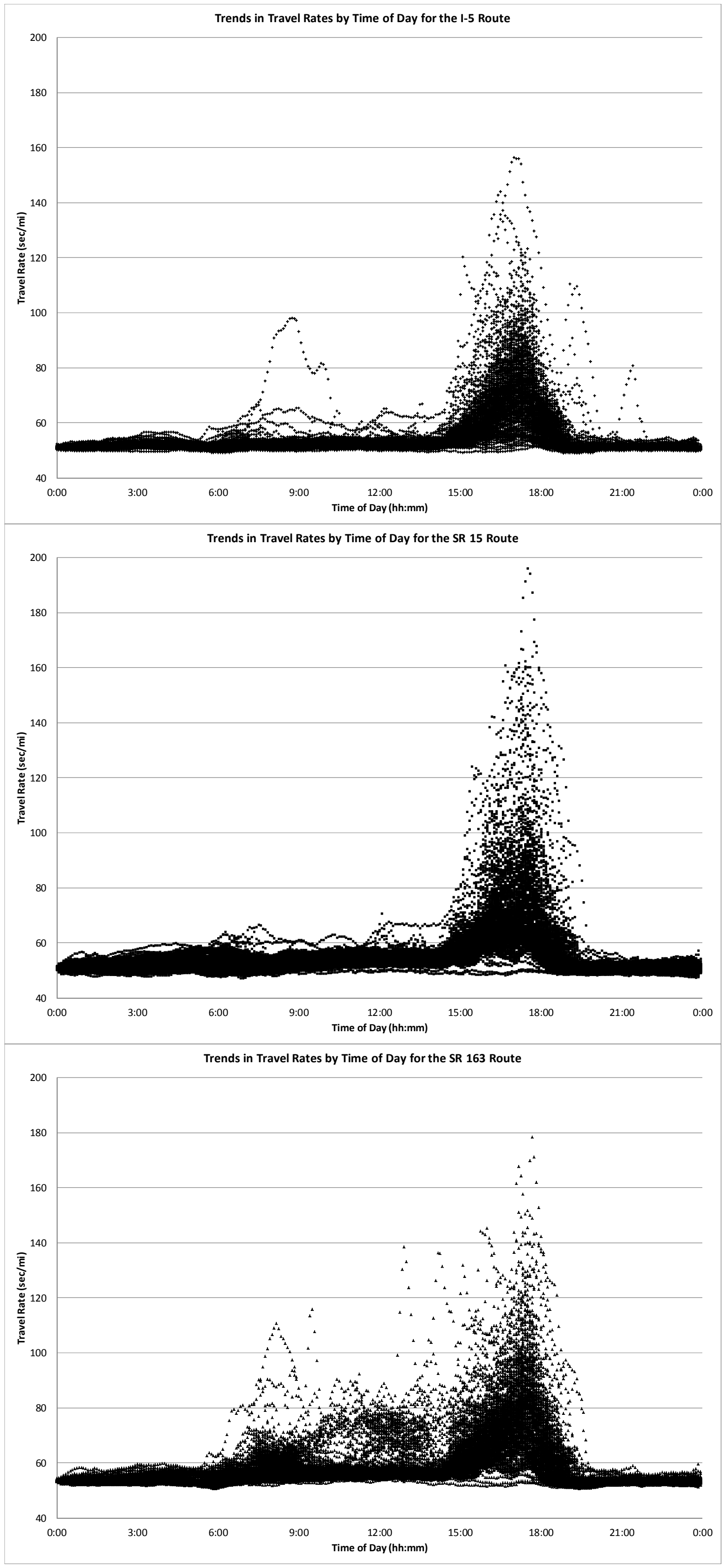


Figure 6-14: Five-Minute Average Weekday Travel Rates for Three Routes in San Diego

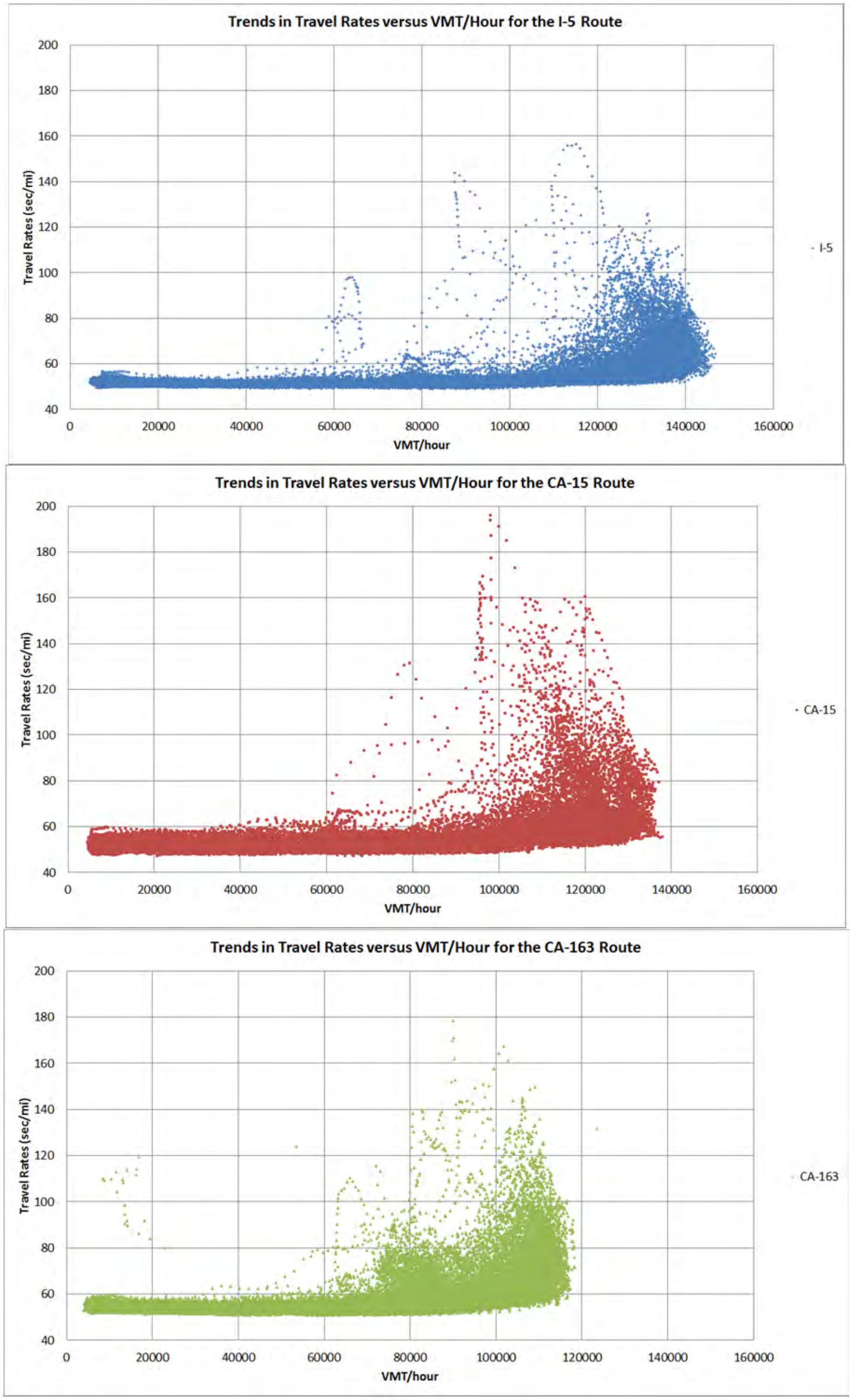


Figure 6-15: Five-Minute Average Weekday Travel Rates plotted against VMT/Hour for Three Routes in San Diego

Sequences of 5-minute time periods should be sought when: the VMT/hr was high and the travel rate was high. (Effectively these were conditions when the D/C ratio was higher than the V/C ratio; implying there were standing queues in the system.) In this particular instance the values used were 75,000 VMT/hr, 80 sec/mile, and 30 minutes. That means that 5-minute time periods were labeled as being in the "Demand" category if their VMT/hr exceeded 70,000 VMT/hr, their travel rate was greater than 80 sec/mile, and at least the next five 5-minute time periods (30 minutes total) were in the same condition.

Note that changing these criteria affects the selection process. Basically, it changes the separation between observations that are considered normal, high congestion and those that are attributed to high demand on top of high congestion. The values were chosen because: 70,000 VMT/hr, especially for the CA-163 route, was the point at which there was a step change in the variability of the travel rates; 80 sec/mile is the same as 45 mph, which is often the speed that arises when freeways are operating at capacity; and 30 minutes was deemed to be a reasonable system recovery time. It is effectively how long one assumes it takes the system to recover from normal high demand and return to a status where the travel rate is less than 80 sec/mile. Higher values imply that it is acceptable for the system to take longer; shorter values assume it should take less time. Setting it at 0, for example, would imply that the system should be able to recover from travel rates above 80 sec/mile in five minutes.

The second sub-step involved labeling each observation based upon the nominal loading of the system expected for each observation. This was done by analyzing the observations that remain once the non-recurring events have been removed. The semi-variance measure was employed. In this instance, semi-variance values were computed for every five minute interval for each of the three routes. Figure 6-16 presents the result. The value of *r* employed for each route was the minimum travel rate observed for the entire year. Moreover, because the number of observations varied from one five-minute period to another, the semi-variances were divided by the number of observations by *n* as shown in the equation above (effectively creating an average per observation so that the results would be comparable among the five-minute time periods.

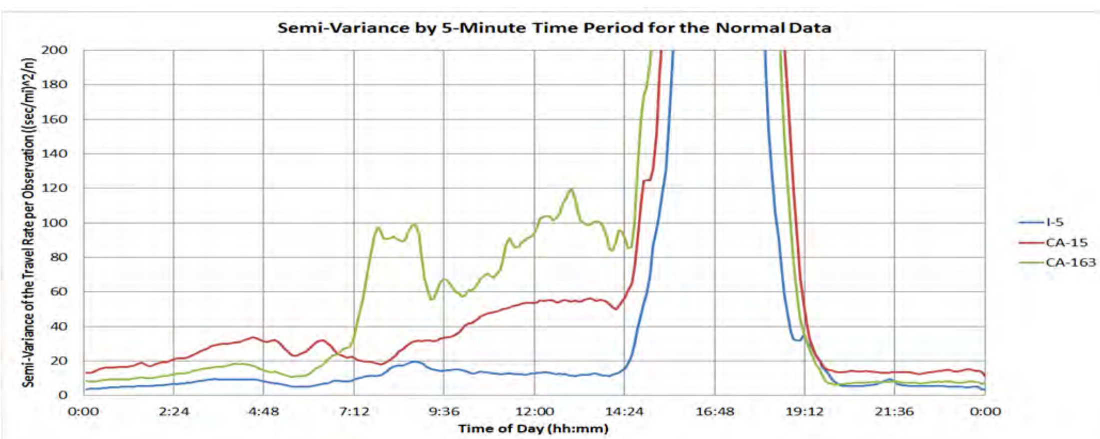


Figure 6-16: Semi-Variance Values for Every Five-Minutes / Weekday Average Travel Rates for the Normal Condition for Three Routes in San Diego

Notice that reliability becomes worse as the traffic levels increase. This should be expected: reliability should be best when the traffic volumes are low – like late at night or early in the morning. It should be poorer during when the traffic volumes are higher – when the vehicles interact more – like during the midday, and it should be poorest when the traffic volumes are the highest – as in the PM peak – when the varying lengths of the standing queues has an impact. The maximum semi-variance values – not shown – reach about 1,000.

While no "right" answer exists for the number of categories to use, four were selected here: Uncongested, Low, Moderate, and High. Uncongested meant the semi-variance was below 20; Low meant 20 to 40; Moderate, 40 to 120; and High, above 120. Thus, the I-5 route was classified as Uncongested all day except from 2:15 p.m. to 6:50 p.m. when it was classified as High. The CA-15 route was classified as Uncongested from midnight to 2:10 a.m.; Low from 2:15 a.m. to 6:45 a.m.; Uncongested from 6:50 a.m. to 8:15 a.m.; Low from 8:20 a.m. to 9:05 a.m.; Moderate from 9:10 a.m. to 2:10 p.m., High from 2:15 p.m. to 7:20 p.m., and Uncongested from 7:25 p.m. to midnight. The CA-163 route was classified as Uncongested from midnight to 6:45 a.m., Moderate from 6:50 a.m. to 2:15 p.m., High from 2:20 p.m. to 7:20 p.m., and Uncongested from 7:25 p.m. to midnight.

Step 5 involved developing TR-CDFs for each regime; that is, each combination of nominal loading (from the analysis above) and non-recurring event (from the first categorical analysis), including "none". The TR-CDFs are created by appropriately binning the 5-minute travel time observations. Figure 6-17 presents the results.

Step 6 involved interpreting the results in terms of the effects on reliability of the various factors. But since that overlaps with the next use case, the results are presented there.

The objective in this use case is to determine how various factors affect system reliability. Such information helps inform decisions about how to improve performance: geometric treatments, capacity enhancements, operational changes, better signage, improved roadway striping, resurfacing, or better lighting. It can also help managers determine which facilities need better real-time traveler information (such as Changeable Message Signs displaying alternate routes and travel times).

Figure 6-16 showed that the three routes have somewhat different daily patterns of reliability. The I-5 route has high reliability (a low semi-variance value) throughout the day except during the PM peak. In contrast, the CA-15 route has an increase in its semi-variance (a drop in reliability) across the midday (a higher semi-variance). The CA-163 route has an even more dramatic increase in its semi-variance across the midday but a lower semi-variance during the early morning hours. In addition, the CA-163 route has a discernible AM peak while the other two routes do not.

From an interpretation standpoint, this means the I-5 route is probably the most reliable. It is still challenged during the peak, but consistently has the lowest semi-variance values except for a few 5-minute periods around 7-9pm. Interestingly, this means that even though Exhibit 3-3 suggests the CA-15 route may have the lowest average travel rates most of the day, the most reliable route is a different one, namely I-5.

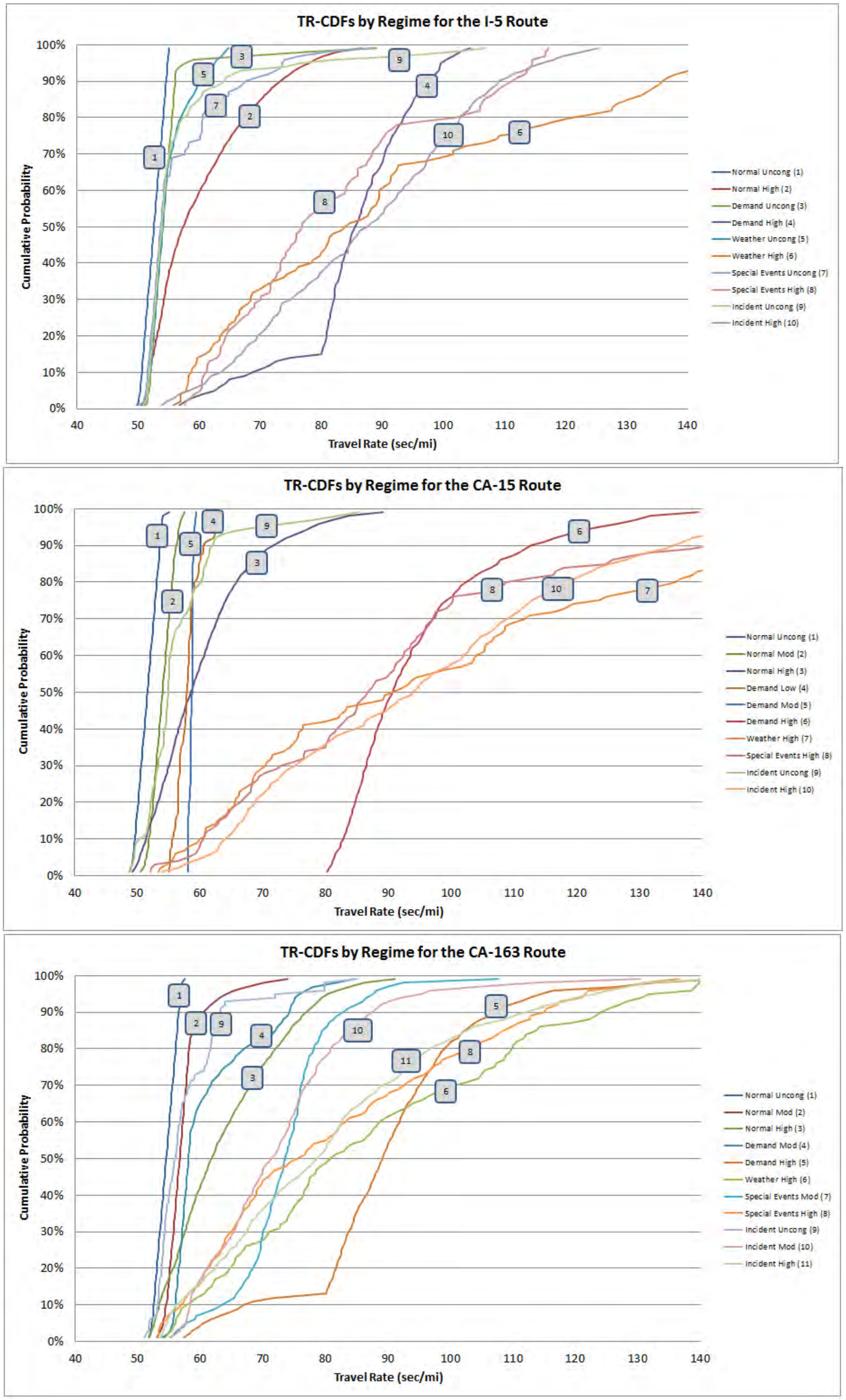


Figure 6-17: CDFs by Regime for the Three Routes in San Diego

Figure 6-16 also suggests that CA-163 is the least reliable route. It has the highest semi-variance during the day – except in the early morning when the CA-15 route has higher values – and its semi-variance is significantly higher – especially during the morning and midday time periods.
Figure 6-17 provides additional insights. While the plots are rather dense, they do tell a story about the performance of these three routes. Looking at I-5 first, its TR-CDF for the uncongested/normal condition is at the far left and it is almost vertical. This means it is very reliable travel rates during this condition. Moreover, during these uncongested conditions, even the non-recurring events affect only the top 30% of the 5-minute periods and in the worst case double the travel rate at the 100th percentile from about 50 sec/mile up to 100 seconds per mile – the fourth from the left and the most jagged of the group – related to incidents.

I-5's performance during the congested conditions is quite different. In Figure 6-17, even when there are no identifiable non-recurring events, larger travel rates are involved as can be seen by the smooth red-colored CDF having travel rates from about 50 to 100. Moreover, when non-recurring events occur during high congestion, the impacts are "severe": the travel rates are substantially higher than for normal, high-congestion conditions. The TR-CDFs for three of these conditions largely overlap – for incidents, special events, and weather –and no one CDF dominates the other. However, the TR-CDF for the Demand condition (under high congestion) is strikingly different. It has much larger travel rates even at low percentiles, a kink at about 82 sec/mile – when the Demand events during high-condition begin to have an impact on the CDF – and a maximum value that is substantially smaller than that for the other three non-recurring categories. The implication is that Demand needs to be a cause for concern, and reducing the rates for low percentile values may be possible – through geometric improvements – reducing the tail may not be that important – it may be more important to focus on the tail for the three other conditions – that involve much higher travel rates – even above the 50th or so percentile.

The story for the CA-15 route is similar. Almost all of the regimes involving no or low congestion have similar TR-CDFs. There is some spread between 50-60 sec/mi, but the TR-CDFs are all nearly vertical – not much variation in the travel rate occurs. The one notable exception is the TR-CDF for uncongested conditions when incidents arise. As with the I-5 TR-CDFs, the incidents produce a major shift for the travel rates at the higher percentile values – in this case above about 90%. The TR-CDF for high congestion during Normal conditions is the very smooth curve on the right-hand edge of the large cluster. Like I-5 it involves a much larger range of travel rates, from 50-85, and more change in the travel rate as the percentiles increase.

The four TR-CDFs that are strikingly different are those for incidents, special events, weather, and demand during periods that would normally involve high congestion. This is not surprising, but it does reinforce the importance of taking actions that help manage the severity of these events when they occur during congested operation. (In this case, for the Demand conditions there is a significant shift in the travel rates from 50 to 80 sec/mile even at the 0th percentile.)

The story for the CA-163 route is quite different. It obviously has problems. Its TR-CDFs are widely scattered, and non-recurring events have an impact under all levels of congestion. The most important details to notice are that: 1) the most significant impacts (the CDFs furthest to the right) - all during high congestion - come from (right to left) weather, special events, and incidents; 2) the next two (light blue and dark red) are for weather under moderate congestion and demand during high congestion; and 3) the next three (right to left) are incidents, special events, and demand under *low* congestion conditions – not moderate.

With these differences noted, this route's reliability performance is otherwise similar to the other two. More specifically, it has a travel rate performance very similar to the other two routes under uncongested-normal conditions, but it struggles to maintain that performance either when the congestion levels get higher or non-recurring events take place.

The fact that the CA-163 route has more significant shifts in the TR-PDFs for various conditions leads to a conclusion that there are problems with this route between I-805 and I-5. It is not too difficult to see why by "driving" the route – either physically or virtually - and observing its physical features and congestion. The highway has many curves, its geometry is tight, and there are closely spaced interchanges. Particularly, between I-8 and I-5, it has tight geometry – it is an old facility – and only has two-lanes wide in each direction. While it is not the purpose of L-02 to determine what geometric and other treatments that would help alleviate reliability problems – that is the focus of other SHRP 2 projects like L-07 – it is obvious that this section of CA-163 is one where geometric improvements and expedient response to incidents would be likely to have a significant impact on reliability.

Step 6 involves rank ordering the facilities based on the relative impacts so that those most affected can receive mitigating treatments. Table 6-5 provides a way to develop the rankings. Columns 3-12 report the average semi-variance values (SV) for each regime as well as the frequency (n) with which that regime occurs. The 13^{th} column shows the semi-variance totals for each congestion condition (e.g., 573,000 for I-5 during uncongested conditions and 4,705,000 during congested conditions). These are based on the sum-product of the SV and *n* values. The last column in the top table reports the total semi-variance in the travel rate for the year (Facility Total).

Inspection of the facility totals suggests that the least reliable facility is CA-163. This is consistent with the impression one gains from the scatterplots shown in Figures 7-15 and 7-16. The CA-15 route is the next most unreliable (9465 versus 9561), but its distribution of the semi-variance is slightly different. As the bottom table shows, a higher percentage can be attributed to incidents and special events during nominally high congestion conditions.

A summary of this analysis is that all three of the routes exhibit variations in reliability depending upon the recurring congestion condition and non-recurring event. Evidence of these differences is most significant for the CA-163 route, and it seems apparent the "problems" it has are due to the geometric conditions on the section of CA-163 from I-805 to I-5. All three routes are significantly affected by high congestion—even under normal conditions—the TR-CDF for that condition is dramatically different from the CDFs for normal operation under lesser congestion conditions. And incidents, weather, special events, and higher-than-normal demand all have a significant effect on reliability during highly congested conditions. Finally, it is clear that these TR-CDFs provide guidance about actions that might be useful to help fix the reliability problems.

Table 6-5: Semi-Variances for Each Regime for Three Routes in San Diego

Route	Cond	Normal		Demand		Weather		Special Events		Incidents		Σ(SV*n) (000)	Facility Total
		SV	n	SV	n	SV	n	SV	n	SV	n		
I-5	Uncong	7	55533	60	1250	46	797	111	135	172	285	573	5278
	High	205	12783	1415	472	2563	175	1399	104	1769	466	4705	
CA-15	Uncong	15	24491	47	147	68	229	29	77	139	55	400	9465
	Low	27	15931	118	102	106	193	0	0	97	25	457	
	Mod	46	14863	127	13	151	271	0	0	93	103	740	
	High	241	13918	2415	665	3751	162	3113	168	3032	587	7868	
CA - 163	Uncong	11	32823	13	1019	61	277	21	29	54	102	386	9561
	Mod	56	20950	169	519	399	333	601	344	684	354	1841	
	High	261	12764	1789	1028	1924	254	1424	243	1385	961	7333	

Route	Cond	Normal	Demand	Weather	Special Events	Incidents	Σ(SV*n) (000)	Facility Total
I-5	Uncong	8%	1%	1%	0%	1%	11%	1
	High	50%	13%	8%	3%	16%	89%	
CA-15	Uncong	4%	0%	0%	0%	0%	4%	1
	Low	4%	0%	0%	0%	0%	5%	
	Mod	7%	0%	0%	0%	0%	8%	
	High	35%	17%	6%	6%	19%	83%	
CA - 163	Uncong	4%	0%	0%	0%	0%	4%	1
	Mod	12%	1%	1%	2%	3%	19%	
	High	35%	19%	5%	4%	14%	77%	

Considerations for Transit

Most of the discussion in this supplement has focused on vehicle (effectively auto) travel times. The exhibits are dominated by auto travel; the discussions about travel time and travel rates predominantly focus on automobile trips; and the commentary about diagnostic ideas relate to automobile trips.

Transit and freight trips are different. Transit passengers do not control what the vehicles do. They board and alight from the vehicles and make transfers. Their travel times are strongly influenced by the headways at which the vehicles operate and the reliability of the transfers.

Freight trips are similar. Packages get picked up and carried from shipper to terminal, terminal to terminal, and terminal to receiver. The travel times they experience are heavily influenced by the operating plans being followed by the freight providers and the reliability of their operations. Packages are very similar to transit passengers in that they ride on one vehicle after another and their travel time is influenced by the headway between pick-ups (not often thought about that way, but often once a day) and the reliability of the connections between vehicles (i.e., trucks). Unlike transit passengers, the packages cannot influence the reliability of their trips. If they get placed on the dock in the loading area for the wrong truck they cannot move themselves to the area for the right truck. Hence, the reliability of their trip times is likely to be worse than that of the transit riders. However, freight companies only earn revenues if they deliver packages on time, so they tend to pay attention to whether the packages are being handled correctly. On the other hand, transit agencies are not particularly sensitive to whether the passengers route themselves correctly - if a transit passenger gets delayed or reaches the wrong destination - culpability rests with the passenger as well as the service provider.

This having been said, a strong similarity exists between transit trips and package trips. They are both dependent on the headway between vehicle arrivals and the reliability of connections.

The observability of the trips is a different issue. Transit trips are largely unobservable. Many transit agencies do not track the movements of their passengers. Even the transit agencies with the most sophisticated data - such as the Washington Area Metropolitan Transit Authority - only know where and when the passengers entered and left the system - akin to AVI-type information. They do not know the path followed unless they were to track Bluetooth devices or cell phones - which they could do.

Packages, on the other hand, tend to be tracked carefully by many freight service providers. The public agencies may not have access to this information, but many carriers know where the packages are at all times. In some instances it is because the package's bar code was just read (i.e., it was picked up or received at a distribution center) or sometimes it is by inference (it was scanned as it was loaded onto the delivery truck and the delivery truck is en-route to the receiver). In this sense, the package data is akin to AVI-type data. Undoubtedly, in selected instances, the packages have RFID tags that are being read constantly, so the data are AVL-like, but in most instances, this is not the case.

Since the carriers rarely share their package-level information except with the stakeholders that have a need-to-know (the shipper and receiver), providing reliable service to freight carriers becomes functionally similar to dealing with reliable travel times for the autos. The trucks need to be able to traverse the highway network with reliable travel times. They do not want to be delayed so their deliveries are late. Unlike person trips, though, they often also do not want to be early because they will then have to wait until they were supposed to arrive - another activity could have been inserted - which represents a lost opportunity for better efficiency, more cost-effective operation, or more revenue.

Hence, this discussion now focuses on the transit trips because they are more often under the purview of the agencies responsible for operating the highway system.

During the case studies, transit data were only obtained during the San Diego case study. However, those data are very representative of the information available to the most progressive transit operators. Selected vehicles were equipped with AVL-like devices that could monitor the Lat/Lon location of the bus in real time, the times at which the bus doors opened and closed, and the number of people who boarded or alighted from the bus.

Were all the buses instrumented, then a technique akin to that used to generate the freeway travel times could have been used. It could have been assumed that hypothetical passengers boarded a bus B1 at time T1 at stop S1 bound for stop S2. By simulating a large number of trips from S1 to S2 during different times of day (operating conditions), PDFs of the transit travel times could have been created. For trips on a single line this would have been simple. For trips that involve transfers, the process would have been slightly more complicated. The hypothetical passenger would have boarded bus B1 at time T1 and stop S1, traveled to transfer location X1, alighted at T2 and waited for a bus B2 that arrived at X1 at some time T3 > T2. The traveler would then board bus B2 travel to S2 and alight at some time T4. The difference T4 - T1 would be the travel time and the reliability of these trips could also be assessed.

In the case of San Diego where not all of the buses were instrumented, a more complex analysis procedure had to be employed. The process involved two steps: 1) pre-processing the bus trip data to develop information needed to conduct the analysis and 2) generating a synthesized set of hypothetical, representative trips through Monte Carlo simulation. (For other techniques see Bertini and El-Geneidy 2003, Bertini and El-Geneidy 2004, Yang *et al.* 2010.)

Developing Transit Rider PDFs for Trips

Figure 6-18 shows the process used to synthesize the trip times. The flow chart at the top of the figure provides an overview. The bottom flow chart provides more detail. The whole figure is annotated with letters from A to J to provide reference markers for the description that follows. It is also couched in the context of a trip on bus routes 11 and 7, but the bus route numbers are not relevant to this discussion. It is sufficient to recognize that two separate bus routes are involved with a transfer between them.

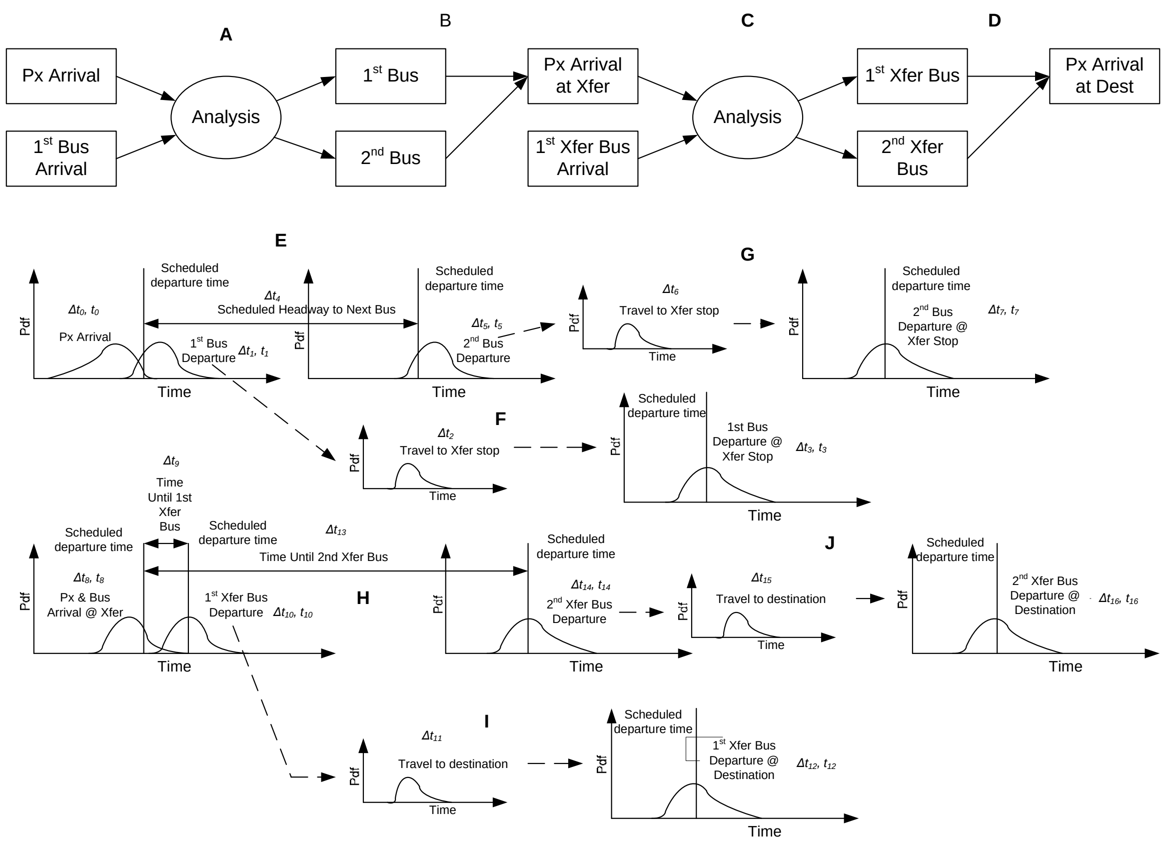


Figure 6-18:Analysis Flow Chart for Transit Trips involving Transfers

The overview starts with Marker "A", focused on the initial bus boarding process. The passenger arrives; as does a bus on Route 11. Depending on when they arrive, the passenger either gets on the 1st Route 11 bus or the next (2nd) one. If he/she gets on the 2nd one, a delay of one-headway is incurred. (Later text will describe what this means in more detail.) In either event, as shown by the blocks near marker "B", the passenger travels to and arrives at the transfer point, as shown near Marker "C". Arriving separately is the 1st Route 7 bus. An analysis of when that bus arrives relative to when the passenger arrives on the Route 11 bus determines whether the passenger gets on the 1st Route 7 bus or has to wait for the next (2nd) one. If he/she gets on the second Route 7 bus, an additional delay is incurred. (Later text will describe this in more detail.) In either event, as shown by the blocks near marker "D", the passenger then arrives at the destination.

The detailed description starts with Marker "E". Near it are shown the PDFs for the arrival of the passenger (Px) and the 1st Route 11 bus. Consistent with Bowman and Turnquist (1981), the passenger PDF (Δt_0) tends to favor early arrivals with a small probability of being late. Separately, consistent with the San Diego data, the Route 11 bus (Δt_1) follows a second PDF. The distribution for the bus indicates a small probability of departing early (earlier than the scheduled departure time) and a much larger probability of departing late. If the passenger arrives before the Route 11 bus departs, then the passenger boards the 1st Route 11 bus. If that happens, the descending dashed line toward marker "F" indicates that the passenger incurs a travel time (Δt_2) to reach the transfer stop and the passenger (on the Route 11 bus) arrives at the transfer stop at t_1, which is at some point in time relative to the scheduled departure time (Δt_3). (Departure times have been used as the reference because they are "worst case" times - we know for sure that the passenger has arrived when the bus departs.) If the passenger misses the 1st Route 11 bus, because he/she arrives after the 1st Route 11 bus departs, then a schedule delay (Δt_4) is incurred until the next Route 11 bus arrives (to the right of marker "E"). A 2nd Route 11 bus arrives (Δt_5), the passenger boards, and Route 11 bus travels to the transfer location (Δt_6), shown by marker "G", and the passenger arrives at the transfer stop at t_2, which is at some time relative to its scheduled departure (Δt_7).

Whichever arrival time governs (t_1 or t_2) becomes the start of the second part of the trip (Marker "H"). Moreover, the corresponding relative arrival time (Δt_4or Δt_7) becomes the basis (Δt_8) for determining which transfer bus is caught. If the passenger's relative arrival time on the Route 11 bus (Δt_8) is less than the sum of the scheduled connection time (Δt_9) and the relative departure time for the Route 7 bus (Δt_{10}), then the 1st Route 7 bus is caught. This leads to a travel time to the destination (Δt_{11}), an arrival time (t_3) and a relative arrival time compared to the schedule (Δt_{12}) (Marker "I"). On the other hand, if the Route 11 bus arrives late (Δt_8) or the Route 7 bus departs early (Δt_9+ Δt_{10}), then the passenger may miss the 1st Route 7 bus, incur a delay (Δt_{13}), until the next Route 7 bus arrives (Δt_{14}), then incur a travel time (Δt_{15}) to the destination and arrive at t4 with a relative arrival time Δt_{16} (Marker "J").

A couple numerical examples help illustrate the analysis. Table 6-6 presents four of them. In the first, no bus is missed. In the second, the connection bus is missed. In the third, the first Route 11 bus is missed but the subsequent connection is made. In the fourth, both the first Route 11 bus is missed and the first Route 7 transfer bus is missed. In all cases the reference time when t = 0 is the scheduled departure time of the first Route 11 bus. All the values are in seconds. Results obtained from actually working with the transit data obtained in the San Diego case study can be found in that supplement.

The first example starts with $\Delta t_0 < \Delta t_1$ (-120 < 30), which means the passenger gets to catch the first Route 11 bus. The starting time for the trip (t0) becomes -120 seconds (i.e., the passenger arrived 2 minutes before the scheduled departure time, which is the reference point for $t = 0$). The travel time to the transfer point is $\Delta t_2 = 1570$, the arrival time is $t_3 = t_8 = 1600$, and the relative arrival time at the transfer point (relative to the scheduled departure at that location) is $\Delta t_3 = \Delta t_8 = 20$.

The next thing to do is to analyze the connection. The relative arrival time is $\Delta t_8 = 20$, the transfer time is $\Delta t_9 = 240$ and the 1st Route 7 bus is late $\Delta t_{10} = 50$, so the passenger has no problem catching the first transfer bus $\Delta t_8 < \Delta t_9 + \Delta t_{10}$). The passenger then departs the transfer stop at $t_{10} = t_8 - \Delta t_8 + \Delta t_9 + \Delta t_{10} = 1600 - 20 + 240 + 50 = 1870$, travels to the destination $\Delta t_{11} = 190$, arrives at the destination at $t_{12} = t_{10} + \Delta t_{11} = 1870 + 190 = 2060$, with an arrival relative to the scheduled arrival time of $\Delta t_{12} = -10$ (10 seconds early) and an overall travel time of $tt = t_{12} - t_0 = 2060 - (-120) = 2180$ seconds (36.3 minutes).

Table 6-6: Four Numerical Examples of Estimating Travel Times for Transit Trips involving a Transfer

Metric	No Miss	Miss 2	Miss 1	Miss Both
Δt_0	-120	-90	-30	50
Δt_1	30	15	-50	-100
Δt_2	1570	1730	-	-
Δt_3	20	350	-	-
Δt_4	-	-	900	900
Δt_5	-	-	-30	40
Δt_6	-	-	1400	1800
Δt_7	-	-	-100	400
Δt_8	20	350	-100	400
Δt_9	240	240	240	240
Δt_{10}	50	-100	70	-100
Δt_{11}	190	-	210	-
Δt_{12}	-10	-	20	-
Δt_{13}	-	720	-	720
Δt_{14}	-	10	-	-30
Δt_{15}	-	180	-	190
Δt_{16}	-	-10	-	30
t_0	-120	-90	-30	50
t_1	30	15	-	-
t_3	1600	1745	-	-
t_5	-	-	870	940
t_7	-	-	2270	2740
t_8	1600	1745	2270	2740
t_{10}	1870	-	2680	-
t_{12}	2060	-	2890	-
t_{14}	-	2365	-	3270
t_{16}	-	2545	-	3460
tt	2180	2635	2920	3410

In the second example, the first Route 11 bus is caught, but the first Route 7 transfer bus is missed. The example starts with $\Delta t_0 \leq \Delta t_1$ (-90 ≤ 15), which means the passenger catches the first

Route 11 bus. The starting time for the trip (t_0) becomes -90. The travel time to the transfer point is Δt_2 = 1730, the arrival time is $t_3 = t_8$ = 1745, and the relative arrival time at the transfer point (relative to the scheduled departure time) is $\Delta t_3 = \Delta t_8$ = 350. The transfer time is Δt_9 = 240 and the 1st Route 7 bus leaves early Δt_{10} = -100, so the passenger misses the first transfer bus ($\Delta t_8 \geq \Delta t_9 + \Delta t_{10}$, or 350 ≥ 240 + (-100)). Hence, the passenger has to wait for the second transfer bus which has a scheduled time Δt_{13} = 720 which is 12 minutes later than the 1st transfer bus, and it arrives a little late Δt_{14} = 10. This means it leaves at $t_{14} = t_8 - \Delta t_8 + \Delta t_9 + \Delta t_{13} + \Delta t_{14}$ = 1745 – 350 + 240 + 720 + 10 = 2365. The Route 7 bus then travels to the destination Δt_{15} = 180 and arrives a little early Δt_{16} = -10 at t_{16} = 2545. The overall trip time is $tt = t_{16} – t_0$ = 2645 (43.9 minutes).

In the third example, the first Route 11 bus is missed and the first Route 7 transfer bus is caught. The example starts with $\Delta t_0 > \Delta t_1$ (-30 > -50), so the passenger misses the first Route 11 bus. (The starting time for the trip (t_0) becomes -30.) The passenger has to wait for the next bus Δt_4 = 900 which is a little early Δt_5 = -30. The travel time to the transfer point is Δt_6 = 1400, the arrival time is $t_7 = t_8$ = 2270, and the arrival time at the transfer point relative to the scheduled departure time is $\Delta t_7 = \Delta t_8$ = -100. The transfer time is Δt_9 = 240 and the 1st Route 7 bus leaves late Δt_{10} = 70, so the passenger catches the first transfer bus ($\Delta t_8 \leq \Delta t_9 + \Delta t_{10}$, or -100 ≤ 240 + 70). The passenger departs the transfer stop at $t_{10} = t_8 - \Delta t_8 + \Delta t_9 + \Delta t_{10}$ = 2270 – (-100) + 240 + 70 = 2680, travels to the destination Δt_{11} = 210, arrives at the destination at $t_{12} = t_{10} + \Delta t_{11}$ = 2680 + 210 = 2890, with an arrival relative to the scheduled arrival time of Δt_{12} = 20 (20 seconds late) and an overall travel time of $tt = t_{12} – t_0$ = 2890 – (-30) = 2920 seconds (48.7 minutes).

In the fourth example, both the first Route 11 bus and the first Route 7 transfer bus are missed. The example starts with $\Delta t_0 > \Delta t_1$ (50 > -100), so the passenger misses the first Route 11 bus. (The starting time for the trip (t_0) becomes 50.) The passenger has to wait for the next bus Δt_4 = 900 which is a little late Δt_5 = 40. The travel time to the transfer point is Δt_6 = 1800, the arrival time is $t_7 = t_8$ = 2740, and the arrival time at the transfer point relative to the scheduled departure time is $\Delta t_7 = \Delta t_8$ = 400. The transfer time is Δt_9 = 240 and the 1st Route 7 bus leaves early Δt_{10} = -100, so the passenger misses this bus ($\Delta t_8 \geq \Delta t_9 + \Delta t_{10}$, or 400 ≤ 240 + (-100)) and has to catch the second one. The added wait for the next bus is Δt_{13} = 720 which is 12 minutes later than the 1st transfer bus, and that bus arrives a little early Δt_{14} = -30. This means the departure time from the transfer stop is $t_{14} = t_8 - \Delta t_8 + \Delta t_9 + \Delta t_{13} + \Delta t_{14}$ = 2740 – 400 + 240 + 720 + (-30) = 3270. The Route 7 bus then travels to the destination Δt_{15} = 190 and arrives a little late Δt_{16} = 30 at t_{16} = 3460. The overall trip time is $tt = t_{16} – t_0$ = 3460 – 50 = 3410 (56.8 minutes).

Summary

Data processing and analysis is essential in using the travel time reliability monitoring system. The ultimate objective is to prepare distributions of the travel times that can be displayed in histograms, PDFs and CDFs. This chapter described the processes whereby raw travel time information can be analyzed and summarized to create the travel time distributions.

An important observation is that no single processing strategy seems to work for all situations. While the methods all culled and summarize the raw data to create the distributions, the detailed manner in which this is done depends on the data sources available.

Another important observation is that the analyst needs to decide "what" should be analyzed. At one extreme, it may be the entire year. At the other, it may be a set of five-minute time slices during the morning or afternoon peak on weekdays in the winter. The methods will work in all instances.

It is obvious that any analysis requires an understanding the causal factors involved. System operators need to know and understand the impacts of congestion, incidents, weather, etc. Hence, deciding how to attribute these influences is a major element of the analysis. This may seem simple at first, but the L02 study team found it was fairly complex. This is because non-recurring events can have impacts on segments well beyond the one on which they occur, including upstream and downstream, in the opposing direction, and on intersecting facilities. For example, an incident on an intersecting freeway can cause back-ups through ramps onto other facilities. Hence, an understanding of the network is critical in determining what events affect what segments.

It is also critically important that influences not be confounded. For example, mixing data from different congestion levels and non-recurring events can confound the analyst's ability to see clear effects. If the impacts of the causal factors were separable and additive, this might not be a problem, but such is not the case. For example, the L02 team found that weather can have a dramatic impact during high congestion, but during times of low or no congestion, the impact is far less dramatic. The L02 team's use of regimes to bin the data was particularly valuable in parsing out the influence of various causal factors.

Finally, it will be very helpful in the future if monitoring system modules can capture data for non-recurring events as they occur, rather than ex post facto. While the L02 team demonstrated that ex post facto analyses can be done, explanatory information is only sought when it is obvious through outlier analysis that the travel times have been effected, so instances are missed when the non-recurring events took place and no travel time impact occurred.

CHAPTER 7: CASE STUDIES

This section describes the case studies and use cases employed to test the ideas being presented in the guidebook for the TTRMS. The case studies were performed in San Diego, California; Northern Virginia; Sacramento/Lake Tahoe, California; New York/New Jersey; and Atlanta, Georgia. Figure 7-1 shows the case study locations. The five main case studies are presented first, followed by additional applications in other locations.

Map data © 2012 Google

Figure 7-1. Case Study Locations

In each of the case studies, sensor data was collected in real-time from a variety of transportation networks, process these data inside a large data warehouse, and generate reports on travel time reliability to help agencies better operate and plan their transportation systems.

The TTRMS realizations used in the case studies were based on the existing Performance Measurement System (PeMS) monitoring system, a web-based software system for the state of California that collects traffic data from over 30,000 loop detectors every 30 seconds, filters and cleans the raw data, computes performance measures, and aggregates and archives them to enable detailed analysis. PeMS is a traffic data collection, processing, and analysis tool that extracts information from real-time intelligent transportation systems (ITS) data, saves it permanently in a data warehouse, and presents it in various forms to users via the web.

PeMS was linked with various existing monitoring systems in the case studies outside California. Because it can calculate many different performance measures, the requirements for linking PeMS with an existing system depend on the features being used. PeMS needs to acquire both the roadway network information and equipment configuration metadata before traffic data can be stored in the database. PeMS has a very strict equipment configuration framework which is described in the Travel Time Reliability Monitoring System Resource Document. Different methodologies were applied and specific use cases were demonstrated in each case study based on the existing data and monitoring systems.

The investigations presented in each case study are categorized as System Integration experiments, Integration of Sources of Non-Recurrent Congestion experiments, and Other Use Cases. Systems Integration experiments relate to activities that occur before development of a probability density function (PDF) for travel time reliability. Integration of Sources of Non-Recurrent Congestion experiments include both system integration aspects and use case demonstrations. Other Use Cases relate to the demonstration of specific use cases after a PDF has been created.

System Integration includes investigations into data integration considerations, comparison with probe data, and development of travel time reliability functions. The Northern Virginia, Sacramento/Lake Tahoe, Atlanta, and New York/New Jersey case studies include System Integration experiments.

Integration of Sources of Non-Recurrent Congestion experiments demonstrate specific use cases related to analyzing the seven sources of congestion. The San Diego, Sacramento/Lake Tahoe, Atlanta, and New York/New Jersey case studies include investigations sources of non-recurrent congestion.

Other Use Case investigations demonstrate specific use cases for various types of users described in the Supplement D: Use Case Demonstrations. The San Diego case study includes investigations of use cases including using planning-based reliability tools.

San Diego

This case study focused on using a mature reliability monitoring system in San Diego, California to illustrate the state of the art for existing practice. Led by its Metropolitan Planning Organization, the San Diego Association of Governments (SANDAG), and the California Department of Transportation (Caltrans), the San Diego region has developed one of the most sophisticated regional travel time monitoring systems in the United States. This system is based on an extensive network of sensors on freeways, arterials, and transit vehicles. It includes a data warehouse and software system for calculating travel times automatically. Regional agencies use these data in sophisticated ways to make operations and planning decisions.

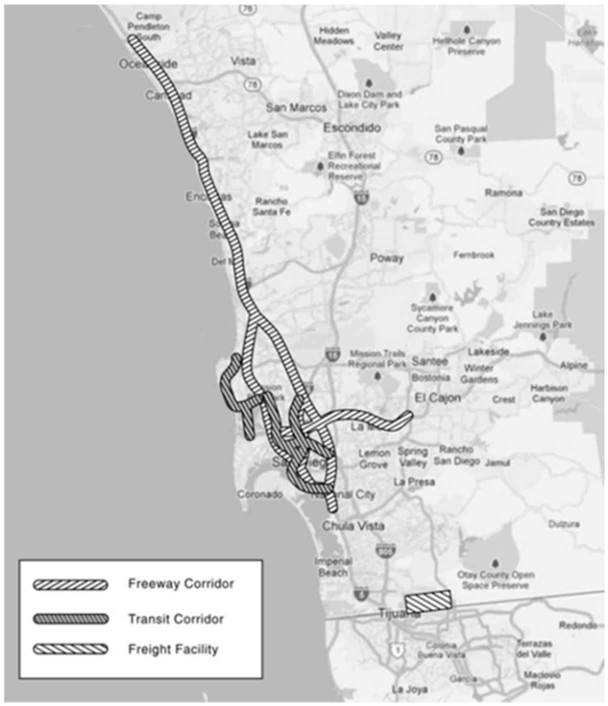
Because this technical and institutional infrastructure was already in place, the team focused on generating sophisticated reliability use case analysis. The rich, multimodal nature of the San Diego data presented numerous opportunities for state of the art reliability monitoring, as well as challenges in implementing guidebook methodologies on real data.

The purpose of this case study was to:

- Assemble regimes and travel time probability density functions from individual vehicle travel times
- Explore methods to analyze transit data from Automatic Vehicle Location (AVL) and Automated Passenger Count (APC) equipment
- Demonstrate high-level use cases encompassing freeways, transit, and freight systems

- Relate travel time variability to the seven sources of congestion

Figure 7-2 shows the study area for the San Diego case study.



Map data © 2012 Google

Figure 7-2. San Diego Case Study Area

The Caltrans District 11 encompasses San Diego and Imperial Counties and the metropolitan area of San Diego. A variety of detection systems are used in the study area to monitor freeways, arterials, and transit fleet. District 11 has 3,592 sensors, which are a mix of loop detectors and radar detectors, located at 1,210 locations on its freeways. District 11 also has 17 wireless vehicle sensors deployed to monitor intersection approaches on its arterials.

On the transit side, the San Diego Metropolitan Transit System (MTS) is currently supplying data from their real-time computer aided dispatch (CAD) system into an archived data user service. To monitor its transit fleet, MTS has equipped over one-third of its bus fleet with Automatic Vehicle Location (AVL) transponders and over one-half of its fleet with Automated Passenger Count (APC) equipment.

All Caltrans districts use the Performance Measurement System (PeMS) for data and performance measure archiving and reporting. District 11 uses an arterial extension of PeMS, the Arterial Performance Measurement System (A-PeMS), to collect and store its arterial data. District 11 also uses a transit extension of PeMS, the Transit Performance Measurement System (T-PeMS), to obtain schedule, AVL, and APC data from its existing real-time transit management system, compute performance measures based on these data, and aggregate and store them for further analysis.

Caltrans uses other management systems in conjunction with PeMS to operate its transportation network. For example, the California Highway Patrol's Computer Aided Dispatch (CAD) System provides an automated incident data feed that is fed into PeMS in real-time. Caltrans also keeps a non-automated database of incidents through its Traffic Accident Surveillance and Analysis System (TASAS). TASAS data are incorporated into PeMS with a two-year lag.

Freeway Analyses

This use case is primarily for the system planner and roadway manager user types. To perform this analysis, methods were developed to create travel time PDFs from large data sets of travel times that occurred under each congested condition. This use case analysis illustrates one potential method for linking travel time variability with the sources of congestion. In this case study, the research team opted to pursue a less sophisticated but more accessible approach than had previously been developed because it provides meaningful and actionable results without requiring agency staff to have advanced statistical knowledge. The application of the methodology to the two study corridors in San Diego revealed key insights into how this type of analysis should be performed, as detailed in the San Diego Case Study Resource Document.

This case study demonstrated an additional five high-level use cases that broadly encompass reliability information of interest to various users of the transportation system. The specific use cases were developed to be well-suited for demonstration using the San Diego data sources. The use cases apply to roadway, transit, and freight users.

Freeway Use Case 2: Using planning-based reliability tools to determine departure time and travel time for a trip. This use case represents a function that would be used by drivers. The use case demonstration showed the route that is the fastest on average is not always the route that consistently gets travelers to their destination on-time.

Freeway Use Case 3: Combining real-time and historical data to predict travel times in real-time. This use case is primarily for the operations manager user type. This use case demonstration described in the San Diego Case Study Resource Document shows that it is possible to provide predictive travel time ranges and expected near-term travel times by combining real-time and archived travel time data. The travel time predictions for both study routes proved very similar to the actual travel times measured on the sample day.

Transit Analyses

The biggest data challenge in this case study was processing the transit data, which is stored in a newly developed performance measurement system. This case study represented the first research effort to use these data and this system. The research team found that data quality is a major issue when processing transit data to compute travel times. Many of the records reported by equipped buses had errors, which had to be programmatically filtered out.

Assembling route-based reliability statistics using a drastically reduced subset of good data presented the next challenge. From this experience, the research team concluded that transit travel time reliability monitoring requires a robust data processing engine that can programmatically filter data to ensure that archived travel times are accurate. Additionally, transit reliability analysis requires a long timeline of historical data, due to the fact that, typically, a subset of buses is monitored and a large percentage of obtained data points will prove invalid.

Seven Sources Analysis: Offline analyses were conducted on the relationship between travel time variability and the seven sources of congestion. This use case serves a function primarily used by transit planners and operators. This use case analysis, described further in the San Diego Case Study Resource Document, illustrates one method for exploring the relationship between travel time variability and the sources of congestion. The application of the methodology to the three San Diego routes revealed key insights into how this type of analysis should be performed.

Transit Use Case 2: Using planning-based reliability tools to determine departure times and travel times for a trip. This use case primarily serves the transit passenger user type. This use case demonstration resulted in departure times and corresponding planning times for two bus routes. The methodology is described in detail in the San Diego Case Study Resource Document. The demonstration of this use case concluded that the most direct analysis would be achieved by restricting the date range to dates with identical schedules.

Transit Use Case 3: Analyzing the effects of transfers on the travel time reliability of transit trips. This use case primarily serves the transit operator user type. It was concluded that unusually long in vehicle travel times can have a larger effect on traditional reliability measures than missed transfers, potentially hiding the existence of missed transfers on a route.

Freight Analyses

Freight Use Case: Using freight-specific data to study travel times and travel time variability across an international border crossing. This use case represents a functionality that would primarily be used by freight service providers. This use case demonstration represented an initial use of truck travel time data from the Otay Mesa border crossing to evaluate travel time reliability for different aspects of a border crossing. By understanding where the bottlenecks are in the border crossing process and how they are impacting travel times and reliability, managers can begin to take steps to improve operations.

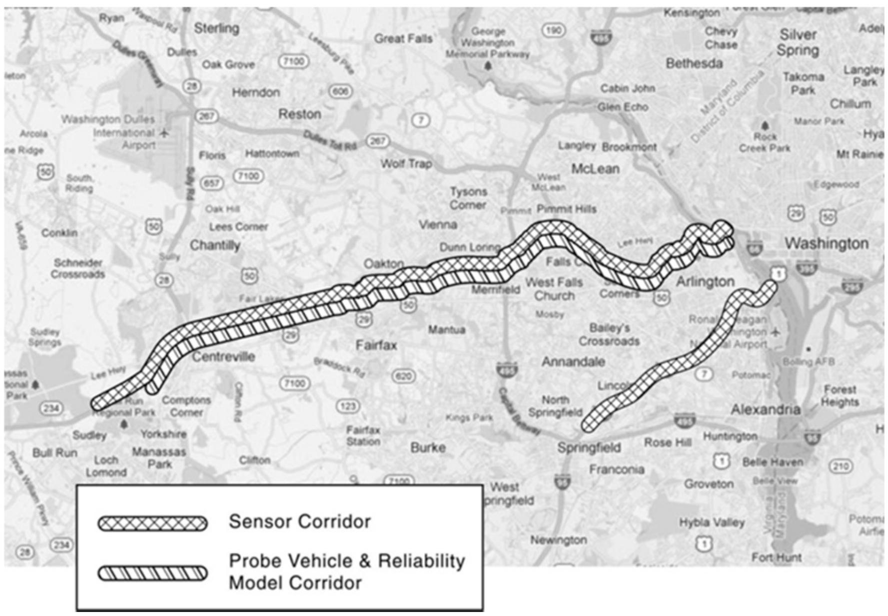
Northern Virginia

This case study provides an example of a more traditional transportation data collection network operating in a mixture of urban and suburban environments. Northern Virginia was selected as a case study site because it provided an opportunity to integrate a reliability monitoring system into a pre-existing, extensive data collection network. The focus of this case study was to describe the required steps and considerations for integrating a travel time reliability monitoring system into existing data collection systems.

The purpose of this case study was to:

- Describe the data acquisition and processing steps needed to transfer information between the existing system and the PeMS reliability monitoring system
- Demonstrate methods to ensure data quality of infrastructure-based sensors by comparing probe vehicle travel times using the procedures described in Chapter 3
- Develop multi-state travel time reliability distributions from traffic data

The study area for this investigation comprises the Interstate 66 (I-66) freeway from Manassas to Arlington, Virginia and the Interstate 395 (I-395) freeway from Springfield to Arlington, Virginia. Figure 7-3 shows the study corridors for the Northern Virginia case study.



Map data © 2012 Google

Figure 7-3: Northern Virginia Study Area

The Northern Virginia (NOVA) District of the Virginia Department of Transportation (VDOT) includes over 4,000 miles of roadway in Fairfax, Arlington, Loudoun, and Prince William counties. Traffic operations in the District are managed from the Northern Virginia Traffic Operations Center, which manages more than 100 miles of instrumented roadways, including HOV facilities on Interstates 95/395, 295, 66, and the Dulles Toll Road. The Northern Virginia Traffic Operations Center has deployed a wide range of technologies to support its activities, including cameras, dynamic message signs, ramp meters, and lane control signals.

In Northern Virginia, VDOT has deployed an extensive network of point-based detectors (primarily inductive loops and radar-based detectors), which are described in Chapter 1, to facilitate real-time collection of volume, occupancy, and (limited) speed data on freeways. A key component of the case study is ensuring data quality of infrastructure-based sensors, as described in Chapter 2.

To monitor regional travel conditions, the Northern Virginia District collects data from a wide range of sources on area freeways, including multiple types of traffic sensors and third parties such as INRIX, Trichord, and Traffic.com. The Northern Virginia Case Study Resource Document contains details about the types of traffic sensors and their specific locations.

Northern Virginia's Freeway Management System (FMS) is operated by VDOT staff located at the Traffic Operations Center (TOC). Staff members use the FMS to monitor and manage traffic, respond to incidents, and disseminate traveler information. In addition to managing freeway-related operations, VDOT staff use the NOVA Smart Traffic Signal System (STSS) to manage surface street and arterial systems in the region, monitoring, controlling, and maintaining over 1,000 traffic signals within their jurisdiction.

System Integration

For purposes of this case study, data from NOVA's data collection network and management system were integrated into a developed archived data user service and travel time reliability monitoring system. The steps and challenges encountered in enabling the information and data exchange between these two large and complex systems are described in detail in the Northern Virginia Case Study Resource Document. The goal of this experiment was to provide agencies with a real-world example of the resources needed to accomplish data collection to monitoring system integration, and the likely challenges that will be encountered when procuring a monitoring system.

NOVA equipment configuration information was obtained from an XML file posted on the RITIS website. The issues with fitting the data into the PeMS configuration related to conflicting terminologies, information required by PeMS that was missing from the configuration file, and equipment types not supported by PeMS. The Northern Virginia Case Study Resource Document describes the issues in more detail. The Resource Document also describes the metadata quality control steps that were used to insert NOVA configuration information into PeMS.

Configuring PeMS to receive NOVA data helped define the requirements for complex traffic systems integration and illustrate what agencies can do to facilitate the process of implementing reliability monitoring. The process of fully integrating the NOVA data with PeMS took several weeks. Assuming that agencies are interested in acquiring PeMS or a similar system, there are a number of steps that agencies can take to make this integration go more smoothly and quickly.

First, it is important that the implementation and maintenance of a traffic data collection system be carried out with a broad audience in mind. Often, increasing access to data outside of an organization can help to further agency goals; for example, providing data to mobile application developers can help agencies distribute information in a way that increases the efficiency of the transportation network.

One of the ways that agencies can facilitate the distribution of data from their data collection system is by establishing one or more data feeds. Since maintaining multiple data feeds can be a challenge, if agencies want to provide a feed of processed data, it will save resources in the long run to document the processing steps performed on the data (see also Soriguera 2011). This will allow implementers of external systems to evaluate them and undo them, if needed.

Aside from the processing documentation, maintaining clear documentation on the format of data files and units of data will greatly facilitate the use of data outside of the agency. Additionally, documentation on the path of data from a detector through the agency's internal systems can be of value to contractors and other external data users. Clearly explaining this information in a text file minimizes the back-and-forth communication between agency staff and contractors and prevents inaccurate assumptions from being made.

Probe Vehicle Comparisons

The team performed a quality control procedure to better understand the implications of the data quality issues on travel times. The primary question the team wanted to answer in this probe-based experiment was: how well do the probe data align with the traffic speed and travel time estimates provided by the sparsely deployed point-based detectors? Probe vehicle runs were conducted along I-66 to amass "ground-truth" data that could be compared with the sensor data. In addition to analyzing speed data, the team conducted an analysis of the differences between the travel times experienced by the probe vehicle during each trip versus the estimated travel times generated from the sensor speeds. It was determined that the steadiness of the travel time estimates from the sensors is not ideal for computing travel time reliability, which relies on the ability of the system to detect variability in traffic conditions over time. As a result, it is highly unlikely that these sensors would provide accurate travel times under most congested conditions.

The authors' analysis of the data available from these sensors has yielded a number of findings of potential interest to a wide variety of agencies, particularly those facing maintenance and calibration issues associated with older sensor systems, as well as those agencies with more sparsely spaced spot sensors. Overall, five primary factors were identified that accounted for differences between the probe vehicle data and speed / estimated travel times generated based on VDOT sensor data. These factors are described in the Northern Virginia Case Study Resource Document. Public agency staff should take the factors into consideration when making decisions concerning the deployment of new data collection infrastructure and the maintenance and expansion of existing systems.

Analyses of PDFs with Multiple Statistical Modes

Because of the type of data available in this case study and investigations done previously in the I-66 corridor, the research team elected to experiment with travel time reliability monitoring ideas that are being developed in SHRP 2 project L10, Feasibility of Using In-Vehicle Video Data to Explore How to Modify Driver Behavior that Causes Non-Recurrent Congestion. In the SHRP 2 project L10, researchers are experimenting with a multi-state travel time reliability modeling framework using mixed mode normal distributions to represent the PDFs of travel time data from a simulation model of eastbound I-66 in Northern Virginia. This case study adopted that technique and applied it to the travel times calculated from the freeway loop detectors on eastbound I-66.

The goal of this study was to generate, for each hour of the day, two outputs: the percent chance that the traveler would encounter a certain condition and the average and 95^{th} percentile travel time for each condition. The methodology to answer these questions and the results of the analysis are described in the Northern Virginia Case Study Resource Document.

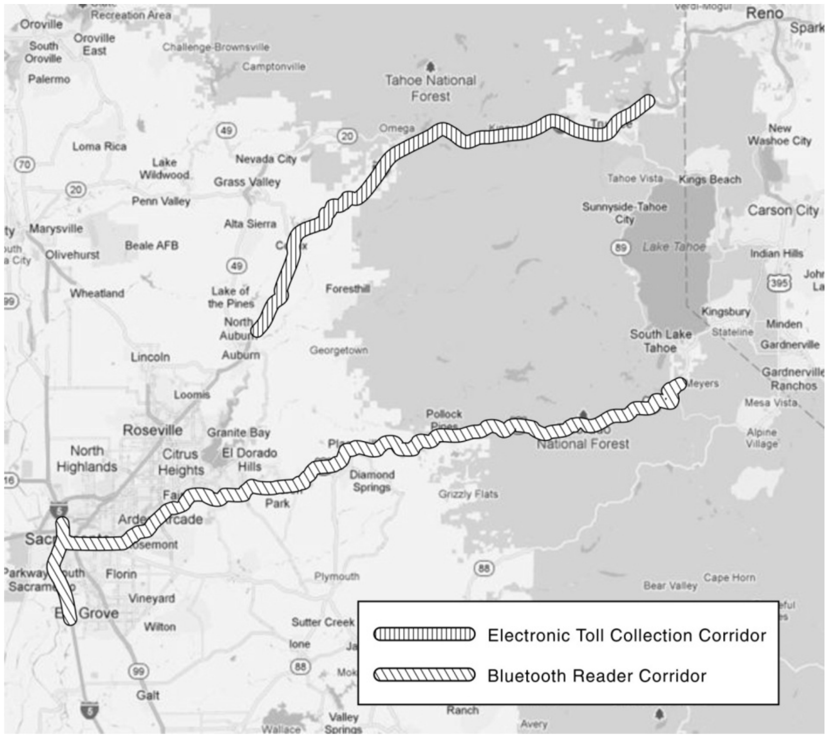
The methodological findings of this investigation are that multi-state normal distribution models can approximate travel time distributions generated from loop detectors better than normal or log-normal distributions. During the peak hours on a congested facility, three states are generally sufficient to balance a good model (distribution) fit with the need to generate information that can be easily communicated to interested parties. During off-peak hours, two states typically provide a reasonable model or distribution fit. The outputs of this method can inform travelers of the percent change that they will encounter moderate or severe congestions and, if they do, what their expected and 95th percentile travel times will be.

Sacramento/Lake Tahoe

This case study illustrates an example of a rural transportation network with fairly sparse data collection infrastructure. The purpose of this case study was to:

- Examine vehicle travel time calculation and reliability using Bluetooth and RFID re-identification systems
- Filter out travel time from trip time collected by Bluetooth and Electronic Toll Collection (ETC) devices
- Explore the following aspects of the ETC and Bluetooth reader networks used in the Lake Tahoe region: (1) detailed locations and mounting structures; (2) lanes and facilities monitored; (3) percentage of traffic sampled; and (4) percentage and number of vehicles re-identified between readers
- Quantify the effects of adverse weather and demand-related conditions on travel time reliability using data derived from Bluetooth and ETC systems

The study area for this case study comprises the Interstate 5 (I-5) freeway through Sacramento, California and the two highways leading east to Lake Tahoe: Interstate 80 (I-80) and US Highway 50 (US-50). Figure 7-4 shows the study corridors for the Sacramento/Lake Tahoe case study.



Map data © 2012 Google

Figure 7-4: Sacramento/Lake Tahoe Study Area.

This case study is located in the Caltrans District 3, which encompasses the Sacramento metropolitan area and the Sacramento Valley and Northern Sierra regions of California. District 3 includes urban, suburban, and rural areas, including areas near Lake Tahoe where weather is a serious travel time reliability concern and there is heavy recreational traffic. Two major interstates pass through the District: I-80, which is oriented generally east/west, and I-5, which is oriented generally north/south along the west side of the Sacramento and San Joaquin Valleys. Other major freeway facilities include US-50, which connects Sacramento and South Lake Tahoe, and State Route 99 (SR-99), which runs north/south along the east side of the Sacramento and San Joaquin Valleys.

Caltrans District 3 currently only collects traffic data along freeway facilities. It operates a total of 2,251 point detectors (either radar detectors or loop detectors) located in over 1,000 roadway locations in the District. To supplement the point detection network, the District has installed 32 non-revenue generating ETC readers (25 on I-80 and 7 on US-50) in rural portions of the Sierra Nevada Mountains near Lake Tahoe. Details about the locations of these ETC readers can be found in the Sacramento/Lake Tahoe Case Study Resource Document.

All Caltrans districts use PeMS for data and performance measure archiving and reporting as described at the beginning of this chapter. Caltrans uses other management systems in conjunction with PeMS to operate its transportation network. The California Highway Patrol's

Computer Aided Dispatch (CAD) System provides an automated incident data feed that is fed into PeMS in real-time. Caltrans also keeps a non-automated database of incidents through its Traffic Accident Surveillance and Analysis System (TASAS). TASAS data are incorporated into PeMS with a two-year lag.

AVI Sensor Deployment

The two sources of data used in support of this case study, based on the movement of vehicles equipped with ETC and Bluetooth devices, are extremely new and not currently integrated into Caltrans District 3’s existing PeMS data feed. Consequently, it was necessary to ingest these data sets into project–specific instances of PeMS for analysis as part of this project. The pre-requisite data collection through monitoring system integration-related activities included ETC data and Bluetooth data is described in the Sacramento/Lake Tahoe Case Study Resource Document.

This case study explored four aspects of the ETC and Bluetooth reader networks used in the Sacramento/Lake Tahoe case study: (1) detailed locations and mounting structures; (2) lanes and facilities monitored; (3) percentage of traffic sampled; and (4) percentage and number of vehicles re-identified between readers. As a whole, it showed that vehicle re-identification technologies are suitable for monitoring reliability in rural environments, provided traffic volumes are high enough to generate a sufficient number of samples.

For rural areas that have heavy recreational or event traffic, vehicle re-identification technologies such as ETC and Bluetooth can provide sufficient samples to calculate accurate average travel times at a fine granularity during high-traffic time periods. During these high-volume periods, vehicle re-identification technologies can be used to monitor travel times and reliability over long distances, such as between the rural region and nearby urban areas.

For agencies deploying vehicle re-identification monitoring networks, it is necessary to understand that the quality of the collected data is highly dependent on the decisions made regarding ETC and Bluetooth technologies during the design and installation process. For agencies leveraging existing networks, it is important to fully understand the configuration of the network before using its data.

Travel Time Calculations

Due to the significant amounts of Bluetooth-based travel time data available for analysis as part of this case study, the research team elected to focus its methodological efforts on this dataset rather than on data generated by the ETC-based system.

The primary goal of Bluetooth reader (BTR) based data analysis is to characterize segment travel times between BTRs based on the re-identification (re-id) of observations derived from unique mobile devices. Generally, the data processing procedures associated with the calculation of BTR-to-BTR travel times can be broadly broken down into three processes, which are discussed in detail in the Sacramento/Lake Tahoe Case Study Resource Document:

- Identification of Passage Times
- Generation of Passage Time Pairs
- Generation of Segment Travel Time Histograms

This case study evaluated various methodological approaches and processes for estimating ground-truth segment travel times based on Bluetooth data, which are described in the Sacramento/Lake Tahoe Case Study Resource Document. A number of factors were identified that influence travel time reliability and guided the development of methods for processing re-id observations and calculating segment travel times. The results show that smart filtering and processing of Bluetooth data to better identify likely segment trips increases the quality of calculated segment travel time data. This approach helps preserve the integrity of the data set by retaining as many points as possible and basing decisions to discard points on the physical characteristics of the system rather than their statistical qualities.

For either of the data collection technologies described in this report to be successful over the long-term, safeguards must be put into place to ensure that the privacy of individual drivers being sampled is protected (see Karr *et al.* 2007 and National Institute of Statistical Sciences 2004, for example). It is recommended that any probe data collection program implemented by public agencies or private sector companies on their behalf adhere to a pre-determined set of privacy principles (e.g., see Briggs and Walton 2000) aimed at maintaining the anonymity of specific users. Additionally, any third party data provider working for a public agency to implement a travel time data collection solution based on either of the technologies described in this case study should be required to submit an affidavit indicating that they will not use data collected on the agency's behalf in an inappropriate manner.

Integration of Sources of Non-Recurrent Congestion

The purpose of this use case was to quantify the impact of adverse weather and demand-related conditions on travel time reliability using data derived from the case study's Bluetooth and ETC-based systems deployed in rural areas. To examine travel time reliability within the context of this use case, methods were developed to generate PDFs from large quantities of travel time data representing different operating conditions. To facilitate this analysis, travel time and flow data from ETC readers deployed on I-80W and Bluetooth readers deployed on US-50E and US-50W were obtained from PeMS and compared with weather data from local surface observation stations. PDFs were subsequently constructed to reflect reliability conditions along these routes during adverse weather conditions, as well as according to time-of-day and day-of-week. The PDFs of travel times under different operating conditions consistently demonstrated the unreliability associated with low visibility, rain, and travel under high-demand conditions.

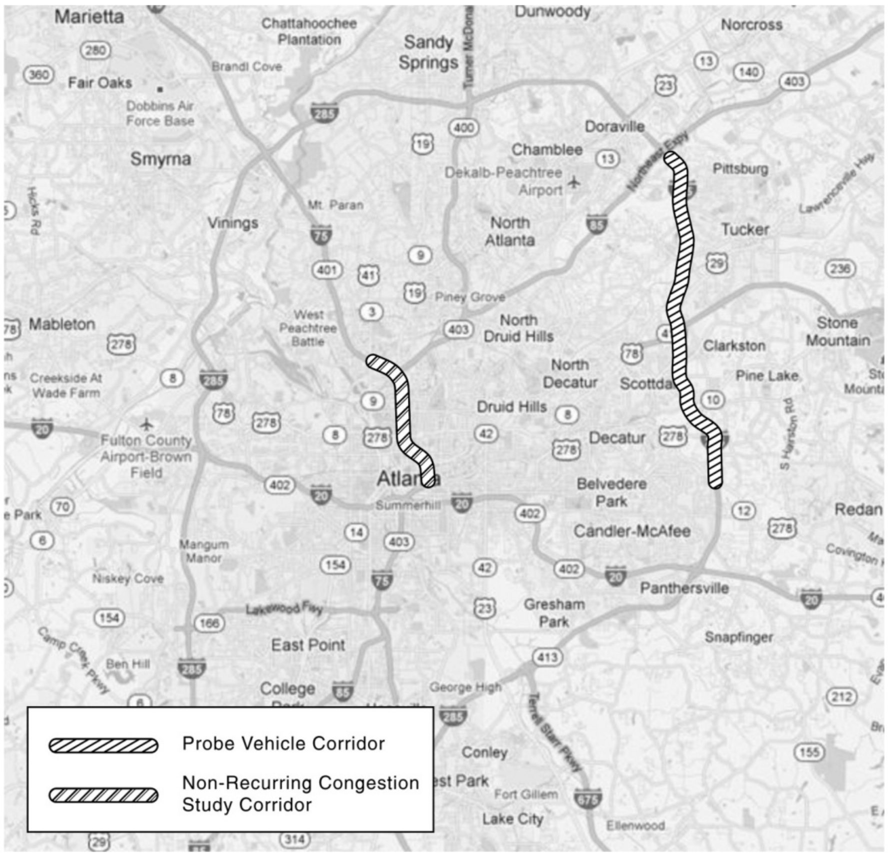
Atlanta

The team selected the Atlanta Metropolitan Region to provide an example of a mixed urban and suburban site that primarily relies on video detection cameras for real-time travel information. The main objectives of the Atlanta case study were to:

- Demonstrate methods to resolve integration issues by using real-time data from Atlanta's traffic management system for travel time reliability monitoring
- Compare probe data from a third-party provider with data reported by agency-owned infrastructure

- Fuse the regime-estimation and non-recurrent congestion analysis methodologies to inform on the reliability impacts of non-recurrent congestion

Figure 7-5 shows the study corridors investigated in the Atlanta case study.



Map data © 2012 Google

Figure 7-5: Atlanta Study Area

In the Atlanta region, the Georgia Department of Transportation (GDOT) collects data from over 2,100 roadway sensors, which include a mix of video detection sensors and radar detectors. Both of these types of sensors consist of single devices that monitor traffic across multiple lanes. The majority of active sensors are monitoring freeway lanes, with some limited coverage of conventional highways. Sensors in the active network are manufactured by four different vendors. In general, the different types of sensors are divided up by freeway. The Atlanta Case Study Resource Document provides more details about the sensor vendors and the location of active mainline sensors in the GDOT network categorized by manufacturer. To deepen the case study analysis and explore alternative data sources, the project team acquired a parallel, probe traffic data set, provided by NavTeq. The data set covers the entirety of the Interstate 285 (I-285) ring road, and is reported by Traffic Message Channel ID. One use case of this case study focuses on comparing probe data from a third-party provider with data reported by agency-owned infrastructure.

GDOT monitors traffic in the Atlanta Metropolitan Area in real-time through its Advanced Traffic Management System (ATMS), called Navigator. The Transportation Management Center (TMC), located in Atlanta, is the headquarters and information clearinghouse for Navigator. GDOT's traffic management system integrates with traffic sensors, CCTVs, changeable message signs (CMS), ramp meters, weather stations, and Highway Advisory Radio (HAR).

Navigator was initially deployed in metropolitan Atlanta in preparation for the 1996 Summer Olympic Games. Navigator collects lane-specific volume, speed, and occupancy data in real-time and stores it in a database table for 30-minutes. Every fifteen minutes, the raw Navigator traffic data samples are aggregated up to lane-specific 15-minute volumes, average speeds, and average occupancies, and archived for each detector station. The data are not filtered or quality-controlled prior to being archived.

Aside from the traffic data, Navigator also maintains a historical log of incidents. When the TMC receives a call about an incident, TMC staff log it as a "potential" incident in Navigator, until it can be confirmed through a camera or multiple calls. Once the incident has been confirmed, its information is updated in Navigator to include the county, type of incident, and estimated duration. This incident information is archived and stored.

For the purposes of this case study, data from GDOT's Navigator system was integrated into PeMS, a developed archived data user service and travel time reliability monitoring system. Two aspects of the Navigator framework presented major challenges for incorporating the traffic data into PeMS:

- The frequency of data reporting differs for different device types
- Many video detection system (VDS) device data samples are missing

As such, one experiment of this case study focuses on resolving these integration issues to ensure data quality.

System Integration

The first system integration experiment details how the integration issues of using *ATMS* data for travel time reliability monitoring were resolved. The experiment showed that unstructured configuration information obtained from ATMS requires careful analysis when mapping to the data model of a reliability monitoring system. It also highlights the importance of understanding the reporting frequency and form of detector data for ensuring accurate aggregation and travel time calculation.

The second experiment compared the speed data reported by agency-owned infrastructure with probe data obtained from a third-party provider on the I-285 ring road. Results showed the speeds between data types to be similar during the peak hours, but that the third-party provider artificially capped speeds to remain below a certain threshold. The experiment also investigated the speed error introduced by the differences in locations between the agency-owned infrastructure and the midpoint of its associated third-party link (defined by Traffic Message

Channel ID). Some difference in reported speeds was attributed to the distance of the agency-owned detection from the mid-point of the third-party provider link.

Integration of Sources of Non-Recurrent Congestion

The use case analysis applied the methodological advancement techniques established and demonstrated in previous case studies to travel time data on a downtown Atlanta corridor to interpret the impact of the seven sources of non-recurrent congestion on travel time reliability.

Two of the main themes of the case study demonstrations are: estimating the quantity and characteristics of the operating travel time regimes experienced by different facilities and calculating the impacts of the seven sources of non-recurrent congestion on travel time reliability. The methodological goal of the Atlanta case study is to fuse the previously-developed regime-estimation and non-recurrent congestion analysis methodologies by using multi-state models to inform on the reliability impacts of non-recurrent congestion. This developed method consisted of three steps:

1) **Regime Characterization**, to estimate the number and characteristics of each travel time regime measured along the facility;
2) **Data Fusion**, to link travel times with the causal factor (such as weather or incident) active during their measurement, and;
3) **Seven Sources Analysis**, to calculate the contributions of each source on each travel time regime.

Analysis showed that the study corridor operates with two regimes during the peak period, with the more congested and variable regime composed of many travel times influenced by traffic incidents. This case study showed that, with proper quality control and integration measures, ATMS data can be used for travel time reliability monitoring, including the linking of travel time variability with the sources of non-recurrent congestion.

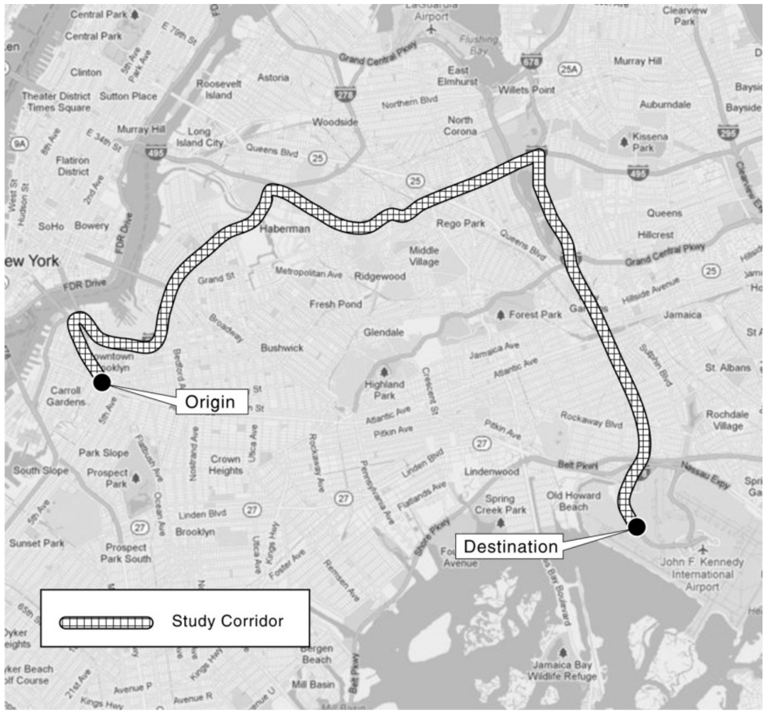
New York/New Jersey

The New York City site was chosen to provide insight into travel time monitoring in a high-density urban location. The 2010 United States census revealed New York City's population to be in excess of 8 million residents, at a density near 28,000 people per square mile. While New York City has a low rate of auto ownership compared to other United States cities, more than half of all commute trips are still made in single-occupancy vehicles. In 2010, these factors contributed to New York City having the longest average commute time of any United States city, at 31.3 minutes.

The main objectives of the New York/New Jersey case study included:

- Obtaining time-of-day travel time distributions for a study route based on probe data
- Identifying the cause of bi-modal travel time distributions on certain links
- Exploring the causal factors for travel times that vary significantly from the mean conditions

The route analyzed in this case study begins in the Boerum Hill neighborhood of Brooklyn and ends at JFK Airport, traversing three major freeways: the Brooklyn-Queens Expressway (I-278), the Queens-Midtown Expressway (I-495), and the Van Wyck Expressway (I-678). Figure 7-6 shows the study route from origin to destination.



Map data © 2012 Google

Figure 7-6: New York/New Jersey Study Area

Another reason the New York/New Jersey site was selected is because it is covered by a probe data set, provided to the research team by ALK Technologies, a third-party data provider. These data are composed of GPS traces collected from mobile devices inside individual vehicles. This detection technology provides high-density information along the vehicle's entire path, as opposed to infrastructure-based sensors which measure traffic only at discrete points. This probe data set was analyzed at two levels: at the individual GPS trace level and through aggregation into single per-link speed values. The raw GPS trace data is the only case study data set that traces the entire path of vehicle trips. The aggregated speeds are similar in format to the Traffic Message Channel path-based data analyzed in the Atlanta case study. The data obtained for this case study covers a rectangular region around the study route.

A static collection of historical probe data provided the basis for analysis in this case study. No real-time data was acquired or analyzed. Unlike other sites, an Archived Data User Service (ADUS) was not specifically deployed for this case study.

System Integration

The first investigation describes how to obtain route travel time distributions from the probe data set. This experiment discusses the data density along the route, presents methods for visualizing individual probe trips within the context of historical conditions, and details three techniques for constructing route-level travel time distributions. The central outcome of this experiment is the comparison of time-of-day travel time distributions along the route constructed using each of the three techniques. Methods were developed to compare a particular probe vehicle's path with the 25th percentile, 75th percentile, and median speed profile along the route by time-of-day. Probe traces are also visualized within historical speed bounds based on location and time-of-day. This methodology makes it possible to simulate the upper and lower bound of expected trip trajectories from a particular point along the route, based on the historical travel times.

The raw ALK probe data is in the form of standard NMEA GPS sentences taken directly from the probe vehicles. These data are further processed by ALK into link-based speed measurements. Although each data point contains rich information, the data set is sparse in that few probe vehicles traverse the entire route from beginning to end. As such, the route travel time distribution must be constructed piecemeal from individual link data. Obtaining composite travel time distributions from vehicles that only traveled on a portion of the route is a complex process, most notably because this project has shown that travel times on consecutive links have a strong linear dependence. This linear dependence must be accounted for when combining individual link travel times into an overall route travel time distribution. This is the core methodological challenge of this case study.

Three methods for computing route PDFs from the available probe data are compared:

1) **Constructing the PDFs carefully from direct measurements.** This method begins by determining the distribution of speed measurements on the first link of the route. This distribution is combined with the travel time distributions of longer trips that also traversed the initial link. Incrementally, longer trips are added to the distribution until a speed distribution for the entire route is obtained. Trips are grouped by time of day, at an hourly granularity when the data density allows.
2) **Constructing the PDFs with a Monte-Carlo simulation.** This method considers consecutive pairs of links along the route, e.g., link 1 and link 2, link 2 and link 3, etc. It constructs the full route PDF out of a large number of simulated trips. Each simulation begins with the sampling of a travel time on the first link. Next, the correlation between travel times on link 1 and link 2 is examined and a travel time sample on link 2 is taken based on this correlation and the original link 1 sample. This procedure is repeated for link 3, based on the previous link 2 sample and the correlation between links 2 and 3, and continues until a single trip along the entire route has been simulated. A large number of these simulated trips form the full travel time distribution for the route.
3) **Constructing the PDFs assuming link speed independence.** This method ignores the linear dependence between consecutive links and directly computes the route travel time distribution as if all link travel times were independent. It works by simply convolving the distributions of travel times on consecutive links. For example, the frequency distribution of travel times on the first link will be added to the frequency distribution of

travel times on the second link, and so on until a full travel time distribution for the entire route is obtained.

This case study showed that it is possible to obtain trip reliability measures based on probe data, even when that probe data is sparse. The travel time distribution for the route is constructed from vehicles that only travel on a portion of the route, and takes into account the linear dependence of speeds on consecutive links. This case study also contributes techniques for creating time-space contour plots based on probe speeds. These contour plots can be made to represent any measured speed percentile, so that contours for the worst observed conditions can be compared with typical conditions.

Travel Time Distributions

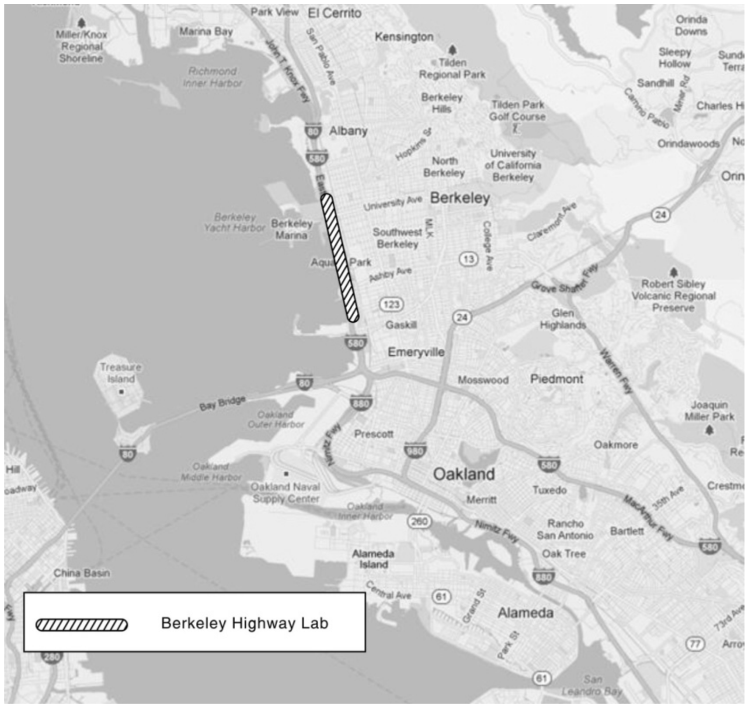
The second experiment details an investigation into the cause of bi-modal travel time distributions on certain links. Time of day, day of week, and non-recurrent congestion sources are explored as a source of the bimodality.

Integration of Sources of Non-Recurrent Congestion

The use case analysis explores the associated factors for travel times that vary significantly from the mean conditions. This use case represents this case study's investigative analysis of the seven sources of non-recurrent congestion on travel time reliability.

Berkeley Highway Lab

One objective of the case studies is to test and refine the methods developed for defining and identifying segment and route regimes for freeway and arterial networks. The team's research to date has focused on identifying operational regimes based on individual vehicle travel times and determining how to relate these regimes to system-level information on average travel times. Since individual vehicle travel times on freeways are not available in the San Diego metropolitan region, data from the Berkeley Highway Laboratory (BHL) was used in this analysis. Details about the Berkeley Highway Lab applications can be found in the San Diego Case Study Resource Document. Figure 7-7 shows the BHL location.



Map data © 2012 Google

Figure 7-7: Berkeley Highway Laboratory Study Area

The Berkeley Highway Laboratory (BHL) is a 2.7-mile section of Interstate 80 (I-80) in west Berkeley and Emeryville. The BHL includes fourteen surveillance cameras and sixteen directional dual inductive loop detector stations dedicated to monitoring traffic for research purposes. The sensors are a unique resource because they provide individual vehicle measurements. The corridor was also temporarily instrumented with two Bluetooth reader stations (BTRs) along eastbound I-80 to record the timestamps and Media Access Control (MAC) addresses of Bluetooth devices in passing vehicles.

System Integration

Data from the Berkeley Highway Laboratory section of I-80 was used in this case study. This section is valuable because it has co-located dual loop detectors and Bluetooth sensors. This dataset provided an opportunity for the team to begin to assemble regimes and travel time probability density functions from individual vehicle travel times. These travel time PDFs are needed to support motorist and traveler information use cases. Since the majority of the case study sites did not provide data on individual traveler variability, it was important for the research team to study the connection between individual travel time variability and aggregated travel times, and whether the former can be estimated from the latter.

Analysis was performed on a day's worth of BHL data from the BTRs and loop detector stations. This analysis examined data from the Berkeley Highway Lab to see if operative regimes for individual vehicle travel times can be identified from Bluetooth data. The research team concluded that this can, indeed, be done. Based on more than 5,000 observations of individual travel times, three different regimes can be identified: (1) off-peak or uncongested; (2) peak or congested; and (3) transition between congested and uncongested. All three can be characterized by 3-parameter Gamma density functions, as demonstrated in the San Diego Case Study Resource Document.

Use Cases

A functioning reliability monitoring system must meet the needs of many different types of users because different users perceive and value deviations from the expected travel time in different ways. Each of these user classes has different motivations for monitoring travel time reliability, and these needs have to be accounted for in the types of analysis that the system can support through the user interface. Use cases are a formal systems engineering construct that transforms user needs into short, descriptive narratives that describe a system's behavior. Use cases capture a system's behavioral requirements by detailing scenario-driven threads through the functional requirements. The collective use cases define the monitoring system by capturing its functionalities and applications for various users.

Supplement D: Use Case Demonstrations provides a series of use cases to help readers of the guidebook determine what information the travel time reliability monitoring system needs to produce and what applications it needs to satisfy their specific situation. Once the appropriate users and their needs for reliability information are defined, the guidebook reader can determine the performance measures, spatial coverage, data interface needs (i.e., weather, crashes, construction activity, special events), and archival requirements for their monitoring system.

The use cases are organized around the various stakeholders that use or manage aspects of the surface transportation system. The use cases for each aspect of the transportation system are also broken down into providers and consumers – supply and demand:

- **Policy and Planning Support**: Agency administrators and planners that have responsibility for and make capital investment decisions about the highway network.
- **Overall Highway System**: Operators of the roadway system (supply), including its freeways, arterials, collectors, and local streets and drivers of private autos, trucks, and transit vehicles (demand).
- **Transit Sub-system**: Operators of transit systems that operate on the highway network, primarily buses and light rail (supply) and riders (demand).
- **Freight Sub-system**: Freight service suppliers (supply) and shippers and receivers that make use of those services (demand).

Supplement D describes several use cases for each user type listed above. The use cases for system administrators and planners are shown in Table 7-1 as an example of the types of use cases considered in this guidebook. The list of use cases in Exhibit 4-8 illustrates the types of functionality that may be desired by administrators and planners. The use cases for other user types are provided in Supplement D.

Table 7-1: Use Cases for System Administrators and Planners

Category	**Subgroup**	**Use Cases**
System Administrators and Planners	Administrators	See How Facilities are Affected by the Seven Factors Assess the Contributions of the Factors View the Travel Time Reliability for a Subarea Assist Planning and Programming Decisions Document Agency Accomplishments Assess Progress Toward Long-Term Reliability Goals Assess the Reliability Impact of a Specific Investment
	Planners	Find the Facilities with Highest Variability Assess the Reliability Trends over Time for a Route Assess Changes in the Hours of Unreliability for a Route Assess the Sources of Unreliability for a Route Determine When a Route is Unreliable Assist Rural Freight Operations Decisions

Each use case in Supplement D is described by specific parameters: a user, a statement of the question being posed, a description of the inputs needed to answer the question, the steps involved in answering the question, and the result to be obtained. Table 7-2 shows a template for the parameters provided for each use case.

Table 7-2: Use Case Template

Parameter	Description
User	The type of TTRMS user posing the question
Question	Description of the question being asked and why it would be posed.
Steps	A list of the actions that have to be performed to answer the question.
Inputs	The data and information that will be used to answer the question. This description helps users understand the inputs required; and programmers understand the data inputs that must be assembled.
Result	The system output at the completion of the use case.

CHAPTER 8: SUMMARY AND CONCLUSIONS

Project L02 within SHRP 2 was undertaken to create methods by which travel time reliability can be monitored, assessed, and communicated to end users of the transportation system. The project developed guidance for operating agencies about how they can put reliability measurement methods into practice by enhancing existing monitoring systems or creating new ones. The project's main product is a guidebook which describes how to develop and use a Travel Time Reliability Monitoring System (TTRMS). A set of supporting documents provide additional detail not found in the guidebook.

Travel time reliability is the absence of variation in travel times. If a system is reliable, people can get to where they want to go, when they want to be there, all the time. If a freeway is reliable, then its travel times are the same under all conditions, all year long. It is similar to a vehicle that always starts when the key is turned on. Of course, in reality no system or roadway is perfectly reliable; this project is intended to address this challenge.

L02 focused on how to measure reliability, how to understand what makes a system unreliable, and how to pinpoint mitigating actions. For example, the TTRMS will indicate the effects of congestion and if operational actions mitigate the impacts. The TTRMS analysis methods will let managers know if and how traffic incidents, weather, and other non-recurring events affect reliability, and the extent of the effect. Moreover, if actions are taken like the shoulders are widened or more roadside assistance trucks are deployed, it will show the impacts of those mitigations. (For a discussion about selecting mitigation strategies see Margiotta 2010. Also see Margiotta *et al*. 2006 for a guide to effective freeway performance measurement.)

Figure 8-1 shows the travel times for a specific trip in the San Diego, California, area that would have been experienced by someone who left at exactly the same time every weekday.

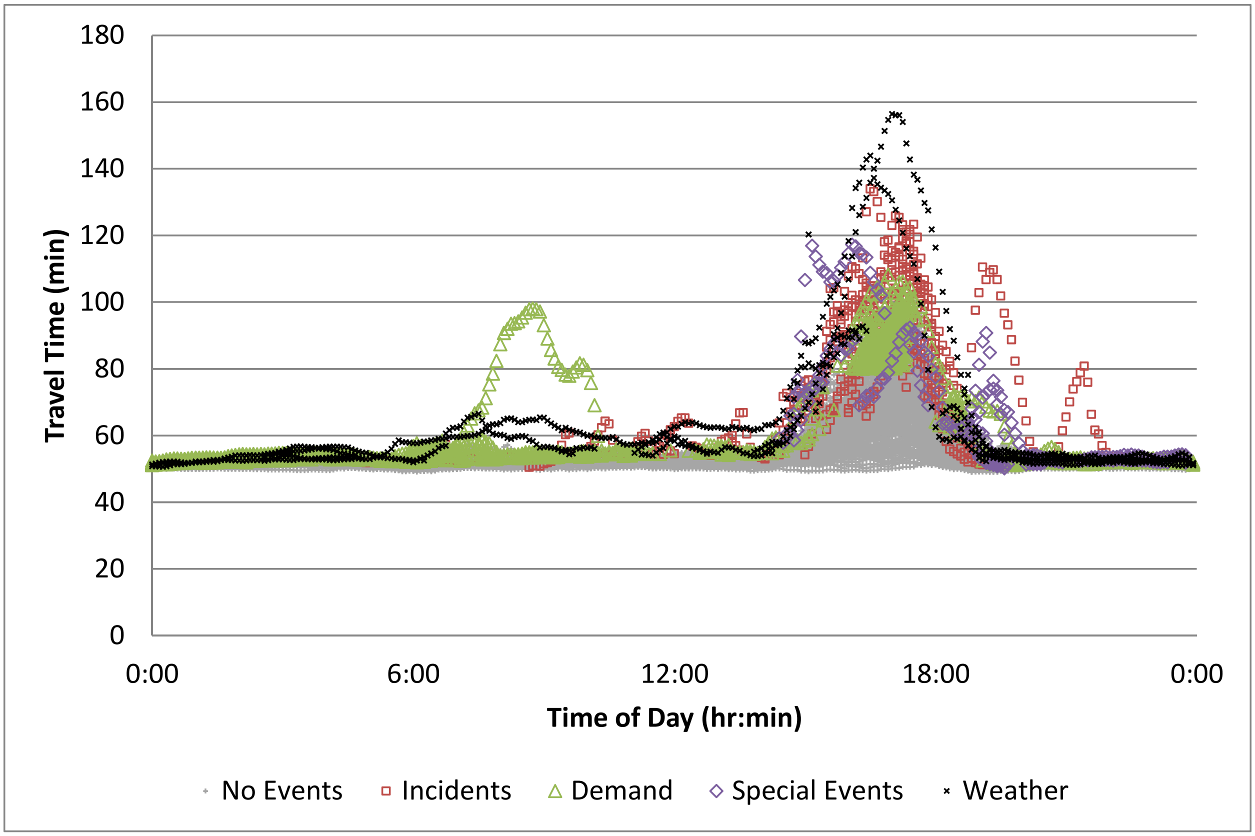


Figure 8-1: Variation in Travel Times by Time of Day Across a Year

It is clear from this exhibit that the travel times on this roadway segment are not always the same; the system *is* unreliable. Not only does the travel time vary but the spread in the times varies. At about midnight, the minimum and maximum are only 5 minutes different (50 minutes versus 55 minutes) but differ by 110 minutes during the weekday afternoon peak (50 minutes versus 160 minutes). It is also clear that non-recurring events have an impact. A good example is adverse weather, especially during the peak period. Traffic incidents also have an effect on travel time reliability, as do special events and unusually high demand. Even when no non-recurring event is happening—the "none" data points—the travel time can vary widely. The TTRMS helps indicate when, why, and by how much travel time will vary.

The TTRMS is designed to be an add-on to an existing traffic management system with a structure as shown in Figure 8-2. Inside the dotted line box are the three major modules of the TTRMS: a data manager, a computational engine, and a report generator. The data manager assembles incoming information from traffic sensors and other systems, such as weather data feeds and incident reporting systems, and places them in a database that is ready for analysis as "cleaned data". The computational engine works off the cleaned data to prepare "pictures" of the system's reliability: when it is reliable, when it is not, to what extent, under what conditions, etc. In the exhibit this is illustrated by "regime TT-PDFs". The report generator responds to inquiries from users—system managers or travelers—and uses the computation engine to analyze the data and provide information that can then be presented back to the inquirer or decision maker.

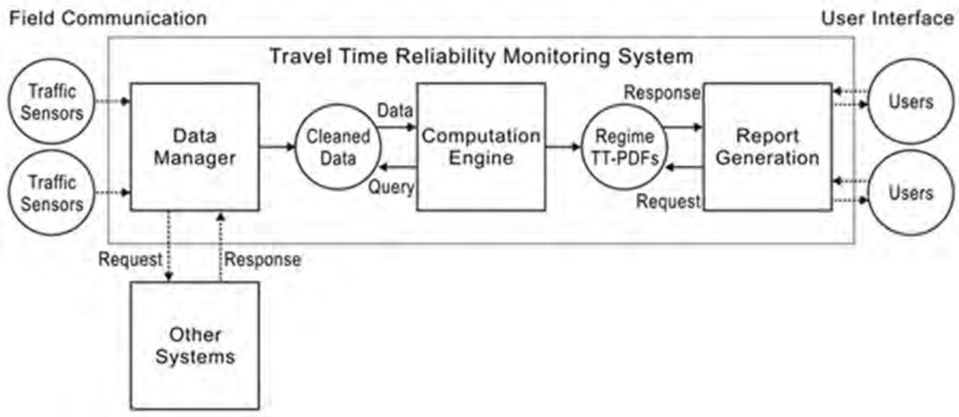


Figure 8-2: Block Diagram for a Travel Time Reliability Monitoring System (TTRMS)

Each of these modules is discussed and described in the guidebook. In addition, case studies and use cases illustrate how these modules work together to produce answers to questions that managers would likely pose. The supplemental material provides further details about how each of the modules should work – together and separately.

Figure 8-3 shows an example of what to expect as a report from the TTRMS. The plot shows the distribution of travel times on Interstate-8 westbound in San Diego across a three-month period under various operating conditions. The distributions are shown in a cumulative fashion; the location of each line shows how many travel times are that value or shorter. For example, when traffic incidents occur during heavy (recurrent) congestion, one half (50%) of the travel rates (seconds per mile) are up to 70 sec/mile. That is, 50% of the travel rates are this long or *shorter/smaller*. The 90^{th} percentile travel rate is 110 seconds per mile. Or put another way, 9 out of every 10 vehicles is traveling at that rate or faster.

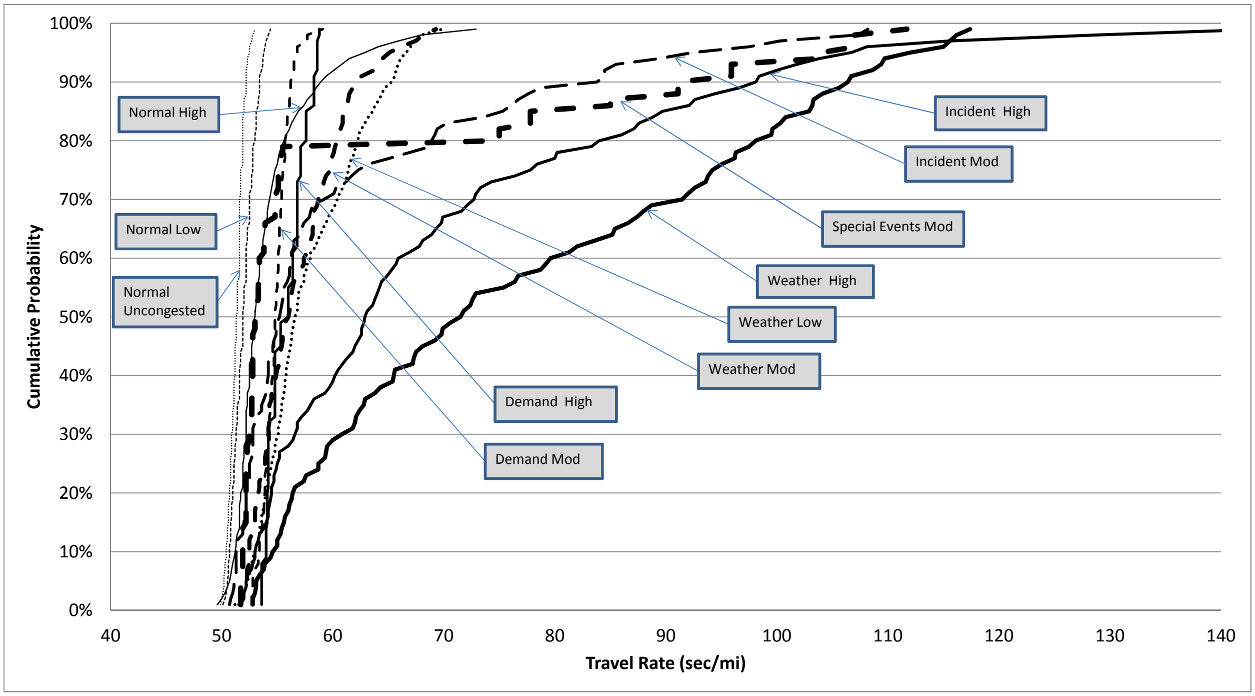


Figure 8-3: How Travel Rates Are Affected by Congestion and Non-Recurring Incidents

The value in the results come from comparing one distribution with another. For example, analysts can compare the distribution for 1) high recurrent congestion *and* traffic incidents with 2) high recurrent congestion without incidents. Without incidents, 50% of the vehicles are traveling at 58 sec/mi instead of 70 sec/mi—considerably faster. And at the 90th percentile, the difference is even more dramatic: 65 sec/mi versus 110 sec/mi.

Not only does the exhibit indicate that the difference between the two conditions is dramatic, but it also suggests that taking actions to mitigate these impacts would produce significant benefits in terms of improving reliability. The mitigating actions would be intended to cause the travel times (or travel rates) during incidents to get much closer to those when there are no incidents. Moreover, after the mitigating actions were taken, the TTRMS would be able to show how reliability improved.

In conclusion, a TTRMS will help an agency understand the reliability performance of their systems and monitor how reliability improves over time:

- What is the distribution of travel times in their system?
- How is the distribution affected by recurrent congestion and non-recurring events?
- How are freeways and arterials performing relative to performance targets set by the agency?
- Are capacity investments and other improvements really necessary given the current distribution of travel times?

- Are operational improvement actions and capacity investments improving the travel times and their reliability?

REFERENCES

Arezoumandi, M., Bham, G.H. 2011. Use of median travel time as measure of central tendency. Integrated Transportation and Development for a Better Tomorrow - Proceedings of the 1st Congress of the Transportation and Development Institute of ASCE, pp. 59-68.

Barkley, T, Hranac, R, and Petty, K. 2012. Using Multistate Models to Relate Travel Time Reliability and Non-recurrent Congestion. *Proceedings of the 91st Annual Meeting of the Transportation Research Board*, Washington, DC.

Bates, J., Polak, J., Jones, P. and Cook, A. 2001. The valuation of reliability for personal travel, *Transportation Research, Part E* , Vol. 37, pp. 191–229.

Batley, R. and Ibanez, J. N. 2009. Randomness in preferences, outcomes and tastes; an application to journey time risk, *Proceedings of the International Choice Modelling Conference*, Harrogate, UK.

Berkow, M, Wolfe, M, Monsere, C, and Bertini, R. 2008. Using signal system data and buses as probe vehicles to define the congested regime on arterials *Proceedings of the 87th Annual Meeting of the Transportation Research Board*, Washington, DC.

Bertini, R L, and Ahmed El-Geneidy, M. 2003. Generating Transit Performance Measures with Archived Data.

Bertini, R L, and Ahmed El-Geneidy, M. 2004. Modeling Transit Trip Time Using Archived Bus Dispatch System Data.

Briggs, V., Walton, C.M. 2000. The Implications of Privacy Issues for Intelligent Transportation Systems (ITS) Data. Research Report 472840-00075, Center for Transportation Research.

Bertini, R L, Cameron, GJ, and Peters, J. 2005. Evaluating traffic signal improvements using archived transit AVL data *ITE Journal*, Vol.75, No.2, pp. 69–75.

Bo, L., Hiroaki, M. 2008. Discussion of traffic signal effect on calculating link travel time and field test evaluation. Proceedings 8th International Conference on Intelligent Transport System Telecommunications, pp. 112-115.

Byon, Y-J, Shalaby, A, and Abdulhai, B. 2006. Travel Time Collection and Traffic Monitoring via GPS Technologies *Proceedings of the IEEE Intelligent Transportation Systems Conference*

Cambridge Systematics, Texas Transportation Institute, University of Washington, and Dowling Associates. 2003. *Providing a Highway System with Reliable Travel Times Prepared for the Future*. Strategic Highway Research Program NCHRP Project 20-58, Transportation Research Board, Washington, DC.

Cambridge Systematics and Texas Transportation Institute. 2005. *Traffic Congestion and Reliability: Trends and Advanced Strategies for Congestion Mitigation*. Prepared for the FHWA Office of Operations, Washington, DC.

Carrion, C. and Levinson, D. 2010. Value of reliability: High occupancy toll lanes, general purpose lanes, and arterials *Proceedings of 4th International Symposiumon Tranportation Network Reliability*, Minneapolis, MN, USA.

Carrion, C. and Levinson, D. 2011. A model of bridge choice across the mississippi river in minneapolis, Presented at the 90th Annual Transportation Research Board Conference, Washington, DC.

Cetin M., List, G.F., and Zhou, Y. 2005. Factors affecting minimum number of probes required for reliable estimation of travel time. *Transportation Research Record*, Vol. 1917, pp. 37-44.

Chakroborty, P., and Kikuchi, S. 2004. Using bus travel time data to estimate travel times on urban corridors. *Transportation Research Record*, Vol.1870, pp. 8-25.

Chang, X. and Stopher, P. 1981. Defining the perceived attributes of travel modes for urban work trips, *Transportation Planning and Technology*, Vol. 7, pp. 55–65.

Chen, C., J. Kwon, A. Skabardonis, and P. Varaiya. 2003. Detecting Errors and Imputing Missing Data for Single Loop Surveillance Systems. *Transportation Research Record*, Vol. 1855, pp. 160-167.

De Fabritiis, C., Ragona, R., Valenti, G. 2008. Traffic estimation and prediction based on real time floating car data. Proceedings of IEEE Conference on Intelligent Transportation Systems, pp. 197-203.

Demers, A., List, G.F., Wallace, W.A., Lee, E.E., and Wojtowicz, J. 2006. Probes as path seekers: a new paradigm. *Transportation Research Record*, Vol. 1944, pp. 107-114.

Demers A., List, G.F., Wojtowicz, W., Kornhauser, A., Wallace, W.A., Lee, E.E., and Salasznyk, P. 2006. Experimenting with real-time ATIS: Stepping forward from ADVANCE. *Proceedings of the 9th International Conference on Applications of Advanced Technology in Transportation*.

Dhaene, J., Denuit, M., Goovaerts, M. J., Kaas, R., and Vyncke, D. 2002a. The concept of comonotonicity in actuarial science and finance: theory. *Insurance: Mathematics and Economics*, Vol. 31, No.1, pp. 3-33.

Dhaene, J & Denuit, M & Goovaerts, M J & Kaas, R & Vyncke, D. 2002b. The concept of comonotonicity in actuarial science and finance: applications, *Insurance: Mathematics and Economics*, Vol. 31, No.2, pp. 133-161.

Dion F., and Rakha, H. 2006. Estimating dynamic roadway travel times using automatic vehicle identification data for low sampling rates. *Transportation Research - Part B*, Vol. 40, No.9, pp. 745–766.

Dong, J., Mahmassani, H.S. 2011. Stochastic modeling of traffic flow breakdown phenomenon: Application to predicting travel time reliability. Proceedings of 14th International IEEE Conference on Intelligent Transportation Systems, pp. 2112-2117.

Enam, E. and Al-Deek, H. 2006. Using Real-Life Dual-Loop Detector Data to Develop New Methodology for Estimating Freeway Travel Time Reliability. *Transportation Research Record*, Vol. 1959, pp. 140-150.

Ernst, J.M., Day, C. M., and Krogmeier, J.V. 2012. Probe Data Sampling Guidelines for Characterizing Arterial Travel Time. *Proceedings of the 91st Annual Meeting of the Transportation Research Board*, Washington, DC.

Federal Highway Administration. 2005. Traffic Congestion and Reliability: Trends and Advanced Strategies for Congestion Mitigation. September 2005.

Federal Highway Administration. 2008. Travel time reliability: Making it there on time, all the time. Available on-line at ops.fhwa.dot.gov/publications/tt_reliability/TTR_Report.htm#WhatisTTR.

Fosgerau, M. and Karlstrom, A. 2010. The value of reliability, *Transportation Research Part B*, Vol. 44, pp. 38–49.

Fosgerau, M. and Engelson, L. 2011. The value of travel time variance, *Transportation Research Part B*, Vol. 45, pp. 1–8.

Fraley C. and Raftery A.E. 2009. MCLUST Version 3 for R: Normal Mixture Modeling and Model-Based Clustering. Technical Report No. 504. Department of Statistics, University of Washington. http://www.stat.washington.edu/fraley/mclust/tr504.pdf.

Feng, Y., Davis, G.A., Hourdos, J. 2011. Arterial Travel Time Characterization and Real Time Traffic Condition Identification Using GPS-equipped Probe Vehicles. *Proceedings of the 90th Annual Transportation Research Board Meeting*, Washington, DC

Feng, Y., Hourdos, J., and Davis, G.A. 2012. A Bayesian Model for Constructing Arterial Travel Time Distributions using GPS Probe Vehicles. *Proceedings of the 91st Annual Meeting of the Transportation Research Board*, Washington, DC.

Fontaine, M.D., and Smith, B.L. 2005. Probe-based traffic monitoring systems with wireless location technology. *Transportation Research Record*, Vol. 1925, pp. 3–11.

Gaver, D. 1968. Headstart strategies for combating congestion, *Transportation Science*, Vol. 2, pp. 172–181.

Guo, F., Rakha, H., and Park, S. 2010. A Multi-State Travel Time Reliability Model. *Transportation Research Record*, Vol. 2188, pp. 46-54.

Guo, F., Li, Q. and Rakha, H. 2012. Multi-state Travel Time Reliability Models with Skewed Component Distributions. *Proceedings of the 91st Annual Meeting of the Transportation Research Board*, Washington, DC.

Haas, R., M. Carter, E. Perry, J. Trombly, E. Bedsole, and R. Margiotta. iFlorida. 2009. Model Deployment Evaluation Report. Prepared for the USDOT. Report No FHWA-HOP-08-050. January.

Haghani, A., M. Hamedi, K.F. Sadabadi, S. Young, and P. Tarnoff. 2010. Data Collection of Freeway Travel Time Ground Truth with Bluetooth Sensors. In Transportation Research Record: Journal of the Transportation Research Board, No. 2160, Transportation Research Board of the National Academies, Washington, D.D., pp. 60-68.

Hainen, A.M., Remias, S.M., Brennan, T.M., Day, C.M., Bullock, D.M. 2012. Probe vehicle data for characterizing road conditions associated with inclement weather to improve road maintenance decisions. Proceedings of IEEE Intelligent Vehicles Symposium, pp. 730-735.

De Fabritiis, C., Ragona, R., Valenti, G. 2008. Traffic estimation and prediction based on real

Hall, R., and Nilesh, V. 2000. Buses as a traffic probe demonstration project. *Transportation Research Record,* Vol. 1731, pp. 96-103.

Haseman, R. J., Wasson, J. S., Bullock, D. M. 2010. Real-time measurement of travel time delay in work zones and evaluation metrics using bluetooth probe tracking. Transportation Research Record, Vol. 2169, pp. 40-53.

Hellinga, B.R., and Fu, L. 2002. Reducing bias in probe-based arterial link travel time estimates. *Transportation Research Part C,* Vol. 10, No.4, pp. 257-273.

Hesham R., Ihab, E-S., Mazen, A., and Dion, F. 2006. Estimating Path Travel-Time Reliability. *Intelligent Transportation Systems Conference Proceedings*, pp. 236-241.

Higatani, A., Kitazawa, T., Tanabe, J., Suga, Y., Sekhar, R., Asakura, Y. 2009. Empirical analysis of travel time reliability measures in hanshin expressway network. Journal of Intelligent Transportation Systems: Technology, Planning, and Operations, Vol. 13, No. 1, pp. 28-38.

Hoeitner, A., Herringy, R., Bayenz, A., Hanx, Y., Moutardex, F., and de La Fortellex, A. 2012. Large scale estimation of arterial traffic and structural analysis of traffic patterns using probe vehicles. *Proceedings of the 91st Annual Meeting of the Transportation Research Board*, Washington, DC.

Ishak, S., Kondagari, S., Alecsandru, C. 2007. Probabilistic data-driven approach for real-time screening of freeway traffic data. Transportation Research Record, Vol. 2012, pp. 94-104.

Jackson,W. and Jucker, J. 1982. An empirical study of travel time variability and travel choice behavior, *Transportation Science*, Vol. 16, pp. 460–475.

Jenelius, E., Mattsson, L. G. and Levinson, D. 2011. Traveler delay costs and value of time with trip chains, flexible activity scheduling and information, *Transportation Research Part B,* Vol. 45, pp. 789–807.

Jintanakul, K., Chu, L., and Jayakrishnan, R. 2009. Bayesian Mixture Model for Estimating Freeway Travel Time Distributions from Small Probe Samples from Multiple Days, *Transportation Research Record*, Vol. 2136, pp. 37-44.

Kaparias, I., Bell, Michael G. H., Belzner, H. 2008. A new measure of travel time reliability for in-vehicle navigation systems. Journal of Intelligent Transportation Systems: Technology, Planning, and Operations, Vol. 12, No. 4, pp. 202-211.

Karr, A. F. 1993. Probability. Springer–Verlag, New York.

Karr, A. F., Sanil, A. P., and Banks, D. L. 2006. Data quality: A statistical perspective. Statistical Methodology, Vol.3, No.2, pp. 137–173.

Karr, A. F., Fulp, W. J., Lin, X., Reiter, J. P., Vera, F., and Young. S. S. 2007. Secure, privacy-preserving analysis of distributed databases. *Technometrics*, Vol.49, No.3, pp. 335–345.

Khattak, A., Youngbin Y., and Linda S. 2003. Willingness to pay for travel information, *Transportation*

Research-Part C, Vol. 11, No. 2, Pergamon Press, pp. 137-159.

Khattak, A. J., Fan, Y., & Teague, C. (2008). Economie impact of traffic incidents on businesses. *Transportation Research Record*, *2067*, pp. 93-100.

Khattak, A. J., Joseph L. S., and Koppelman, F. S. 1994. Effect of Traffic Information on Commuters' Propensity to Change Route and Departure Time, *Journal of Advanced Transportation*, Vol. 29, No. 2, pp. 193-212.

Kwon, J., Coifman, B.,and Bickel, P. 2000. Day-to-day travel-time trends and travel-time prediction from loop-detector data. *Transportation Research Record 1717*, pp. 120-129.

Kwon, J., Petty, K., and Varaiya, P. 2007. Probe Vehicle Runs or Loop Detectors? Effect of Detector Spacing and Sample Size on Accuracy of Freeway Congestion Monitoring, *Transportation Research Record 2012*, pp. 57-63.

Leng, J., Zhang, Y., Leng, Y. 2009. Assessment methodology for road network travel time reliability under ice and snowfall conditions. Proceedings of the 9th International Conference of Chinese Transportation Professionals, Vol. 358, pp. 1124-1130.

Li, R., Rose, G., and Sarvi, M. 2006. Using automatic vehicle identification data to gain insight into travel time variability and its causes. *Transportation Research Record*, Vol. 1945, pp. 24-32.

Lin, W., Kulkarni, A. and Mirchandani, P. 2003. Arterial travel time estimation for advanced traveler information systems. *Proceedings of the 82nd Annual Meeting of the Transportation Research Board*. Washington, DC.

List, G.F., Wallace, W.A., Demers, A., Salasznyk, P., Lee, E.E., and Wojtowicz, J. 2005. Field experiment with a wireless GPS-based ATIS system. *Proceedings of the 12th World Congress on ITS*, San Francisco, CA.

List, G.F., Demers, A., Wallace, W.A., Lee, E.E., and Wojtowicz, J. 2005. ATIS via wireless probes: smart vehicles for smart travelers. *INFORMS Annual Meeting*.

List, G.F., and Demers, A. 2006. Estimating highway facility performance from AVL data. *Proceedings of the Fifth International Symposium on Highway Capacity and Quality of Service*, pp. 319-328, Osaka, Japan.

Liu, K., Yamamoto, T., Morikawa, T. 2007. Feasibilities and challenges of probe technologies for real-time traffic data collection. Proceedings of the 7th International Conference of Chinese Transportation Professionals Congress, pp. 328-340.

Liu, H., Sang, L., Zhang, K., Yuan, Y. 2010. An evaluation of obtaining travel time observations via new technology a case study in Tianjin. Proceedings of the 10th International Conference of Chinese Transportation Professionals, Vol. 382, pp. 2380-2388.

Lomax, T., Schrank, D., Turner, S., and Margiotta, R. 2003. *Selecting Travel Reliability Measures*. http://ttitamuedu/documents/474360-1pdf.

Lyman, K., Bertini, R.L. 2008. Using travel time reliability measures to improve regional transportation planning and operations. Transportation Research Record, Vol. 2046, pp. 1-10.

Ma, X., and Koutsopoulos, H.N. 2010. Estimation of the automatic vehicle identification based spatial travel time information collected in Stockholm. *IET Intelligent Transport Systems*, Vol. 4, No. 4, pp. 298-306.

Ma, Y., Chowdhury, M., Sadek, A., Jeihani, M. 2009. Real-time highway traffic condition assessment framework using vehicleInfrastructure integration (VII) with artificial intelligence (AI). IEEE Transactions on Intelligent Transportation Systems, Vol. 10, No. 4, pp. 615-627.

Margiotta, R. 2010. *Analytic Procedures for Determining the Impacts of Reliability Mitigation Strategies*. SHRP-2 L03 Final Report Transportation Research Board, Washington, DC.

Margiotta, R., Lomax, T., Hallenbeck, M., Turner, S., Skabardonis, A., Ferrell, C., and Eisele, B. 2006. *Guide to Effective Freeway Performance Measurement*, Final Report, NCHRP Project 3-68, Transportation Research Board, Washington, DC.

Martchouk, M., Mannering, F., Bullock, D. 2011. Analysis of Freeway Travel Time Variability Using Bluetooth Detection. Journal of Transportation Engineering, Vol. 137, No. 10, pp. 697-704.

Morris A.G., Kornhauser, A.L., and Kay, M.J. 1998. Urban freight mobility: collection of data on time, costs, and barriers related to moving product into the central business district. *Transportation Research Record*, Vol. 1613, pp. 27-32.

National Institute of Statistical Sciences. 2004. Data Confidentiality, Data Quality and Data Integration for Federal Databases: Foundations to Software Prototypes, Available on-line at www.niss.org/dgii.

Noland, R. and Small, K. 1995. Travel-time uncertainty, departure time choice, and the cost of morning commutes, *Transportation Research Record*, Vol. 1493, pp. 150–158.

Pan, C., Lu, J., Wang, D., and Ran, B. 2007. Data collection based on global positioning system for travel time and delay for arterial roadway network. *Transportation Research Record*, Vol. 2024, pp. 35-43.

Park, M., Kim, S., Park, C., Chon, K. 2007. Transportation network design considering travel time reliability. Proceedings of the 10th International IEEE Conference on Intelligent Transportation Systems, pp. 496-502.

Prashker, J. 1979. Direct analysis of the perceived importance of attributes of reliability of travel modes in urban travel, *Transportation*, Vol. 8, pp. 329–346.

Quiroga C.A., and Bullock, D. 1998. Travel time studies with global positioning and geographic information systems: an integrated methodology. *Transportation Research-Part C*, Vol. 6, No.1/2, 101–127.

Pu, W. 2011. Analytic Relationships between Travel Time Reliability Measures. *Transportation Research Record*, Vol. 2254, pp. 122-130.

Ramezani, M., and Geroliminis, N. 2012. Estimation of arterial route travel time distribution with Markov chains. *Proceedings of the 91st Annual Meeting of the Transportation Research Board*, Washington, DC.

Rice, J., and Van Zwet, E. 2004. A simple and effective method for predicting travel times on freeways. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 5, No. 3, pp. 200-207.

Rosenblatt, M. 1956. Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics*, Vol. 27, pp. 832–837.

Shen, L., Hadi, M. 2012. Practical approach for travel time estimation from point traffic detector data. Journal of Advanced Transportation.

Silverman, B.W. 1986. *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, London.

Small, K. 1982. The scheduling of consumer activities: Work trips, American *Economic Review* , Vol. 72, The American Economic Association, pp. 467–479.

Small, K., Winston, C. and Yan, J. 2005. Uncovering the distribution of motorists' preferences for travel time and reliability, *Econometrica*, Vol. 73, pp. 1367–1382.

Small, K. and Verhoef, E. 2007. The Economics of Urban Transportation, Routledge, part of the Taylor & Francis Group.

Soriguera, F., Thorson, L. 2007. Travel time measurement using toll infrastructure. Transportation Research Record, Vol. 2027, pp. 99-107.

Soriguera, F. 2011. Highway travel time accurate measurement and short-term prediction using multiple data sources. Transportmetrica, Vol. 7, No. 1, pp. 85-109.

Sun, H., Gao, Z. 2012. Stochastic traffic equilibrium based on travel time robust reliability. Journal of Transportation Systems Engineering and Information Technology, Vol. 12, No. 2, pp. 76-84.

Susilawati S., Taylor, M.A.P., and Somenahalli, S.V.C. 2011. Distributions of travel time variability on urban roads. *Journal of Advanced Transportation*, doi: 101002/atr192.

Tilahun, N. and Levinson, D. 2010. Amoment of time: Reliability in route choice using stated preference, *Journal of Intelligent Transportation Systems*, Vol. 14, pp. 179 –187.

Transportation Research Center. 2009: *Improving Reliability on Surface Transportation Networks: Summary Document*, OECD

Transportation Research Center. 2010: *Improving Reliability on Surface Transportation Networks*, OECD

Tseng, Y. and Verhoef, E. 2008. Value of time by time of day: a stated-preference study, *Transportation Research Part B*, Vol. 42, p. 607–618.

Tsubota, T., Kikuchi, H., Uchiumi, K., Warita, H., Kurauchi, F. 2011. Benefit of Accident Reduction Considering the Improvement of Travel Time Reliability. International Journal of Intelligent Transportation Systems Research, Vol. 9, No. 2, pp. 64-70.

Tu, H., Van Lint, H., Van Zuylen, H. 2008. The effects of traffic accidents on travel time reliability. Proceedings of IEEE Conference on Intelligent Transportation Systems, pp. 79-84.

Uno, N., Kurauchi, F., Tamura, H., Iida, Y. 2009. Using bus probe data for Aanalysis of travel time variability. Journal of Intelligent Transportation Systems: Technology, Planning, and Operations, Vol. 13, No. 1, pp. 2-15.

Vanajakshi, L., Subramanian, S.C.; Sivanandan, R. 2009. Travel time prediction under heterogeneous traffic conditions using global positioning system data from buses. IET Intelligent Transport Systems, Vol. 3, No. 1, pp. 1-9.

Van Hinsbergen, C.P.I.J., Van Lint, J.W.C. 2008. Bayesian combination of travel time prediction models. Transportation Research Record, Vol. 2064, pp. 73-80.

Van Zwet, E., Chen, C., Jia, Z., and Kwon, J. 2003. A statistical method for estimating speed from single loop detectors, Freeway Performance Measurement System (PeMS).

Van Lint, J.W.C., and van Zuylen, H.J. 2005. Monitoring and Predicting Freeway Travel Time Reliability Using Width and Skew of Day-to-Day Travel Time Distribution. *Transportation Research Record*, Vol. 1917, pp. 54-62.

Van Lint J.W.C., van Zuylen H.J., and Tu, H. 2008. Travel time unreliability on freeways: Why measures based on variance tell only half the story. *Transportation Research Part A*, Vol. 42, No. 1, pp. 258-277.

Vaziri, M. and Lam., T. N. 1983. Perceived factors affecting driver route decisions, *Journal of Transportation Engineering*, Vol. 109, p. 297–311.

Vickrey,W. 1969. Congestion theory and transport investment, *The American Economic Review*, Vol. 59, pp. 251–260.

Wang, J., Zou, N., Chang, G. 2008. Travel time prediction: Empirical analysis of missing data issues for advanced traveler information system applications. Transportation Research Record, Vol. 2049, pp. 81-91.

Wang, J., Sun, G., Hu, X. 2009. Analysis of city transportation networks' travel time reliability during adverse weather. Proceedings of the 9th International Conference of Chinese Transportation Professionals, Vol. 358, pp. 1536-1542.

Wang, J., He, J., Wu, L. 2011. Evaluating approach of travel time reliability for highway network under rain environment. Journal of Transportation Systems Engineering and Information Technology, Vol. 11, No. 6, pp. 117-123.

Wasson J.S., Sturdevant, J.R., and Bullock, D.M. 2008. Real-time travel time estimates using media access control address matching. *ITE Journal*, Vol. 78, No. 6, pp. 20-23.

Wojtowicz J., Murrugarra, R.I., Bertoli, B., Wallace, W.A., Manuel, P., He, W., and Body, C. 2008. RFID technology for AVI: field demonstration of a wireless solar powered E-ZPass tag reader. *Proceedings of the 15th World Congress on Intelligent Transport Systems*.

Wosyka, J., Pribyl, P. 2012. Real-time travel time estimation on highways using loop detector data and license plate recognition. Proceedings of 9th International Conference, ELEKTRO, pp. 391-394.

Xiong, Z., Shao, C., Yao, Z. 2007. The framework of assessment on travel time reliability. Proceedings of International Conference on Transportation Engineering, pp. 223-228.

Xiaoliang, M., Koutsopoulos, H.N. 2008. A new online travel time estimation approach using distorted automatic vehicle identification data. Proceedings of IEEE Conference on Intelligent Transportation Systems, pp. 204-209.

Yamamoto, T., Liu, K., and Morikawa, T. 2006. Variability of travel time estimates using probe vehicle data. Proceedings of the Fourth International Conference on Traffic and Transportation Studies, pp. 278-287.

Yan, Y., Guo, X., Li, Y., Kong, Z., He, M. 2012. Bus transit travel time reliability evaluation based on automatic vehicle location data. Journal of Southeast University, Vol. 28, No. 1, pp. 100-105.

Yang, M., Liu, Y., You, Z. 2010. The reliability of travel time forecasting. IEEE Transactions on Intelligent Transportation Systems, Vol. 11, No. 1, pp. 162-171.

Yang, Y., Yao, E., Qu, D., Zhang, Y. 2011. Study on travel time reliability of probe vehicle system based on minimum sample size analysis. Multimodal Approach to Sustained Transportation System Development - Information, Technology, Implementation - Proceedings of the 1st International. Conference. on Transportation Information and Safety, pp. 1680-1686.

Yamazaki, H., Uno, N.; Kurauchi, F. 2012. The effect of a new intercity expressway based on travel time reliability using electronic toll collection data. IET Intelligent Transportation Systems, Vol. 6, No. 3, pp. 306-317.

Zou, N., Wang, J., Chang, G. 2008. A reliable hybrid prediction model for real-time travel time prediction with widely spaced detectors. Proceedings of IEEE Conference on Intelligent Transportation Systems, pp. 91-96.

Zou, N., Wang, J., Chang, G., Paracha, J. 2009. Application of advanced traffic information systems: Field test of a travel-time prediction system with widely spaced detectors. Transportation Research Record, Vol. 2129, pp. 62-72.