September, 2017

SHRP 2 Implementation Assistance Program- Arizona

Implementation of L02 and L07 Design Guides

Reza Karimvand, P.E. Vahid Nikou Goftar, P.E.

ADOT TSMO
2302 W Durango Street
Phoenix, AZ 85009

Prepared by the **University of Arizona**

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

The Arizona Department of Transportation implemented two SHRP 2 products on three study corridors along Interstate 10. A prototype travel time reliability monitoring system (TTRMS) was established, following the L02 design guide, to evaluate travel time reliability for the study corridors using data sources collected throughout Arizona. Areas of interest were identified and countermeasures evaluated using the L07 spreadsheet tool. The website created utilizing the data collected in this project displays performance measures across the state of Arizona. Recommendations based on the project's finding include the following:

- Additional data should be collected, including weather and special event data;
- The L02 published guidelines could be more specific on data sources and how to combine them;
- The TTRMS established was beneficial for understanding travel time reliability along the corridors and should be implemented statewide;
- Although the L07 design guide provided useful information on appropriate countermeasures for different observed issues, the spreadsheet tool was not user friendly and should be improved;
- A dedicated data server should be established to facilitate the long-term implementation of these products statewide;
- To achieve the system's full potential, the use of additional performance measures calculated using the data collected in the project should be explored;
- A dedicated website could provide a user-friendly way to display the results of the TTRMS.

To implement the L02 design guide, extensive data collection efforts were required. AZTech's Regional Archived Data Server (RADS) provided the 20-second loop detector data, the FHWA's National Performance Management Research Data Set (NPMRDS) the probe vehicle travel times, the Highway Condition Reporting System (HCRS) the freeway closure data, the Accident Location Identification Surveillance System (ALISS) the crash data recorded by the police, and ADOT's Multimodal Planning Division the vehicle volumes and classifications. These datasets were collected and stored on a single server housed at the University of Arizona for this project.

Travel time data were available in 5-minute intervals directly from the NPMRDS dataset and could be calculated using 20-second loop detector data from RADS. The NPMRDS dataset is substantially easier to use, since the travel times have already been calculated, but the algorithm is proprietary so ADOT must take the results on trust and cannot adjust the algorithm. Access to the dataset is also subject to a recurring subscription cost, although the historical version is paid for by FHWA. Using the RADS dataset is substantially more complex, however, as it requires a complex algorithm to obtain travel times using point speeds. This algorithm is computationally expensive and complex to implement but gives ADOT more control over the sensors collecting the data and the algorithm used to estimate travel times using point-based speeds. Comparing the two, the difference between the 5-minute travel times obtained from the RADS loop data and the NPMRDS data was significantly different. Further study is thus required to determine which dataset yields the most accurate results.

The NPMRDS data was used to implement the L02 design guide and establish a TTRMS for the three study corridors, all of which were located in urban areas along I-10. The TTRMS components consist of the server, currently housed at the University of Arizona, code utilizing R and Microsoft SQL to retrieve data and make calculations, and outputs including reliability metrics such as a Travel Time Index, Misery Index, and Buffer Index as well as visual representations of Cumulative Density Functions (CDFs). The reliability metrics were calculated for each direction on the three study corridors, with the worst metrics being recorded for I-10 westbound between 35th Avenue and Sky Harbor Boulevard in the Phoenix metropolitan area. The CDFs showed peak hour traffic in the afternoon and traffic incidents were the primary causes of these poor reliability metrics.

Focusing on the study corridor with the worst performance on the TTRMS, the L07 design guide was applied to identify possible countermeasures to address the congestion observed along I-10 westbound between 35th Avenue and Sky Harbor Boulevard. Although the use of the spreadsheet tool was found to be difficult and not particularly user friendly for a variety of reasons the results allowed users to quickly identify possible alternatives to ameliorate recurring problems. This indicates that if the tool were implemented in a more user-friendly way, with improved directions, it could serve as a valuable tool to support the engineering process by identifying possible alternatives for further evaluation.

Finally, the data collected for the project was used to create a website that displays performance measures statewide (www.adottsmopm.org). The website is currently a prototype but has already demonstrated its utility as an easy-to-use reporting system for TTRMS results as well as additional performance measures covering mobility, safety, freight, and asset management.

1 INTRODUCTION

1.1 Background

The Second Strategic Highway Research Program (SHRP 2) has devoted significant resources to the development of guidance in four focus areas: Safety, Renewal, Reliability, and Capacity. The Implementation Assistance Program (IAP) is meant to assist state and local governments by deploying research findings and applications from the projects funded by SHRP 2. The objective of the SHRP 2 IAP project discussed in this report was to address reliability on I-10 in the state of Arizona by examining travel time reliability and both recurrent and nonrecurrent congestion.

In 2014, SHRP 2 published two documents designed to guide efforts to develop performance metrics for reliability on the nation's roadways:

1. Guide to Establishing Monitoring Programs for Travel Time Reliability (L02)
2. Design Guide for Addressing Nonrecurrent Congestion (L07)

L02 defines the essential components of a travel time reliability monitoring system (TTRMS) including why a TTRMS is necessary, what data sources are needed to establish a state of the art system, the components of a TTRMS, how to utilize the raw data to arrive at the intended product, and how to make decisions using the TTRMS that will improve travel time reliability across a region. L07 provides agencies with a collection of potential countermeasures that can be used to address nonrecurrent congestion, along with a spreadsheet-based tool that agencies can utilize to evaluate potential countermeasures. The two design guides are highly interrelated, as results from the TTRMS can provide agencies with the information needed to select the most effective countermeasures for their specific issues.

1.2 Study Corridors

Figure 1.1 shows the three segments of Interstate 10 in the State of Arizona used to demonstrate the implementation of L02 and L07 in this project. The segments are:

1. I-10 between US Route 60 and I-17 in Phoenix
2. I-10 between 35th Avenue and Sky Harbor Boulevard in Phoenix
3. I-10 between Ina Road and I-19 in Tucson

Figure 1.1: Study Corridors

The Phoenix metropolitan area is located in central Arizona and is the 11th largest urban area in the United States, with a population of almost 4 million. In 2014, an estimated 155,730,000 person-hours of delay were experienced in the city and as much as 36% of peak hour vehicle-miles travelled (VMT) were driven under congested conditions (Texas Transportation Institute, 2014a). The climate in the area is relatively dry, with annual rainfall of 6-9 inches for the past 3 years.

I-10 runs through southern, central, and western Phoenix. The two study segments in the city have four general purpose lanes and an HOV lane running in each direction. The segment between US Route 60 and I-17 is just over 5 miles long and the segment between 35th Avenue and Sky Harbor Boulevard is about 7 miles long.

The Tucson metropolitan area is located in southern Arizona. In 2014, the urban population was 865,000, ranking 52nd in the United States. An estimated 35,993,000 person-hours of annual delay was experienced in 2014 and 28% of peak hour VMT experiences congested conditions (Texas Transportation Institute, 2014b). The study segment of I-10 between I-19 and Ina Road runs north-south through downtown Tucson and is roughly 11 miles long. There are four general travel lanes in each direction.

2 LITERATURE REVIEW

2.1 Travel Time Reliability and L02

Travel time and travel time reliability are two of the most commonly used metrics for performance in transportation. While the concept of travel time is relatively simple, travel time reliability is more difficult to quantify in terms of definition, visualization, and how it impacts decision making. L02 states that “a system is reliable if each traveler or shipper experiences actual time of arrival (ATA) that match the desired time of arrival (DTA) within some window”. A variety of metrics to quantify and visualize travel time reliability have been proposed. Probability density functions (PDFs) and cumulative density functions (CDFs) are suggested by L02 for visualization and comparisons as they are simple to create and understand, especially for predefined links (List et al., 2014a). Measures associated with PDFs and CDFs such as mean, standard deviation, and variance can also be used to describe reliability.

Metrics to quantify travel time reliability on a situational basis can be valuable. Common metrics for travel time include (List et al., 2014b; Lomax, Schrank, Turner, & Margiotta, 2003):

- Variability Index
- Buffer Index
- Planning Time Index
- Skewness
- Misery Index

These measures all result in different values and provide answers to different questions from different points of view, which can lead to some confusion regarding which metric is best for a particular situation. Some metrics, such as variance-based metrics, are unable to account for certain aspects or reliability, while others may provide valuable insights (Van Lint, Zuylen, & Tu, 2008). Several novel methods to quantify reliability have been proposed that apply more advanced techniques (Ma, Koutsopoulos, Ferreira, & Mesbah, 2017; Uchida, 2014; Yang, Malik, & Wu, 2014; Yang & Wu, 2016), but although these methods often show improved travel time reliability estimates, they are often hard to use in practice due to their complexity.

Reliability can be used to address questions regarding the value travelers place on reliable travel times (Alemazkoor, Burris, & Danda, 2012; Bates, Polak, Jones, & Cook, 2001; Carrion & Levinson, 2012) and to evaluate roadway performance under different conditions (Hojati, Ferreira, Washington, & Charles, 2016; Wright, Zou, & Wang, 2015), but selection of the correct metric is critical. In research applications, travel time reliability is a useful metric for measuring performance and the impact of proposed changes to existing systems and policies. Zheng et al. considered travel time reliability in their proposed strategy for optimizing urban signal controls (Zheng, van Zuylen, Liu, & Vine, 2016), while Li et al. used travel time and travel time reliability as metrics for evaluating the value of different timing plans for a signalized corridor in Indiana (Li, Lavrenz, Day, Stevens, & Bullock, 2013). The case studies presented in Section 2.3 provide additional examples of studies that used travel time reliability for the evaluation of systems and infrastructure.

2.2 Nonrecurrent Congestion and L07

Nonrecurrent congestion is one of the primary causes of unreliable travel times. L07 identifies six primary causes of nonrecurrent congestion: Incidents, Weather, Special Events, Work Zones, Demand Fluctuations, and Traffic Control Devices. Each will have a different effect on the nature of the congestion, so understanding the most frequent cause of nonrecurrent congestion is critical for determining the correct countermeasure or set of countermeasures to deploy (Potts, Harwood, Fees, Hutton, & Kenzel, 2014.).

The impact of different sources of nonrecurrent congestion on operations and safety, ways to detect the root cause of nonrecurrent congestion, and how best to implement countermeasures to address those sources have all been studied extensively. For example, Xu et al. and Yeo et al. evaluated crash risk and crash involvement, respectively, under different traffic states on freeways and found congested flow to be the most unsafe (Xu, Liu, Wang, & Li, 2012; Yeo, Jang, Skabardonis, & Kang, 2013). Dailey correlated the rate of rainfall to speed and created a function that predicts speed reduction based on Doppler radar (Dailey, 2006). Kwon et al. provide a method that divides congestion into six components, namely incidents, special events, weather, lane closures, potential reduction by implementing ramp metering, and demand (Kwon, Mauch, & Varaiya, 2006), while Liu et al suggest a strategy for optimizing diversion rates during incident management periods (Liu, Chang, & Yu, 2011). The *Highway Safety Manual* is just one of many resources that suggest safety-specific countermeasures to reduce crashes based on crash type and facility type (AASHTO, 2010).

2.3 Case Studies using L02 and L07

L02 and L07 were both published in 2014 and have since been implemented in some manner by a variety of state and local governments. L02 itself includes five case studies – San Diego, Northern Virginia, Sacramento-Lake Tahoe, Atlanta, and New York/New Jersey, highlighting specific aspects of different states' TTRMSs and the investigations conducted by each system.

San Diego, California, was used to showcase an established, state of the art freeway sensor network with a corresponding data warehouse and software system that is used for automatic travel time calculation. San Diego Metropolitan Transit System data is collected through automated vehicle location and computer-aided dispatch. San Diego's primary investigations involved the systematic integration of transit data and tracking all sources of nonrecurrent congestion.

Northern Virginia has an extensive network of loop and radar detectors deployed along the study corridors. In addition to the systems deployed by the state's DOT they acquired data from a variety of third parties. The Northern Virginia case focused on the configuration of a Performance Measurement System, a comparison of probe vehicles to examine travel time data quality, and a study of travel time reliability under different conditions.

The Sacramento-Lake Tahoe case utilized traffic operations data, sensor deployment data, and incident data to evaluate the performance of a primarily rural area between Sacramento and Lake Tahoe in California. The primary goals of their investigation were to evaluate the quality and feasibility of using Bluetooth and Electronic Toll Collection readers for travel time calculations and to understand the effects of weather-related and demand-related conditions on nonrecurrent congestion based on travel times.

Atlanta, Georgia, was used to demonstrate a mixed environment comprising both suburban and urban settings. The study used data generated by video detection systems, radar, and probe vehicles. The research team integrated these data sources and used the results to evaluate sources of nonrecurrent congestion in Atlanta.

The New York/New Jersey case implemented a TTRMS in a densely-populated urban location. Utilizing probe data from a third-party, the research team explored how to construct PDFs directly using Monte Carlo simulations, assuming link speed independence. This case study demonstrated that a TTRMS can indeed be implemented successfully using probe data (List et al., 2014a).

L07 also provides specific examples of cataloged countermeasures already deployed across the United States. The extensive list focuses primarily on success stories recounting where relatively new and emerging countermeasures were being deployed and tested at the time of publication (Potts et al., 2014).

Following the L02 and L07 projects, SHRP 2 conducted four independent pilot studies of reliability analysis projects falling under the designation L38. L38 was first conducted in four locations: Southern California (Williges et al., 2015), Minnesota (Sobolewski, Polum, Morris, Loos, & Anderson, 2015), Florida (Hadi et al., 2015), and Washington (Nisbet et al., 2015). The L38 projects implemented L02, L07 and other travel time reliability products and recommendations.

The Southern California L38 pilot study tested and provided feedback on five SHRP 2 products including L02 and L07. L02 methods were implemented for two specific applications stated in the guide:

1. Examine which factors affect reliability
2. Assess the contributions of these factors

For the implementation of the TTRMS, the process of data processing and aggregation was described in detail, along with a discussion of how some of the ambiguous aspects of portions of the guide were handled. For the L07 tool, the Southern California research team experienced difficulties in acquiring some of the data required as input for the tool. There were also problems with calibrating the results for the region and the tool's predicted travel time index for various treatments was relatively insensitive (i.e. the TTI remained roughly the same under most treatments). However this may actually be an accurate result, since congestion is often so bad that many treatments are ineffective (Williges et al., 2015).

The pilot study in Minnesota also tested both L02 and L07. For the L07 tool, the research team stressed that it should not be applied across the entire system because when it is being applied to a specific site or segment it cannot account for other factors up or downstream, such as bottlenecks. The team suggested using findings from the TTRMS established in L02 to identify target sites for the L07 tool and noted that some default values are more important than others as they affect the outcomes more if changed. The researchers concluded that the model seemed especially sensitive to crashes and incidents (Sobolewski et al., 2015).

The L38C project in Florida expanded on some of the findings presented in L02 by suggesting some additional steps and metrics for monitoring reliability. The results from the Florida team's analysis of L07 showed inconsistent and inaccurate estimations of travel time reliability for existing corridors, but when the team altered the underlying models in the spreadsheet tool the results became more accurate and useable for practical applications (Hadi et al., 2015).

The team in Washington used sensor data from a variety of sources in the Puget Sound region to implement a TTRMS based on L02 methodology and visualization techniques. The L07 tool was also used to evaluate the impact of installing an Extra High Median Barrier but failed to draw any meaningful conclusions, although only one instance was tested (Nisbet et al., 2015).

3 DATA COLLECTION

The implementation of both L02 and L07, is very data-intensive. As ADOT provides several options regarding the data selected for the implementation of the two design guides, data was queried from a variety of sources, described in this chapter, for both 2015 and 2016. These years were chosen as they represent the most recent period for which complete datasets could be acquired at the time of this study. This section discusses the data sources that were acquired and describes why they are needed to implement one or both of L02 and L07.

Given that travel time and travel time reliability are two critical metrics when evaluating freeway operations, both the L02 and L07 design guides spend a significant amount of time discussing travel times. The travel time data used for this IAP project were primarily collected from three major sources: AZTech's Regional Archived Data Server (RADS), the MS2 database managed by ADOT Multimodal Planning Division and the FHWA's National Performance Management Research Data Set (NPMRDS).

Incidents, work zones, and other events that result in lane closures are often responsible for nonrecurrent congestion. Incidents include crashes but can also refer to other issues such as a disabled vehicle that have a similar effect on the flow of traffic. Meanwhile, work zones can be planned for ahead of time to minimize the impacts they have on peak hour traffic flow, although completely eliminating the impact is seldom an option. L07 lists a variety of countermeasures that can make managing incidents and lane closures more feasible. This data can also be used to identify travel times collected under different conditions, as required to implement L02. Data from the Highway Condition Reporting System (HCRS) was acquired for this study.

Crashes are a major source of non-recurring congestion for a variety of reasons. There are two primary ways that crashes can be addressed: through prevention (e.g. countermeasures), and through management (e.g. emergency response). The L07 *Design Guide for Addressing Nonrecurrent Congestion* includes an extensive discussion of both approaches to managing crashes. Crash report data was acquired from the Accident Location Identification Surveillance System (ALISS) for this study.

3.1 Regional Archived Data Server (RADS)

The RADS system houses a variety of data archived by both ADOT and the Maricopa county DOT. The traffic data collected from ADOT's Freeway Management System (a collection of ITS devices deployed in Arizona) and stored in RADS is used to calculate travel times in 5-minute intervals. This data is used to provide traveler information to message signs and the 511 system.

In this project, "RADS data" will refer to 20-second loop detector data acquired from the RADS system. The ADOT traffic operation center (TOC) archives the 20-second data from all the dual loop detectors located along freeways in the Phoenix area, more than 400 stations and 8,000 loop detectors, every day. Each station covers several mainlines and an HOV lane; each lane has a dual loop detector consisting of two loops: leading and lagging. The data provides lane-by-lane speed, volume, and loop occupancy for I-10 in the Phoenix area.

3.2 National Performance Management Research Data Set (NPMRDS)

The NPMRDS database contains probe data collected by HERE (2017) through a mobile app, portable navigation devices and, occasionally, the vehicles' hardware. Freight data is included through a link with the American Transportation Research Institute (ATRI, 2012). This dataset contains travel time information in 5-minute intervals, or "epochs", for Traffic Management Channels (TMCs) (predefined road segments). The data was collected across the entire national highway system.

Travel time data is critical for identifying congestion and quantifying travel time reliability. The NPMRDS database can be used for both and is thus especially relevant for implementing a travel time reliability monitoring system such as the one described in L02 (UMRRC, 2014). Data for 2015 and 2016 were acquired for the entire state of Arizona and the relevant TMCs in the study corridors used for this implementation.

3.3 Highway Condition Reporting System (HCRS)

The HCRS database contains data on lane closures, incidents, and construction that occur on ADOT's highways and surface streets. Key fields include the start/end time, location, and type for each. The three most commonly observed types of entries are Incidents/Crashes, Lane Restrictions, and Closures, although there other types are interspersed throughout the data. The data is entered by TOC operators, state agencies/contractors, after validating issues using CCTV cameras, or imported via computer-aided dispatch. Lane restrictions and closures are generally due to work zones. In order to implement L07, the number of incidents, and their duration and type are all required.

3.4 Accident Location Identification Surveillance System (ALISS)

The ALISS database is a Model-Minimum Uniform Crash Criteria (MMUCC) standard-compliant crash database composed of three tables: Crash, Vehicle, and Person. The crash table is utilized in the current implementation of L02 and L07. It has one entry per reported crash and includes key information about the crash event including date, time, location, and severity.

In the L07 design guide, the total number of fatal/major injury, minor injury, and property-damage only (PDO) crashes are required for a variety of treatments. This data can also be utilized in conjunction with the HCRS incident data and the travel time data sets to determine the impact a crashe has on congestion. Data for 2015 and 2016 was once more acquired for the three study corridors.

3.5 Multimodal Planning Division Data

ADOT's MPD archived traffic data was collected from Automatic Traffic Recorders (ATRs) in the ADOT Transportation Data Management System (TDMS) server. Here, 15-minute volume and vehicle classification data is aggregated into speed bins by link. This data was utilized in L07 to obtain truck percentages as well as volumes for use in calculating the performance measures.

3.6 Data Not Included

ADOT was not collecting two data sources that are suggested by either or both of L02 and L07. First, weather data was not available at the resolution needed to confidently include it in the analysis for the study years. High-resolution, hourly precipitation data for earlier years was therefore collected through the National Oceanic and Atmospheric Administration (NOAA), while lower resolution, daily precipitation data was could be found through weather underground (Weather Underground, 2017). Finally, the NOAA's storm event database provides information on major weather events like thunderstorms and dust storms; however as these tend to be geo-located to either the Greater Phoenix area or Maricopa county, which is among the largest counties in the country, it was deemed unfeasible to use the NOAA storm database events for this project due to their low spatial resolution.

The second data source that was not included is special event data. Special events such as festivals and major sporting events are not recorded in the state of Arizona even though they typically result in irregular traffic patterns. In other states they are often recorded in the state's equivalent of the HCRS database and the project team recommends that ADOT consider incorporating both weather data and special event data in their future data collection efforts.

4 IMPLEMENTATION

The implementation of both L02 and L07 followed the guidebooks wherever possible. However certain adjustments needed to be made based on a lack of data or differences in the data that was available for the project. Any adjustments to the methodology due to a lack of availability of data or other reasons are documented in this report.

4.1 L02 Implementation.

The L02 design guide suggests a travel time reliability monitoring system (TTRMS) that follows the system design shown in Figure 4.1, which outlines the acquisition and flow of data required for a model system to be established. The specific system design proposed and implemented for this project is shown in Figure 4.2. For both figures, modules are represented by boxes and circles represent the inputs and outputs.

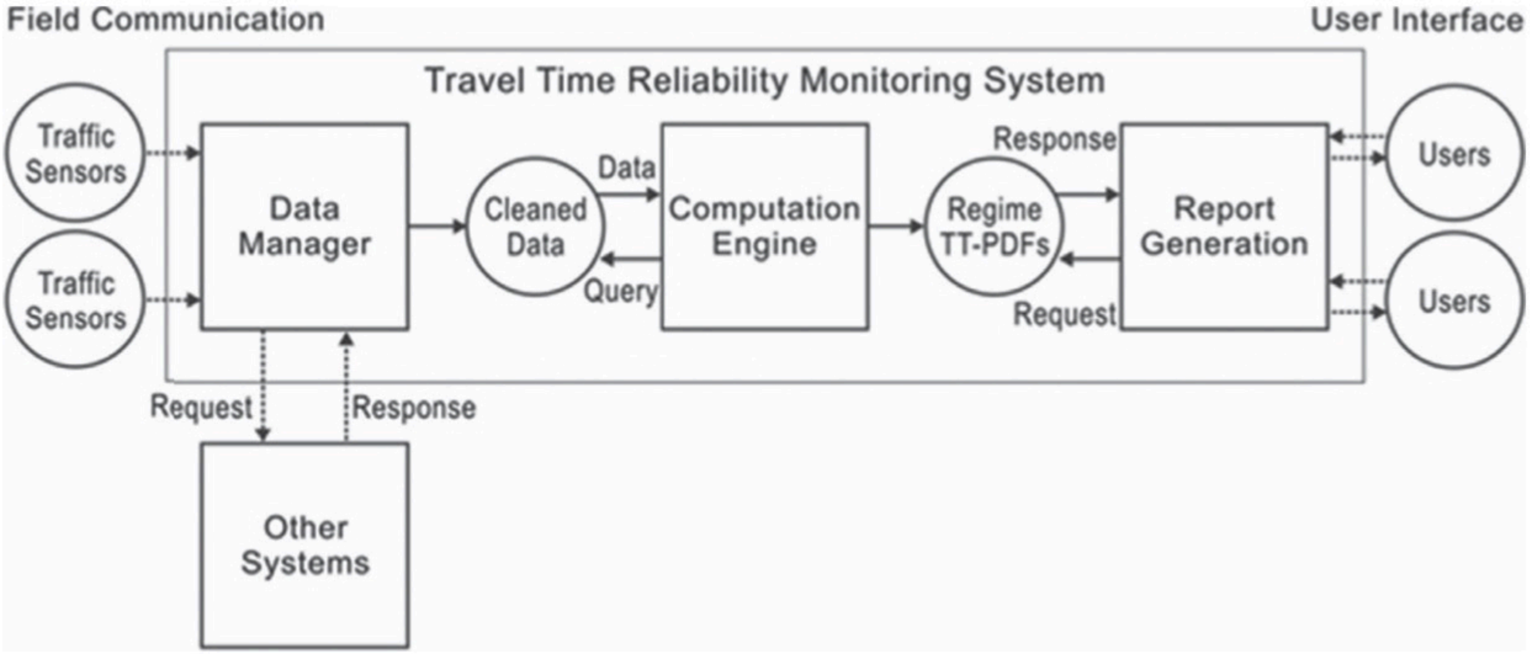


Figure 4.1: Model System Design for TTRMS (Source: L02 Guide (Potts et al., 2014))

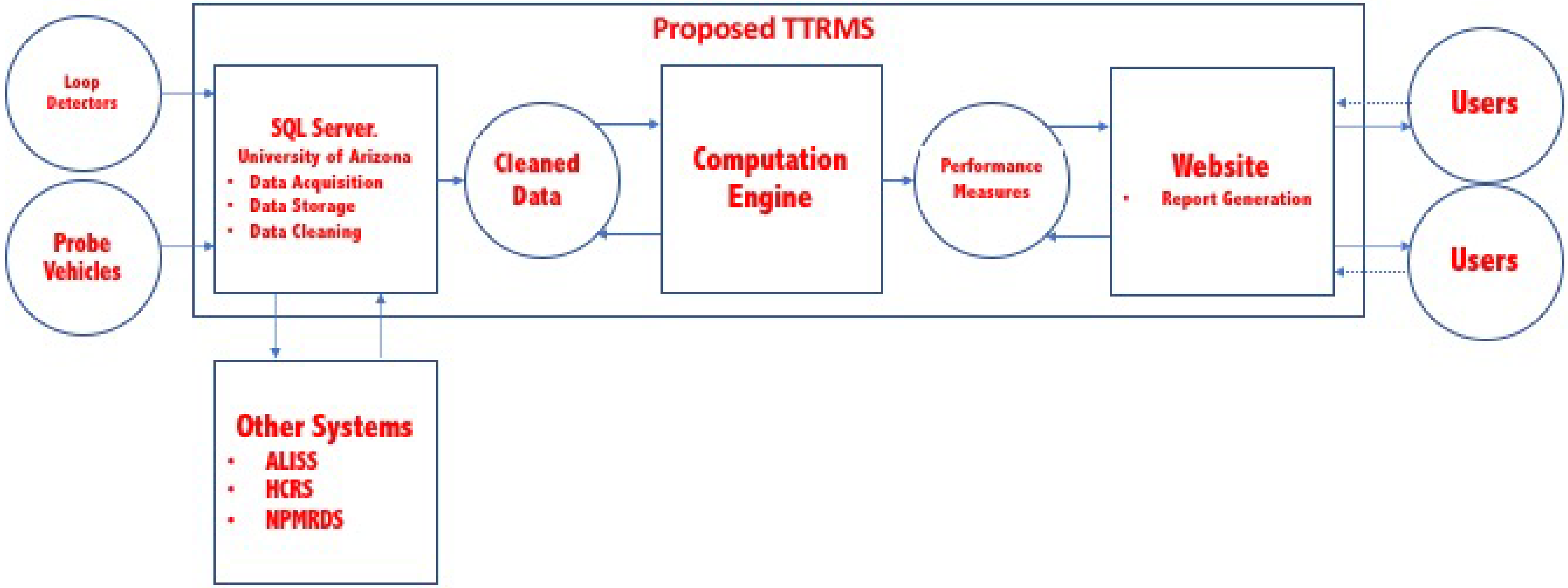


Figure 4.2: Proposed TTRMS System Implementation for ADOT

As illustrated in Figure 4.2, a new SQL server was purchased and installed at the University of Arizona to archive the data for this project. The research team developed a set of algorithms to clean the data, taking steps to identify outliers and conduct data imputation. The computation engine uses SQL queries and R, a programing language for statistical computing (R Core Team, 2017), to calculate performance measures. These measures can be accessed by users via either R or the associated website. A user with experience in R will have full access to the entire TTRMS developed, while the website currently includes only a limited number of performance measures. The website will be expanded in the future depending on the available resources and ADOT's strategic goals.

The roadmap for executing the framework is shown in Figure 4.3. This roadmap shows the seven steps required to establish the TTRMS, and how each step is related to the data processes listed in the L02 guide.

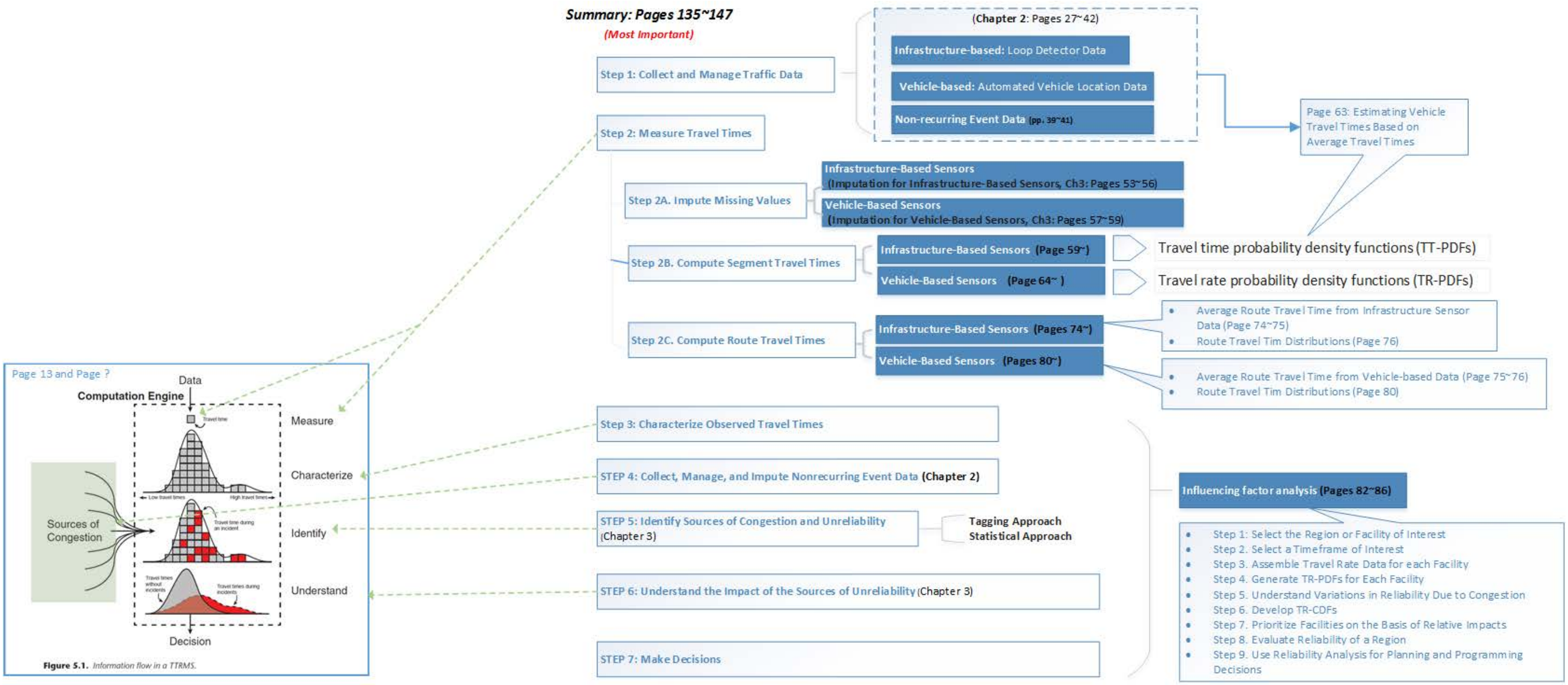


Figure 4.3: Roadmap for Establishing a TTRMS. Adapted from the L02 guide

The seven steps suggested in L02 are:

1. Collect and Manage Traffic Data
2. Measure Travel Times
3. Characterize Observed Travel Times
4. Collect, Manage, and Impute Nonrecurring Event Data
5. Identify Sources of Nonrecurring Congestion
6. Understand the Impact of the Sources of Congestion and Unreliability
7. Make Decisions

These are considered in turn below.

4.1.1 Collect and Manage Traffic Data

The data collected is stored on the STL-04 Server in the Smart Transportation Lab at University of Arizona. The data is continually updated as new data becomes available. For a complete description of the data collected, see Section 3: Data Collection.

4.1.2 Measure Travel Times

ADOT makes both probe data (from the NPMRDS) and infrastructure-based sensor data (from RADS) available for travel time measurement. One goal of this IAP is to determine which data set should be utilized in the TTRMS. In step two, both data sets are processed and travel times are measured. L02 does suggest some processing measures to be used but we found that the methods suggested were not completely applicable to the data available. As a result, the following methodologies were applied to the available datasets. Processing consists of 3 steps:

A. Imputation and Cleaning
B. Segment Travel Time Calculation
C. Corridor Travel Time Calculation

The NPMRDS data set was provided in the form of travel times calculated via data fusion. The original travel times come from both probe vehicles and infrastructure-based sensors. The travel times are in 5-minute epochs (time intervals) covering an entire year divided by pre-defined segments, labelled Traffic Message Channels (TMCs). The exact algorithm used to acquire these travel times is not available, so the accuracy is difficult to assess. Utilizing the data set involved a minimal level of imputation to compensate for some missing values; a moving median was used to impute missing or 0 values. Acquiring corridor travel times was then a matter of simply selecting the TMCs within the study corridor and using the methods described in L02 to convert segment TT to corridor TT.

The infrastructure-based sensor data from RADS is composed of speeds, volumes, and occupancies for each detector station over the course of each day. This makes it more complex to process, which is a downside to using RADS in this implementation. To process the data, four steps were taken using a java-based computer program written by the University of Arizona.

The computer program takes about three days to run a month of data for a corridor using a desktop computer with a Core i7 CPU and 16 GB RAM.   Long processing time is another downside of using the RADS data for the system.

The four steps are as follows; the entire process is described in more detail in Appendix 1:

1. *Generate Tables by Route.* RADS data is currently aggregated in a single table for each day for each station in the network. When monitoring travel time reliability, multiple days of data are required for a single location, rather than every location for a single period of time. A database table was generated using predefined routes in the RADS data and these tables were stored in a new database. In this process, data was also reformatted from MySQL to MSSQL, allowing for the entire data table to be migrated into MSSQL while maintaining the original structure.

2. *Quality Assurance/Quality Control (QA/QC).* Once the data has been reorganized, the data must be cleaned for use. This involved flagging and removing bad data, time alignment, and imputation. First, general criteria developed using traffic flow theory, are applied to check whether the observed data are theoretically feasible. Any issues are tracked and assigned into categories depending on the error. Following the feasibility checks, time alignment is implemented to obtain a complete time stamp with consistent times across each sensor. Finally flagged data (missing data or data that did not meet the original feasibility check) are removed and replaced using a simple imputation method. More on the QA/QC process can be found in Appendix 1.

3. *Spatio-Temporal Aggregation.* Although the data from step 2 is organized in 20-second intervals by detector station, the TTRMS developed considers corridor travel times in 5-minute intervals. Temporal aggregation takes the 20-second data and aggregates it into 5-minute data while maintaining the lane-by-lane features. Spatial aggregation takes the 5-minute temporally aggregated station-by-station data and aggregates it into segment-based averages. High Occupancy Vehicle (HOV) lanes and general purpose (GP) travel lanes are kept separate as the speeds and travel times are often different. Data QA/QC is repeated after each aggregation step to ensure quality.

4. *Travel Time Calculation.* Once the data has been successfully cleaned and aggregated, the travel times and rates can be calculated. The travel time is calculated using the instantaneous travel time model presented in Li et al. (2006). The results are then used in the analysis and implementation of the TTRMS.

While this process is computationally expensive and difficult to implement, it provides transparency and flexibility with the data. The source and quality of the data is known to ADOT and as new methods are developed the methods in this process could feasibly be adjusted to reflect that improvement. In contrast, the NPMRDS data set is a result of a proprietary data-fusion algorithm. This algorithm is neither transparent nor flexible but can be used as received with minimal QA/QC. This trade-off will be discussed further in the Application section.

4.1.3 Characterize Travel Times & Collect, Manage, and Impute Nonrecurring Event Data

Steps 4 and 5 recommended by the L02 guidelines were found to be somewhat ambiguous by the research team. Therefore, the steps taken in this section cover both steps 4 and 5. Travel times were characterized based on operating conditions and time of day. Using the available data, travel times were characterized based on the following aspects:

- Time-of-Day (Off-peak, Morning Peak, Evening Peak)
- Incident
- Closure

Each 5 minute epoch was characterized by joining the HCRS and ALISS databases to the calculated corridor travel times. For any entries that did not have a defined end time, two hours was used as the default value for the length of the event. This needs further guidance in L02, as does how to include incidents and closures that occur on nearby facilities that may be impacting the observed travel times.

4.1.4 Remaining Steps

The remaining steps are intended to analyze the results, identify sources on nonrecurrent congestion and make decisions with the resulting information. The steps taken to accomplish this are described in sections 5 and 6.

4.2 L07 Implementation

4.2.1 Data Compilation and Entry

The user interface for L07 has three sections – a user-input section, treatment data and calculation section, and a results section. Figure 4.4 shows the three interfaces.

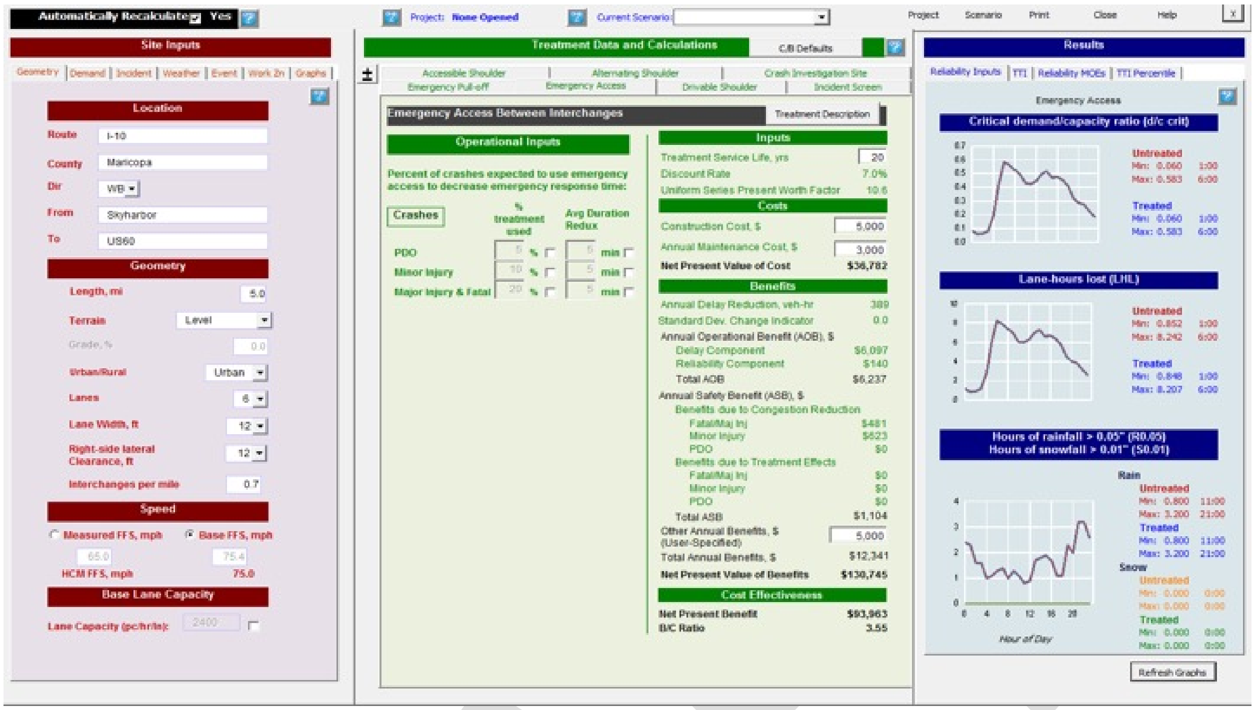


Figure 4.4: L07 Spreadsheet Interface

The user is responsible for adding site inputs in six categories:

- Geometry
- Demand
- Incident
- Weather
- Event
- Work Zones

The first four of those categories are condensed into Figure 4.5. The Geometry, Demand, and Incident inputs require data to be input manually from the following data sources: RADS, MPD, ALISS, and HCRS. The weather data input allows the user to select a proxy site from a large list, and provides default values. The event and work zone information are not entered since they require details that are not available in the HCRS database. To get some idea of the importance of these two fields in the tool's results, a fake work zone and fake event were both added, resulting in minimal changes to the resulting TTIs. Additional assumptions had to be made including:

- Level Terrain,
- Right-side lateral clearance is 12 ft.,
- 0.7 interchange per mile,
- PHF=1,
- The percentage of recreational vehicles (%RV)=1,

- Incapacitating injury and fatal injury (ALISS) are treated as major injuries and fatalities
- Possible injury and non-incapacitating injury (ALISS) are treated as minor injuries.

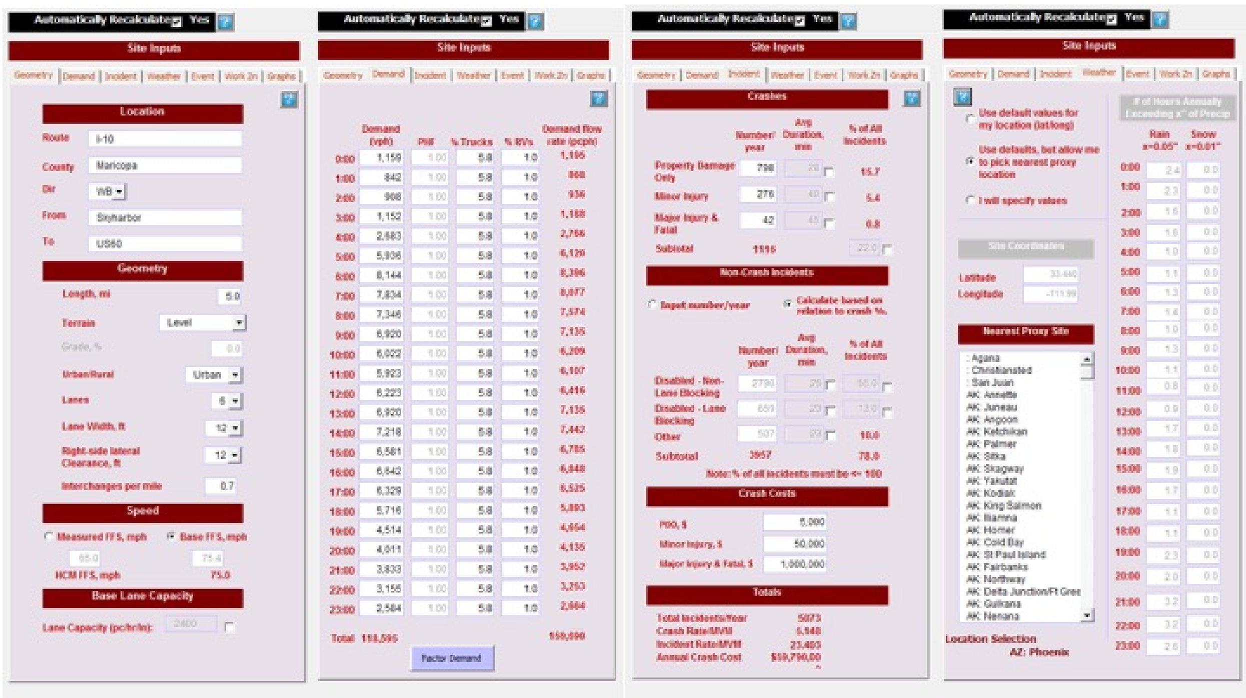


Figure 4.5: Sample of Relevant Inputs for Phoenix Corridors

The treatment designs are meant to demonstrate the impact of different treatments on various aspects of travel time and travel time reliability. Relevant information is already pre-loaded into the green area, but should be checked carefully if major decisions need to be made based on the results. This obviously further increases the data collection burden to include cost estimates, safety performance functions, and other information. For this project, defaults were left in place. Seven potential countermeasures were deemed reasonably appropriate and selected for analysis:

- Accessible Shoulder
- Alternating Shoulder
- Crash Investigation Site
- Emergency Pull-Off
- Emergency Access
- Drivable Shoulder
- Incident Screen

5 APPLICATION

This section will discuss the results of implementing the SHRP 2 guides in the state of Arizona. First the travel times from different data sources were compared, then the resulting travel time data were used for the implementation of the two products.

5.1 Travel Time Data Comparison for L02

One goal of this project was to compare the travel times output by NPMRDS to the travel times obtained using loop data from RADS. Both qualitative comparison and quantitative comparisons were made between the two data sets and statistical tests were used to determine whether the datasets differ significantly. If the statistical tests show the datasets are effectively interchangeable, an objective comparison can be used to select the simplest dataset to use for a specific application. However, if they are not, other aspects must be taken into consideration when selecting the data source for the travel time calculations.

The comparison of travel times was made on the corridor on I-10 in Phoenix between US-Route 60 and I-17 over a 3-month timespan between January 1st and March 31st, 2015. Data were divided into 5-minute epochs for both datasets and then paired for evaluation. Figure 5.1 shows a visual comparison of the two data sets. Although the two datasets do exhibit a linear relationship, the values are not equal. The NPMRDS data tends to consistently predict longer travel times than the RADS data.

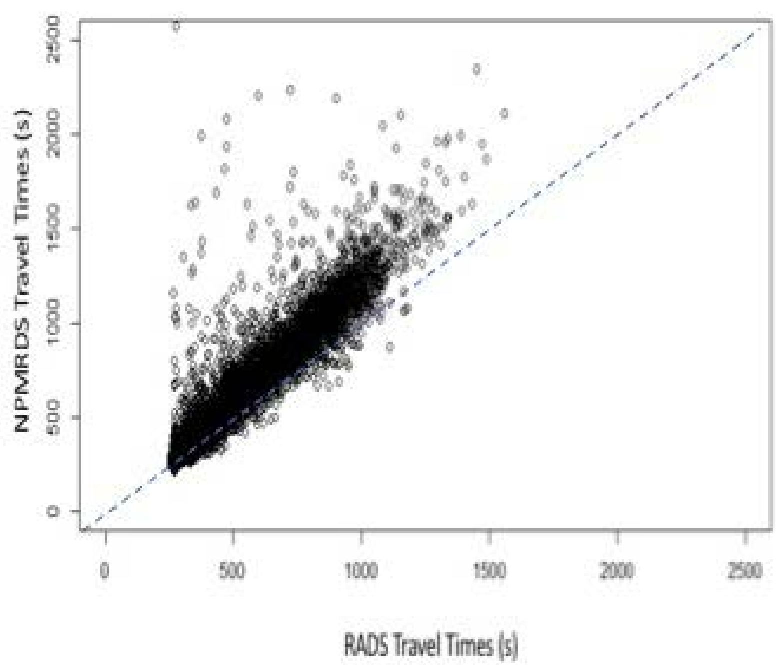


Figure 5.1: Comparison of RADS and NPMRDS Datasets

A set of paired t-tests were conducted over a 3-month period to test the difference between the travel times from each data source. The first test examined the overall difference in travel times between NPMRDS and RADS, while the subsequent tests checked the same differences for peak periods and off-peak periods individually. In the two-tailed paired t-test, the hypotheses are:

$$H_0\colon \mu_d = 0$$
$$H_A\colon \mu_d \neq 0$$

where μ_d is the average difference between the paired observations. If the null hypothesis is rejected, it means there is a non-zero difference between the two sources of data, while if the test is not rejected, the datasets are interchangeable. The test statistic, $t*$, is defined as:

$$t^* = \frac{\bar{X}_d - \mu_d}{s_d/\sqrt{n_d}}$$

where $\bar{X}_d$ is the average sample difference, s_d is the standard deviation of the difference, and n_d is the sample size. Using an $\alpha = 0.05$, Table 5-1 shows the result of the paired t-test.

Table 5-1: t-test Results

	All Travel Time Difference	Off-Peak Hour Travel Time Difference	Peak Hour Travel Time Difference
$\boldsymbol{\mu_d}$	20.9 s	12.85	44.1 s
t*	55.7	31.8	52.5
Test Result	Reject	Reject	Reject

Rejecting the Null Hypotheses signifies that the travel times are different with a 95% confidence level. However, even though the difference is significant from a statistical perspective, the average difference in the measured travel time is only 20 seconds for all travel times. This is a fairly small difference given the length of the corridor, so from a practical perspective the estimated times are not all that different, on average.

While the t-test looks at the average difference between the samples, it is also useful to test whether the samples have similar variances. An F-test was conducted under the following hypotheses:

$$H_0\colon \sigma_1^2 = \sigma_2^2$$
$$H_A\colon \sigma_1^2 \neq \sigma_2^2$$

The F-Statistic is:

$$F^*_{n_1-1,n_2-1} = \frac{s_1^2}{s_2^2}$$

where s_1^2 and s_2^2 are the sample variances. For $\alpha = 0.05$, Table 5-2 shows the result of the F-Test for the three scenarios.

Table 5-2: F-test Results

	All Travel Times	Off-Peak Hour Travel Times	Morning Peak Hour Travel Times
Var(NPMRDS)	33531	12910	79670
Var(RADS)	16096	4307	42230
F*	2.01	2.99	1.88
Test Result	Reject	Reject	Reject

The results of the F-test imply that the NPMRDS travel times have a much greater variance than the RADS travel times. This could be due to the large number of outliers in the dataset or it could be that the NPMRDS estimations are more sensitive to traffic conditions than the RADS estimations. Based on the results of the two tests, we can conclude that from a statistical perspective the travel times from the two data sources are different. The NPMRDS dataset appears to be the more conservative of the two datasets, providing higher estimated travel times when congestion is possible. However, the NPMRDS dataset did contain some outliers relative to the RADS dataset. In particular, one value was close to 10,000 seconds in NPMRDS but only 400 seconds in RADS.

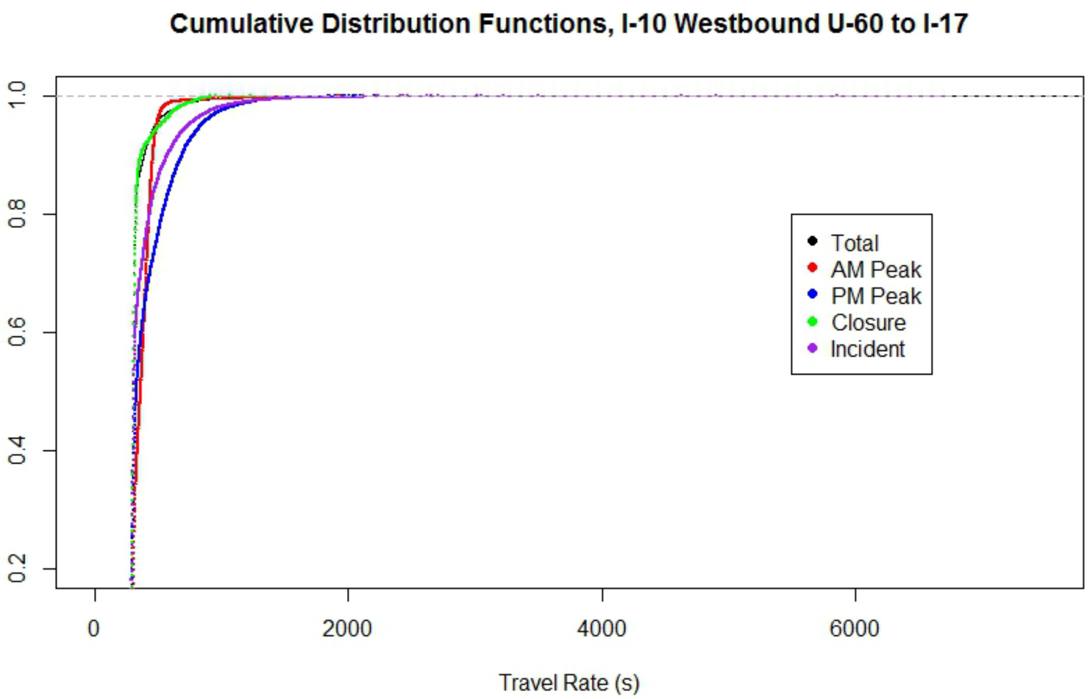
As discussed in section 4.1, both datasets have inherent benefits and drawbacks. Using RADS is computationally expensive and difficult to implement, but provides transparency and flexibility with the data handling. The source and quality of the data are both known to ADOT and as new methods are developed the methods in this process could feasibly be adjusted to reflect that improvement. Meanwhile the NPMRDS data set is the result of a private data-fusion algorithm. Although this is neither transparent nor flexible, it is useable as received with minimal QA/QC. There is a cost associated with obtaining and cleaning the datasets and this should be considered when making this decision.

5.2 L02 Results

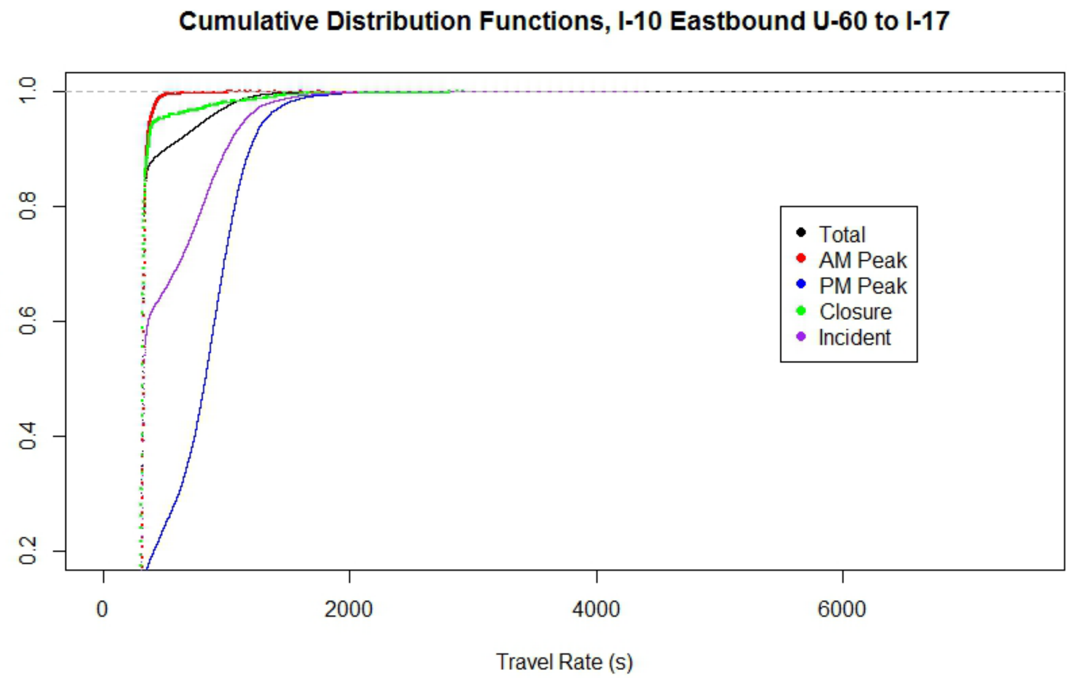
The purpose of this section is to present the remaining steps of the TTRMS process. These steps are: Identify Sources of Nonrecurring Congestion, Understand the Impact of Nonrecurring Congestion on Unreliability, and Make Decisions. Based on an analysis of the results from the NPMRDS travel times, the findings will be presented for all three study corridors.

Travel Time Reliability can be quantified and visualized in multiple ways within the L02 framework. The simplest way to quickly and objectively evaluate travel time reliability is to plot the Cumulative Distribution Function (CDF). L02 recommends evaluating the CDFs under different conditions known to affect the reliability of travel. CDF represents the percentage of travel rates equal to or below a specific value. In the series of graphs in Figure 5.2 for travel along both directions in all three of the study corridors, the x-axis shows the travel rate and the y-axis the percentile value for that travel rate. The conditions considered for this calculation were:

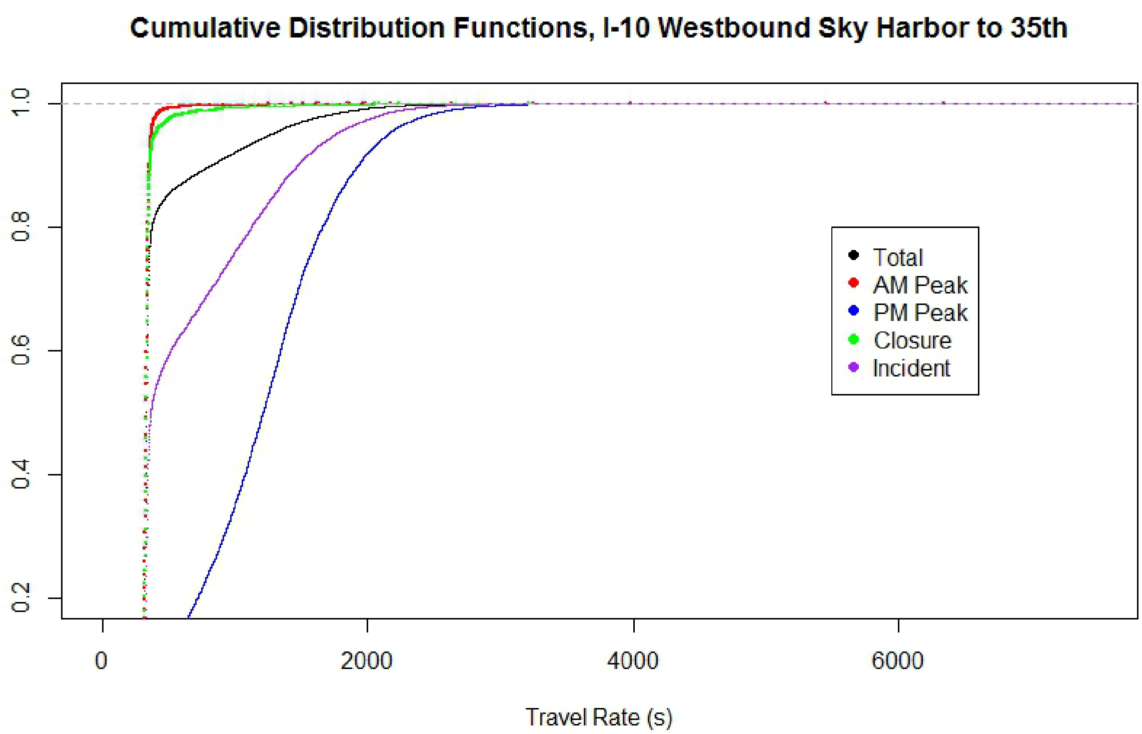
- Peak-Hours
- Crash/Incident
- Closures



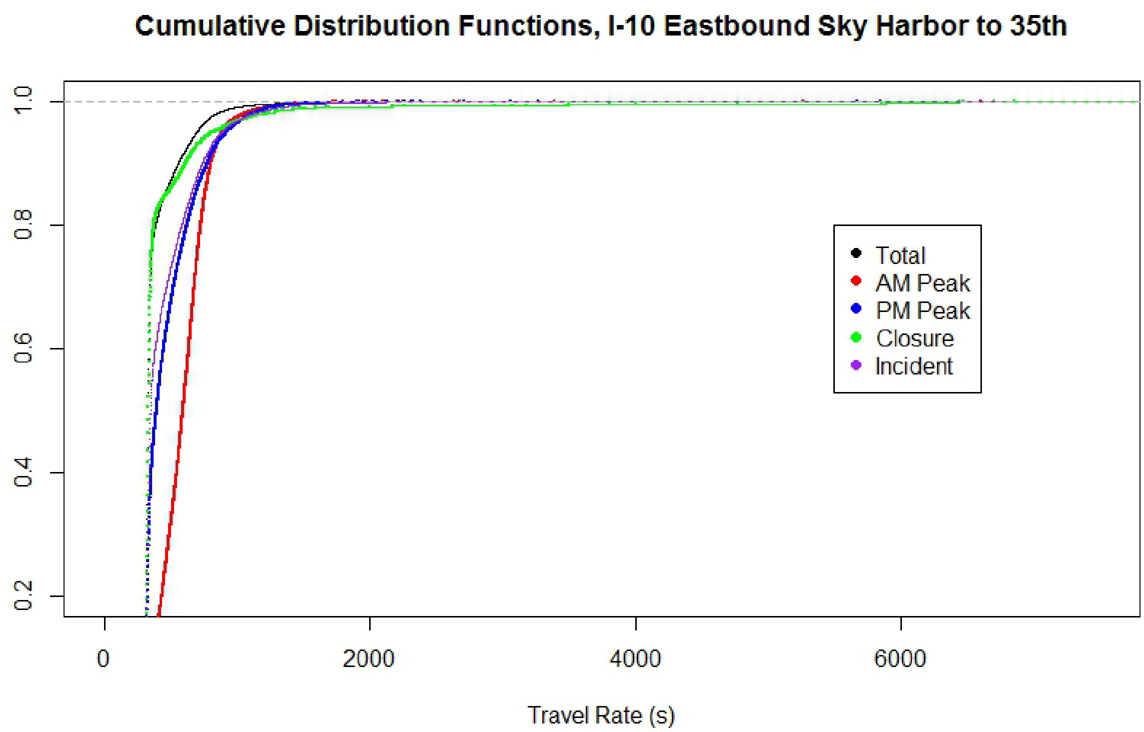
(a)


Cumulative Distribution Functions, I-10 Eastbound U-60 to I-17
Total
AM Peak
PM Peak
Closure
Incident
0.2
0.4
0.6
0.8
1.0
0
2000
4000
6000
Travel Rate (s)

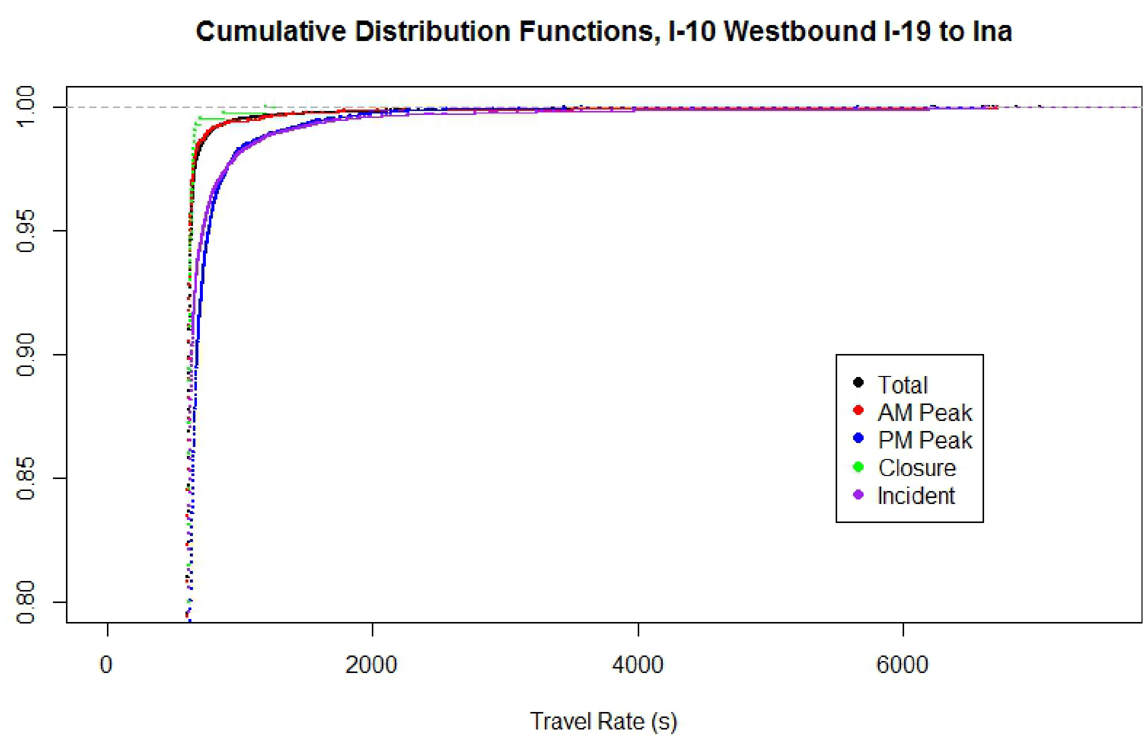
(b)


Cumulative Distribution Functions, I-10 Westbound Sky Harbor to 35th
Total
AM Peak
PM Peak
Closure
Incident
0.2
0.4
0.6
0.8
1.0
0
2000
4000
6000
Travel Rate (s)

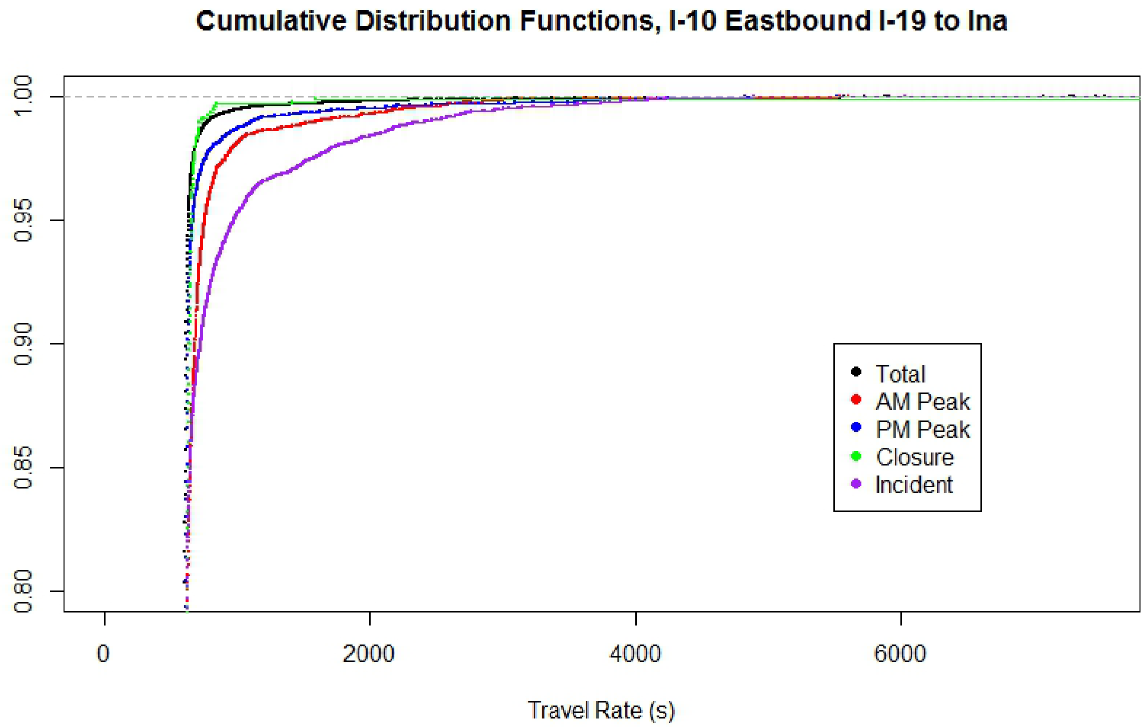
(c)


Cumulative Distribution Functions, I-10 Eastbound Sky Harbor to 35th
1.0
0.8
0.6
0.4
0.2
Total
AM Peak
PM Peak
Closure
Incident
0
2000
4000
6000
Travel Rate (s)

(d)


Cumulative Distribution Functions, I-10 Westbound I-19 to Ina
1.00
0.95
0.90
0.85
0.80
Total
AM Peak
PM Peak
Closure
Incident
0
2000
4000
6000
Travel Rate (s)

(d)



(f)

Figure 5.2 CDFs for Study Corridors

For the most part, the CDFs in Figure 5.2 show results that would be expected by anyone familiar with those corridors. Figures 5.2 (a) and (b) show the results for the corridor between US-60 and I-17. The eastbound direction looks normal, with increased travel times during the PM peak and incidents, but the westbound CDF does not show a similar increase during the AM peak. It is unclear why this should be the case, but possible reasons could include data quality, code bugs, or some other issue that requires further investigation. It is, for example, possible that the AM peak simply has more reliable travel times for that corridor, in the absence of incidents. The CDFs for the other two corridors exhibit normal and unsurprising results, with the AM peak going into the city having slightly increased issues with travel time and travel time reliability and the PM peak leaving the city exhibiting the same trend. As before, incident conditions are the least reliable.

In addition to the CDFs, other metrics can be calculated to help quantify different aspects of travel time reliability. Three metrics were chosen for this application, most of which are simply implementing an equation for the distribution of the travel times. The selected attributes are the misery index, buffer index, and travel time index. The misery index (MI) compares the worst trips to the expected travel times under free flow conditions and the buffer index (BI) represents the additional time most travelers should add to their trip to arrive on-time. The 95th percentile is a typical value used to calculate the buffer index. The equation for the buffer index is shown below. Finally, the travel time index (TTI) measures the average trip length and compares it to the expected travel time under free flow conditions.

The measures were calculated using NPMRDS travel times. Free flow speed was assumed to be the 25th percentile speed, which was slightly over the posted speed limit on the corridor. The results are presented in Table 5-3.

$$Misery\ Index = \frac{Travel\ Time_{97.5\%}}{Travel\ Time_{Free\ Flow}}$$

$$Buffer\ Index = \frac{Travel\ Time_{95\%} - Travel\ Time_{Mean}}{Travel\ Time_{Mean}}$$

$$Travel\ Time\ Index = \frac{Travel\ Time_{Mean}}{Travel\ Time_{Free\ Flow}}$$

Table 5-3: TTR Metrics

Corridors on I-10	Misery Index	Buffer Index	Travel Time Index
US 60 – I-17 (WB)	2.03	0.44	1.12
US 60 – I-17 (EB)	3.26	1.21	1.21
35th Ave – Sky Harbor (WB)	4.86	1.79	1.40
35th Ave – Sky Harbor (EB)	2.51	0.79	1.20
Ina – I-19 (WB)	1.14	0.04	1.04
Ina – I-19 (EB)	1.15	0.05	1.04

The travel time reliability metrics indicate varying levels of reliability issues. The corridors in Tucson are quite reliable and almost never experience substantial delays, as shown by the misery index close to 1 and the buffer index near 0. However in Phoenix the situation is more problematic: both directions of I-10 between 35th Avenue and Sky Harbor experience frequent delays and major delays can cause travel times to be 4-5 times worse than normal.

Under the framework developed for this project, ADOT will be able to identify locations where particular sites could benefit most from interventions to improve reliability. By comparing CDFs and important indices, sites that suffer from the worst congestion and reliability issues can be easily identified. Since CDFs present different conditional travel time reliabilities, additional clues suggesting potentially successful approaches for interventions could be revealed that were not previously available.

5.3 L07 Results

L07 Results are presented for the two corridors in Phoenix. The corridor in Tucson did not show any indication of travel time reliability issues and thus was not analyzed using the L07 spreadsheet. The other two corridors showed enough evidence in the L02 analysis to warrant the use of the L07 spreadsheet. The results can be viewed either in tabs such as those shown in Figure 5.3, which are similar to the input tabs, or printed in the form of a report (Figure 5.4). The benefit of reviewing the results in the tabs is that the treatments can be varied and the charts change dynamically, while the report format provides larger figures and more detail. The report output does require each tab and each treatment to have a separate report generated and output into a pdf format, however, and the resulting pdfs are in black and white and are not labelled for optimal viewing; the tabs are in color and provide more information. Finally the PDFs do print tables with the raw data used in the Travel Time Index plots so if a user was sufficiently motivated they could conduct a manual comparison.

Figure 5.3: 35th Ave to Sky Harbor Results (WB)

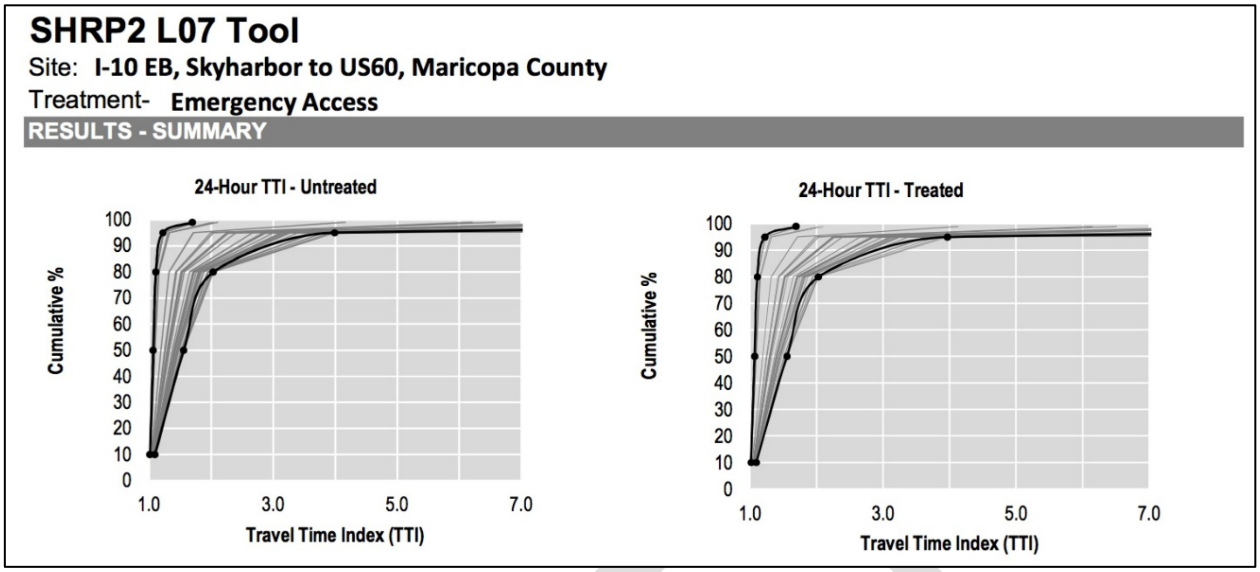


Figure 5.4: 35th Ave to Sky Harbor (EB)

Cost Benefits		
Construction Cost/Unit Cost	$	5,000
Annual Maintenance/Op Cost	$	3,000
Service Life		20
Present Value of Cost	$	36,782
Annual Delay Reduction, veh-hr		21404
Standard Dev. Change Indicator		1.6
Annual Operational Benefit (AOB) $		
VOT Component	$	335,610
VOR Component	$	25,359
Total	$	360,970
Annual Safety Benefit (ASB), $		
Benefits due to Congestion Reduction		
Fatal/Injury	$	9,887
PDO	$	17,976
Benefits due to Treatment Effects		
Fatal/Injury	$	0
PDO	$	0
Total	$	68,227
Other Annual Cost Benefits	$	5,000
Total Annual Benefit	$	434,197
Present Value of Benefits	$	4,599,887
B/C Ratio		125.06
Net Present Benefit	$	4,563,105

Figure 5.5: Cost Analysis of Selected Treatment (35th-Sky Harbor)

Overall, the treatments did not seem to have a noticeable impact on the predicted travel time indexes. This is likely due to the fact that many of these treatments have only a marginal effect on segments with such high demand. While the graphs look similar, the Net Present Benefit of the example treatment, “Emergency Access”, on the eastbound direction shown in Figure 5.5 was $4.5 million, but it is unclear if the default values for costs are reasonable or how accurate the underlying models are when quantifying the dollar value of the benefits. Neither corridor showed promising results from the treatments tested.

5.4 Incorporating TTRMS into Performance Measurement

Our project further extends the concept of TTRMS by developing a new online, statewide performance measurement prototype system. Figure 5.6 shows the statewide performance measurement framework initially developed for this project to demonstrate the feasibility of the integration of TTRMS with other performance measures. This framework contains a total of 27 measures in 4 categories. The specific list of measures included in the prototype system is shown in Figure 5.7. These 27 measures can all be obtained from automated databases. The 27 measures are intended to utilize the statewide data collected by ADOT to address issues with Freeway Management, Freight, Incident Management, and ITS Management. This database is currently in the process of being integrated into the online system, to produce a website that allows users to design their own queries. The database will enable advanced users to quantify measures in the categories of safety, mobility, freight, and ITS management directly using the data sources described in Chapter 3.

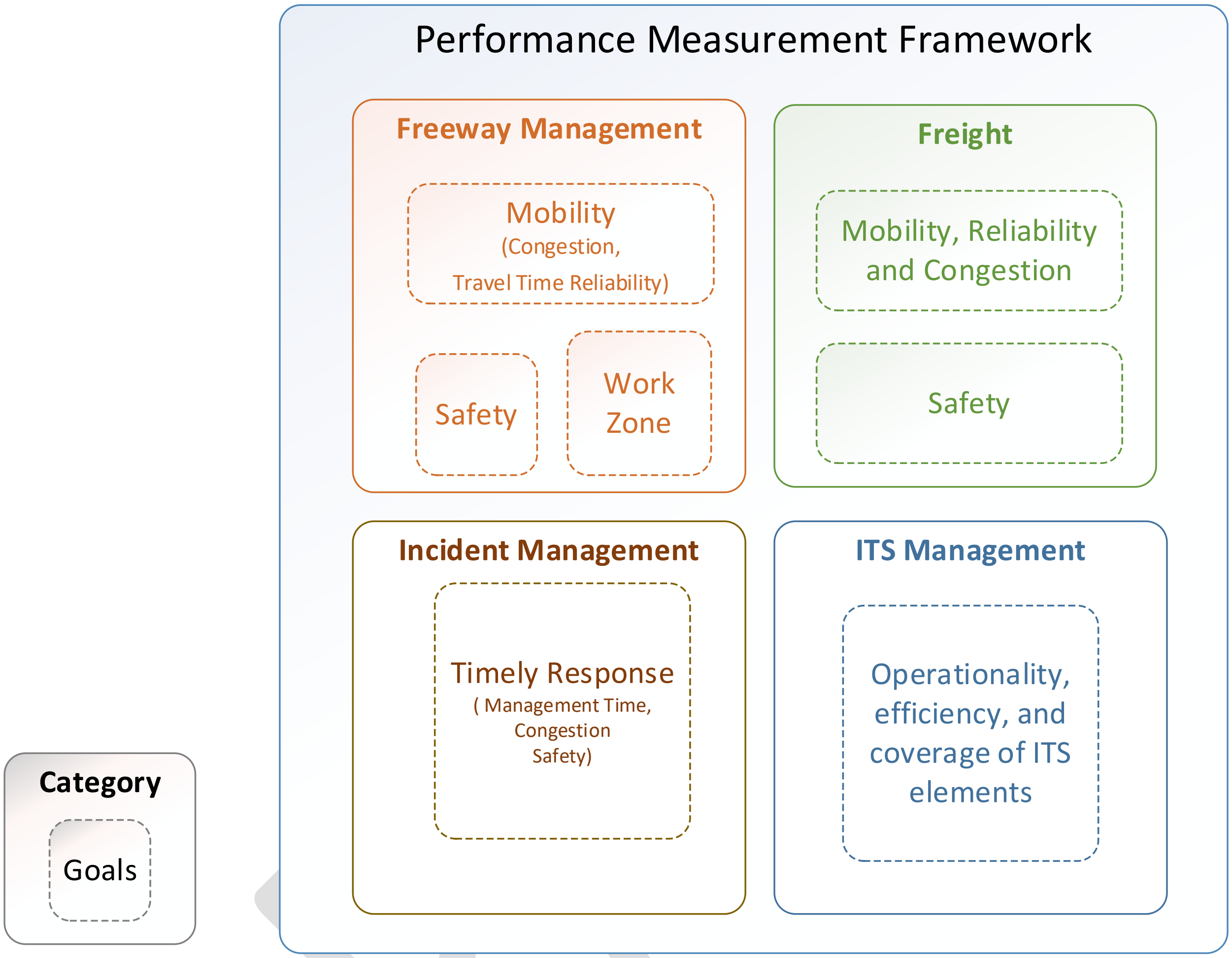


Figure 5.6: ADOT Performance Measurement Framework

Figure 5.7. Performance Measures to be examined

Since most transportation engineers are unfamiliar with database procedure operations, a user-friendly website (www.adottsmopm.org) was developed for this project to enhance interactivity between the statewide data and transportation practitioners. In general, the website allows users to select different attributes such as roadway, direction, milepost range, date range and time interval and then creates an SQL query that is sent to the server at University of Arizona's Smart Transportation Lab. At present, because this website was developed as a proof-of-concept, only one measure has been fully implemented. However, ADOT should be able to complete the development of the website based on the proposed framework and software architecture.

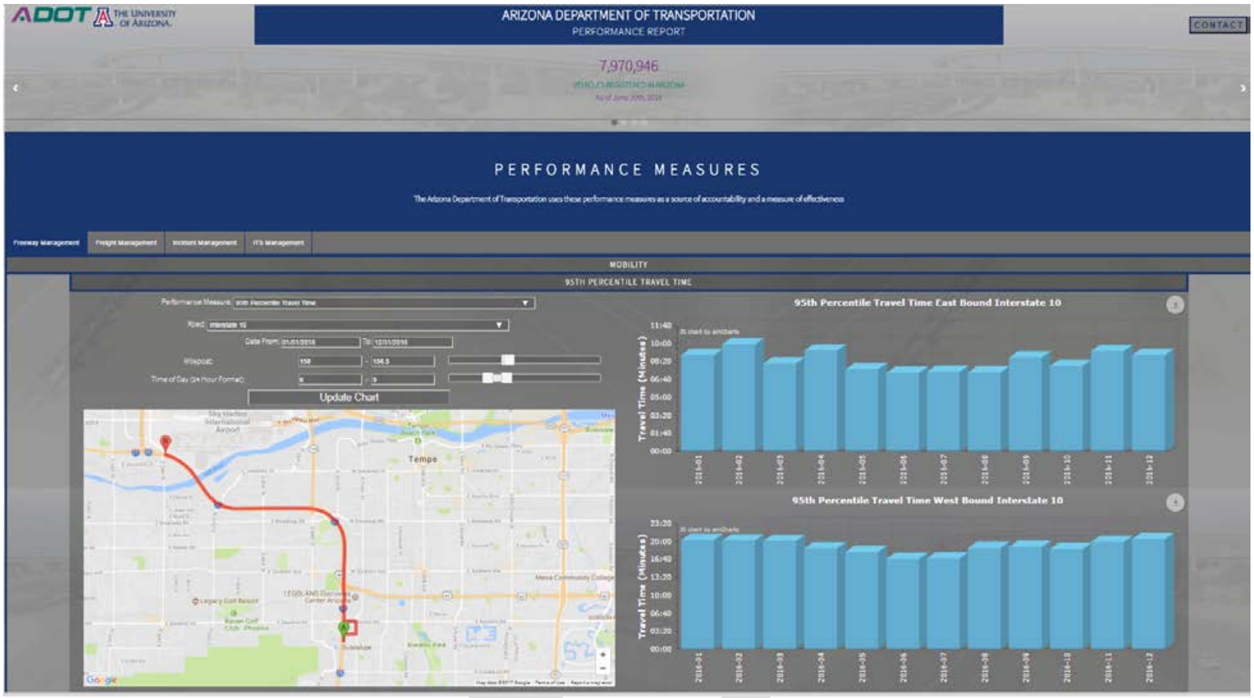


Figure 5.8 ADOT Performance Measurement Website for Transportation System Management and Operations (TSMO)

6 CONCLUSION AND RECOMMENDATIONS

6.1 Conclusion

This project successfully implemented both L02 and L07 design guides from SHRP 2 for three study corridors in the state of Arizona. The tasks accomplished for this project are summarized below.

- In order to implement the design guides, significant efforts were made to collect a wide variety of data and explore the possibility of incorporating the data from each source into the Travel Time Reliability Management System (TTRMS). The travel time data used for this IAP project were primarily collected from three major sources: AZTech's Regional Archived Data Server (RADS), the MS2 database managed by ADOT Multimodal Planning Division and the FHWA's National Performance Management Research Data Set (NPMRDS). Other datasets examined included crash data from the Accident Location Identification Surveillance System (ALISS) and Computer-Aided Dispatch (CAD) data from the Highway Condition Reporting System (HCRS). Weather and special event data was excluded in this study due to data quality issues.
- For the implementation of the L02 guide:
    - A prototype TTRMS system was created based on the L02 guidelines. The new system includes a database containing ADOT traffic data and R code that can be utilized to create CDFs and other Travel Time Reliability metrics using the database.
    - Novel approaches to processing the loop detector and NPMRDS data were developed for the TTRMS because the original guide provides only general guidance for data processing.
    - The two corridors with the worst travel times utilized L07 to identify and evaluate possible treatments. Applying the Cumulative Density Functions (CDFs) and Travel Time Reliability (TTR) metrics revealed that I-10 between Sky Harbor and 35th Avenue in Phoenix suffered from the worst reliability among the three study corridors under nearly all conditions, including regular demand fluctuations at peak hours and unexpected fluctuations during incidents.
- For the implementation of the L07 guide:
    - Due to TTR issues identified for the two study corridors in Phoenix during the implementation of L02, the collected data were input in the L07 spreadsheet to develop potential counter measures. A total of seven countermeasures were evaluated based on the spreadsheet results.

Two additional tasks were completed in addition to the implementation of the design guides:

- A comparison was made between the performance achieved using two travel time data sources, namely the NPMRDS and the RADS-based travel times. The comparison showed that RADS and NPMRDS data do not estimate the same values for travel times along the study corridor, with the NPMRDS data producing more conservative estimates. Ultimately, a decision on which dataset is more accurate demands further study
- An automated performance measure prototype database system was developed to process the data collected statewide. 27 statewide measures were developed to

demonstrate the feasibility of the proposed measures, including measures provided by the proposed TTRMS. An interactive prototype website (www.adottsmopm.org) was created to display one measure, the 95th percentile travel time, to demonstrate proof-of-concept. Currently the website has limited coverage and performance measures; however it has shown its potential to be expanded to cover a variety of measures in real time.

6.2 Recommendations

The project team identified six key recommendations. Some can be implemented by ADOT, while others are directed toward the expansion of the SHRP2 products. The recommendations are as follows:

- Additional data should be collected, including weather and special event data;
- The TTRMS established was beneficial for understanding travel time reliability along the corridors and should be implemented statewide;
- Although the L07 design guide provided useful information on appropriate countermeasures for different observed issues, the spreadsheet tool was not user friendly and should be improved;
- A dedicated data server should be established to facilitate the long-term implementation of these products statewide;
- To achieve the system's full potential, the use of additional performance measures calculated using the data collected in the project should be explored;
- A dedicated website could provide a user-friendly way to display the results of the TTRMS.

There are two data collection-related recommendations. HCRS was used in both L02 and L07 for this project, but an improved dataset could create additional opportunities for ADOT in the L02/L07 projects as well as in other applications. The HCRS data can be expanded to include additional inputs, and would benefit from some added consistency in the fields. The data can be added by a variety of agencies via computer-aided dispatch, but mistakes were common in entries and some of the fields were ambiguous. Additional training in data entry and quality control could improving the dataset substantially and ensure the findings and decisions made when using the dataset are more reliable. Additionally, HCRS should be expanded to include special events in the database; many states already do this in their incident databases. HCRS currently labels incidents as "Closures", which may or may not include the management of traffic during special events; there is no way to tell within HCRS. At present, events will only be labelled if they requires lane closures, but an event that does not require lane closures but induces a surge in demand cannot be identified in HCRS. More meticulous data collection and the inclusion of additional categories such as special events would make HCRS a much more useful dataset for applications such as those considered here.

The other dataset that ADOT should collect is more detailed weather data along the state's major interstates and arterial roads. While this is perhaps less urgent, given the climate in Arizona, monsoons and dust storms throughout the state, as well as major snow events in Northern Arizona, can have a major impact on traffic at certain times of the year. The currently available weather datasets provide insufficient data for any meaningful decision making purposes. If ADOT opts to include weather in any of its traffic management applications, a database should be created for their system with hourly data on various aspects of weather. This will require weather stations to be deployed along major roadways statewide in order to adequately capture this information.

The second data collection recommendation concerns the storage of data. A significant portion of this project was devoted to gathering data from different sources and formatting them so they could be combined. To implement a TTRMS system statewide that can be utilized on a monthly or yearly basis would require this process to be centralized and semi-automated. Establishing a dedicated server that retrieves the datasets, automatically reformats the data, and cleans it will support better repeatability of the process over time and along different corridors.

We recommend fully developing the proposed TTRMS in L02 for statewide use. The measures were found to contribute useful information regarding travel time reliability along the study corridors and helped improve our understanding of the underlying causes of traffic congestion . Currently, all output results of the TTRMS (e.g. CDFs and other metrics) require knowledge of the R-programing language and SQL database techniques, thus requiring traffic engineers to acquire some basic computer coding skills.

Finally, using the data collected in the project, a website to display performance measures statewide was created (www.adottsmopm.org). We recommend that TTRMS should be fully integrated with the prototype performance measurement website to provide a user-friendly interface that ADOT staff can use to quickly access information on specific corridors and obtain accurate results. This will greatly increase the usability and applicability of TTRMS. Given the effort required to compile the data, once this has been done it should be used to obtain further measures in addition to the TTRMS, including performance measures covering mobility, safety, freight, and asset management, all of which could be integrated into the website.

There were two recommendations on ways to improve the products developed by SHRP 2. For L02, additional guidance on the data sources would be useful for many applications. Also, crashes have knock-on effects that affect adjacent roadways along the same corridor for an extended period of time, depending on their severity, but most MMUCC compliant crash databases do not record the length of time involved for such incidents. If 100% of crashes were also in HCRS this would not be an issue, but this is generally not the case. Additional guidance should be included on how long crashes are likely to impact travel times based on the time-of-day they occur and the severity of the event.

Two major recommendations were made regarding the L07 spreadsheet tool. First, the user interface was very difficult to use in practice. Macro-related errors were common in both the Mac and PC versions of Microsoft Excel. Even though most of these errors could be handled adequately through trial-and-error, they made the use of the tool tedious and time consuming. It is therefore recommended that FHWA address the errors in the spreadsheet tool. Alternatively, developing an online tool where the user uploads pre-formatted data and a dedicated host server runs the calculations may be preferable to ensure compatibility and smooth operations.

Another issue with the L07 spreadsheet tool was the lack of a user guide. The L07 design guide document provided a comprehensive overview of potential countermeasures but contained little information for users seeking to use the spreadsheet tool. The best resources found for implementing the spreadsheet were actually the L38 Pilot Studies rather than the original L07 study itself. It would be desirable to have improved explanations built in to the guide or tool.

The project team successfully incorporated the proposed TTRMS into a prototype performance measurement system, but with very limited functionality as yet. To further develop both systems and ready them for regular use as a routine part of ADOT’s daily operations, it is strongly recommended that FHWA provide additional resources to help ADOT complete the systems. Additionally, ADOT and/or FHWA need to consider budgeting additional maintenance costs for operations and to support the sustainability of the newly developed systems.

7 REFERENCES:

AASHTO. (2010). *Highway Safety Manual* (1st ed.). Washington, D.C.

Alemazkoor, N., Burris, M. W., & Danda, S. R. (2012). Using Empirical Data to Find the Best Measure of Travel Time Reliability. *Transportation Research Record*, *2530*, 93–100. https://doi.org/10.3141/2530-11

ATRI. (2012). American Transportation Research Institute. Retrieved from http://atri-online.org

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

Carrion, C., & Levinson, D. (2012). Value of travel time reliability : A review of current evidence. *Transportation Research Part A*, *46*(4), 720–741. https://doi.org/10.1016/j.tra.2012.01.003

Dailey, D. J. (2006). *The Use of Weather Data to Predict Non-recurring Traffic Congestion*. Olympia, WA.

Hadi, M., Xiao, Y., Wang, T., Hu, P., Jia, J., Edelstein, R., & Lopez, A. (2015). *Pilot Testing of SHRP 2 Reliability Data and Analytical Products : Florida*. Washington, D.C. https://doi.org/10.17226/22331

HERE. (2017). HERE. Retrieved from HERE.com

Hojati, A. T., Ferreira, L., Washington, S., & Charles, P. (2016). Modelling the impact of traffic incidents on travel time reliability. *Transportation Research Part C*, *65*, 49–60. https://doi.org/10.1016/j.trc.2015.11.017

Kwon, J., Mauch, M., & Varaiya, P. (2006). Components of Congestion Delay from Incidents , Special Events , Lane Closures , Weather , Potential Ramp Metering Gain , and Excess Demand. *Transportation Research Record: Journal of the Transportation Research Board*, (1959), 84–91.

Li, H., Lavrenz, S. M., Day, C. M., Stevens, A. L., & Bullock, D. M. (2013). Use of Both Travel Time and Travel Time Reliability Measures to Quantify Benefits of Signal Timing Maintenance and Optimization. *Transportation Research Record: Journal of the Transportation Research Board*, (January). https://doi.org/10.3141/2487-05

List, G. F., Williams, B., Rouphail, N. M., Hranac, R., Barkley, T., Mai, E., … Khattak, A. (2014a). *Guide to Establishing Monitoring Programs for Travel Time Reliability*. Washington, D.C.

List, G. F., Williams, B., Rouphail, N. M., Hranac, R., Barkley, T., Mai, E., … Khattak, A. (2014b). *Handbook for Communicating Travel Time Reliability Through Graphics and Tables Reliability Through Graphics and Tables*. Washington, D.C.

Liu, Y., Chang, G., & Yu, J. (2011). An Integrated Control Model for Freeway Corridor Under Nonrecurrent Congestion. *IEEE Transactions on Vehicular Technology*, *60*(4), 1404–1418.

Lomax, T., Schrank, D., Turner, S., & Margiotta, R. (2003). *Selecting travel reliability measures*.

Ma, Z., Koutsopoulos, H. N., Ferreira, L., & Mesbah, M. (2017). Estimation of trip travel time distribution using a generalized Markov chain approach. *Transportation Research Part C*, *74*, 1–21. https://doi.org/10.1016/j.trc.2016.11.008

Nisbet, J., Bremmer, D., Yan, S., Murshed, D., Wang, Y., Zou, Y., … Zhang, Y. (2015). *Pilot Testing of SHRP 2 Reliability Data and Analytical Products : Washington*. Washington, D.C. https://doi.org/10.17226/22254

Potts, I. B., Harwood, D. W., Fees, C. A., Hutton, J. A., & Kenzel, C. (2014). *Design Guide for*

Addressing Nonrecurrent Congestion. Washington, D.C.
R Core Team. (2017). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/
Sobolewski, M., Polum, T., Morris, P., Loos, R., & Anderson, K. (2015). *Pilot Testing of SHRP 2 Reliability Data and Analytical Products : Minnesota*. Washington, D.C. https://doi.org/10.17226/22255
Texas Transportation Institute. (2014a). *Performance Measure Summary - Phoenix-Mesa AZ The Mobility Data for Phoenix-Mesa AZ*.
Texas Transportation Institute. (2014b). *Performance Measure Summary - Tucson AZ The Mobility Data for Tucson AZ*.
Uchida, K. (2014). Estimating the value of travel time and of travel time reliability in road networks. *Transportation Research Part B*, *66*, 129–147. https://doi.org/10.1016/j.trb.2014.01.002
Upper Midwest Reliability Resource Center (UMRRC). (2014). Travel Time Reliability Reference Manual. Retrieved from https://en.wikibooks.org/wiki/Travel_Time_Reliability_Reference_Manual/About_the_Travel_Time_Reliability_Reference_Manual
Van Lint, J. W. C., Zuylen, H. J. Van, & Tu, H. (2008). Travel time unreliability on freeways : Why measures based on variance tell only half the story. *Transportation Research Part A*, *42*, 258–277. https://doi.org/10.1016/j.tra.2007.08.008
Weather Underground. (2017). Available at https://www.wunderground.com.
Williges, C., McCullough, B., Chu, Y. Y., Amataya, N., Kuo, R., Lin, M., & Chu, L. (2015). *Pilot Testing of SHRP 2 Reliability Data and Analytical Products : Southern California*. Washington, D.C. https://doi.org/10.17226/22332
Wright, B., Zou, Y., & Wang, Y. (2015). Impact of Traffic Incidents on Reliability of Freeway Travel Times. *Transportation Research Record: Journal of the Transportation Research Board*, *2484*, 90–98. https://doi.org/10.3141/2484-10
Xu, C., Liu, P., Wang, W., & Li, Z. (2012). Evaluation of the impacts of traffic states on crash risks on freeways. *Accident Analysis and Prevention*, *47*, 162–171. https://doi.org/10.1016/j.aap.2012.01.020
Yang, S., Malik, A., & Wu, Y. (2014). Travel Time Reliability Using the Hasofer – Lind – Rackwitz – Fiessler Algorithm and Kernel Density Estimation. *Transportation Research Record: Journal of the Transportation Research Board*, *2442*, 12–15. https://doi.org/10.3141/2442-10
Yang, S., & Wu, Y. (2016). Mixture Models for Fitting Freeway Travel Time Distributions and Measuring Travel Time Reliability. *Transportation Research Record: Journal of the Transportation Research Board*, *2594*, 95–106. https://doi.org/10.3141/2594-13
Yeo, H., Jang, K., Skabardonis, A., & Kang, S. (2013). Impact of traffic states on freeway crash involvement rates. *Accident Analysis and Prevention*, *50*, 713–723. https://doi.org/10.1016/j.aap.2012.06.023
Zheng, F., van Zuylen, H. J., Liu, X., & Vine, S. Le. (2016). Reliability-Based Traffic Signal Control for Urban Arterial Roads. *IEEE Transactions on Intelligent Transportation Systems*, 1–13.

8 APPENDIX 1: WORK FLOW SUMMARY FOR DATA CLEANING AND MIGRATION

The following section describes the RADS data cleaning process.

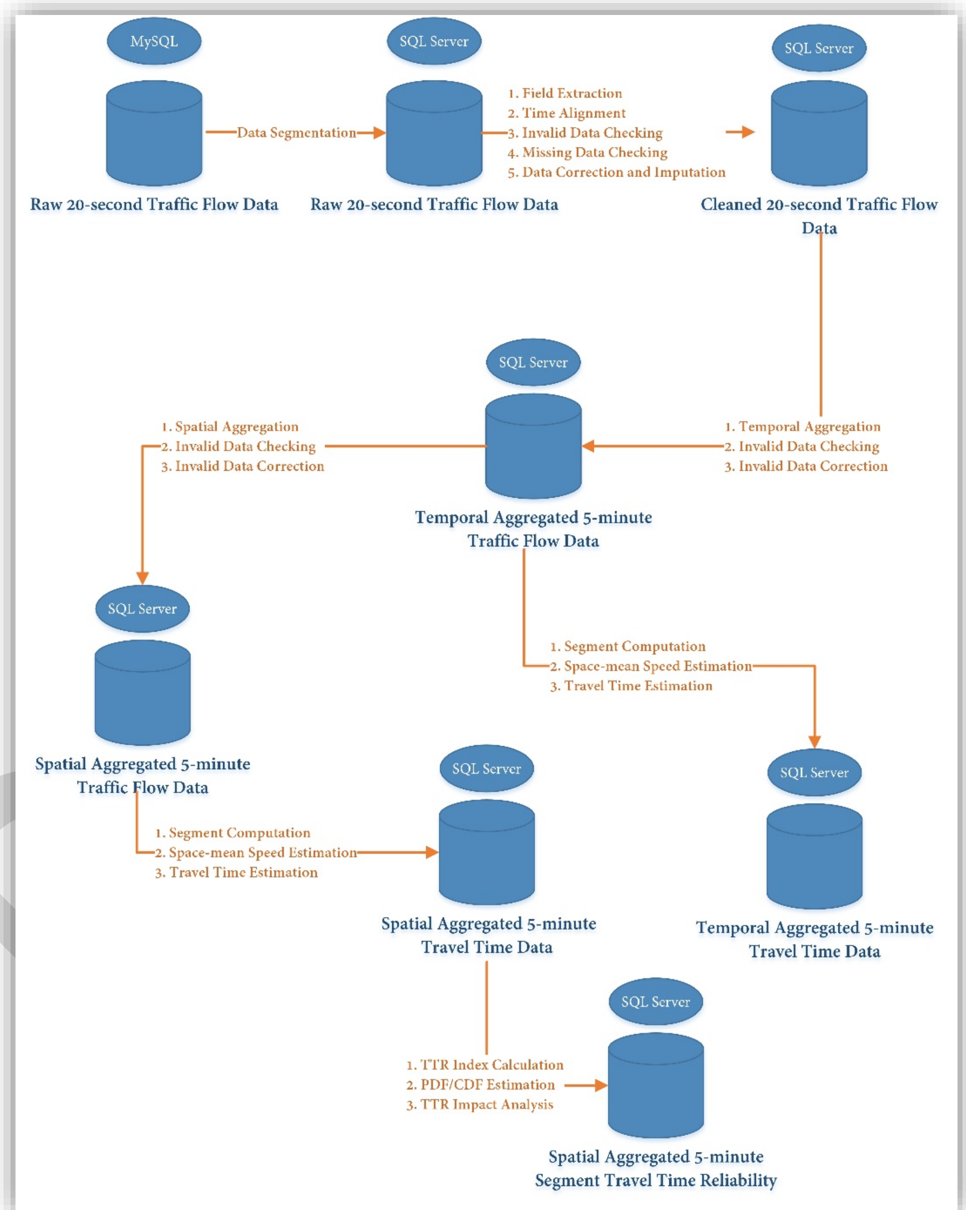


Figure 1 Work Flow of SHRP2 Project

1. Development Steps for SHRP2

a) Complete Data Migration for 20-second data

Since raw 20-second dual-loop detector data were saved in the MySQL database, the SHRP2 project should be implemented based on the SQL Server database (for consistency with other data sources) in order to migrate data from MySQL to SQL Server. The whole migration step was implemented as follows:

1) Automatically create a new table in terms of date collecting data records to save the migrated data records;
2) Read and insert data records into the created table in batch mode;
3) Repeat the above two steps until all data records in MySQL database are migrated to SQL Server database;

b) Re-organizing Data by Route for 20-second Data

To read and operate data faster without adding any optimizing complexity, the data records for every day were segmented route by route, i.e. each table stores data from just one route. Thus, separate data migration was conducted as follows:

1) Separate the target road network into multiple routes;
2) Automatically create a new table in terms of date collecting data records to save the migrated data records;
3) Read and insert data records into the created table in batch mode;
4) Repeat step 2) and step 3) until all data records in MySQL database are migrated to SQL Server database;

Figure 2 shows the results of an individual data migration step.

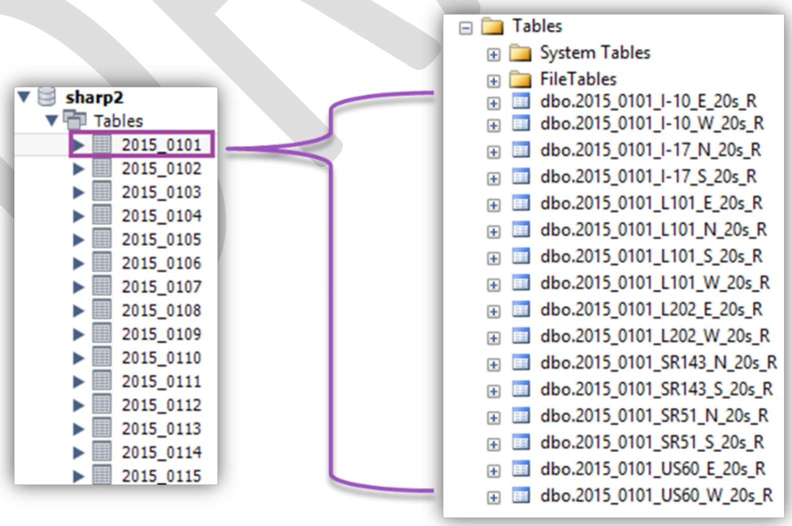


Figure 2 Data Migration for an individual 20-second data point

c) Traffic Flow Data Checking for 20-second Data

Due to communication failures, detector malfunctions, and so forth, there were some erroneous and missing data records in the collected data set. Thus, it is vital to conduct data quality assurance (DQA) before analyzing and modeling. Data checking is a key step of DQA. In this project, two operations involving data checking were conducted:

1) Time Alignment

The timestamps logged by the data collection system were not consistent for every day. When conducting travel time or travel time reliability estimation, it is necessary to make a normal form for each given timestamp read from the database management system. Since all successive analysis operations will be built on the 20-second interval, the main goal of the time alignment operation is to ensure the given timestamp drops into the appropriate 20-second interval. This is especially important when processing cases where there are missing data records or repeating data records. The implementation of the time alignment operation was conducted as follows:

- Create hash mapping between corrected timestamp and data records
- Compare the preset time with record time of the collected data
- Assign proper timestamp according to above comparison

Figure 3 shows an example of time alignment for a 20-second interval

	Logtime	userdatetime	det_stn_id	slot_number
1	2015-01-02 00:00:14	2015-01-02 00:00:08	679	39
2	2015-01-02 00:00:34	2015-01-02 00:00:28	679	39
3	2015-01-02 00:00:54	2015-01-02 00:00:49	679	39
4	2015-01-02 00:01:14	2015-01-02 00:01:08	679	39
5	2015-01-02 00:01:34	2015-01-02 00:01:28	679	39
6	2015-01-02 00:01:54	2015-01-02 00:01:48	679	39
7	2015-01-02 00:02:14	2015-01-02 00:02:08	679	39
8	2015-01-02 00:02:34	2015-01-02 00:02:28	679	39
9	2015-01-02 00:02:54	2015-01-02 00:02:48	679	39
10	2015-01-02 00:03:14	2015-01-02 00:03:08	679	39

	record_time	det_stn_id	slot_number
1	2015-01-02 00:00:00	679	39
2	2015-01-02 00:00:20	679	39
3	2015-01-02 00:00:40	679	39
4	2015-01-02 00:01:00	679	39
5	2015-01-02 00:01:20	679	39
6	2015-01-02 00:01:40	679	39
7	2015-01-02 00:02:00	679	39
8	2015-01-02 00:02:20	679	39
9	2015-01-02 00:02:40	679	39
10	2015-01-02 00:03:00	679	39

Figure 3 Time Alignment for a 20-second interval

2) Checking Traffic Flow Variables

Volume, speed and occupancy are three traffic flow variables that are most important and useful in travel time and travel time reliability estimation. Therefore, how the quantity of these variables is checked is a significant step in DQA. For one data record, it is possible for the volume to be invalid or for the volume and speed to be invalid simultaneously; all three variables may even be missing completely. When processing these cases, one should know which case it belongs to. As illustrated in Figure 4, an encoding scheme was designed to indicate invalid variables.

x x x	-1	missing data record
0 0 0	0	valid data record
0 0 1	1	invalid occupancy
0 1 0	10	invalid speed
0 1 1	11	invalid speed and occupancy
1 0 0	100	invalid volume
1 0 1	101	invalid volume and occupancy
1 1 0	110	invalid volume and speed
1 1 1	111	invalid volume, speed, occupancy

Figure 4 Encoding Scheme in Data Checking step

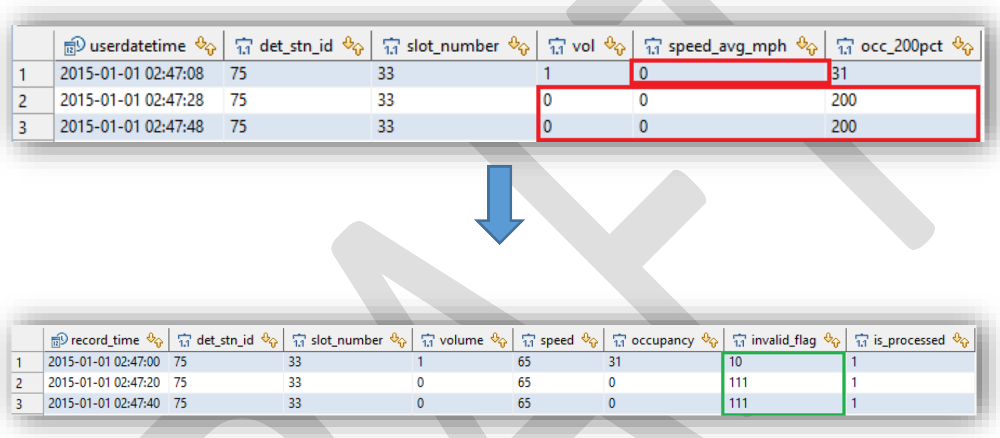
The whole traffic flow variable checking process is described as follows:

1) Extract useful fields from raw traffic flow data
2) Conduct time alignment
3) Check data quality and assign proper invalid flag for each data record according to checking rules, i.e., the volume in 20-second interval should be dropped into [0, 17], the speed over 20-

second interval should be dropped into [3, 100], the occupancy over 20-second interval should be dropped into [0, 0.95], and so forth.

4) Identify bad slots and discard them
5) Create new dynamic table to save data
6) Save checked data into the new created table

Figure 5 showed an example for checking 20-second data.



	userdatetime	det_stn_id	slot_number	vol	speed_avg_mph	occ_200pct
1	2015-01-01 02:47:08	75	33	1	0	31
2	2015-01-01 02:47:28	75	33	0	0	200
3	2015-01-01 02:47:48	75	33	0	0	200

	record_time	det_stn_id	slot_number	volume	speed	occupancy	invalid_flag	is_processed
1	2015-01-01 02:47:00	75	33	1	65	31	10	1
2	2015-01-01 02:47:20	75	33	0	65	0	111	1
3	2015-01-01 02:47:40	75	33	0	65	0	111	1

Figure 5 Data checking for 20-second data

d) Traffic Flow Data Imputation for 20-second data

After checking the traffic flow data, the next step is data imputation. The main purposes of this step are to impute missing data and correct erroneous data. As mentioned earlier in the proposed encoding scheme, there are nine cases that should be treated. To improve the reliability and robustness of the imputation system, multiple imputation schemes should be put forward to ensure finding at least one solution when conducting data imputation. We implemented six solutions for each case; the six imputation solutions for the case 001, where the occupancy variable is invalid or missing, is shown below. Other cases are conducted in a similar way.

The six imputation solutions for CASE 001 are described below, in descending order of priority:

1) Find valid occupancy records in the past 10 weeks at the same station and slot (lane type) with the same time (time of day) and with the same volume and speed values, and calculate the median of these occupancies to yield the imputation value;

2) Find valid occupancy records in the past 10 weeks at the same station and slot (lane type) with the same time (time of day), and calculate the median of these occupancies to yield the imputation value;
3) Find valid occupancy records in the past 10 weeks at the same station and slot (lane type) with the time between [current time - 10min, current time + 10min] and with the same volume and speed values, and calculate the median of these occupancies to yield the imputation value;
4) Find valid occupancy records in the past 10 weeks at the same station and slot (lane type) with the time between [current time - 10min, current time + 10min], and calculate the median of these occupancies to yield the imputation value;
5) Find valid occupancy records in the past 10 weeks at the same station and slot (lane type) with the time between [current time - 30min, current time + 30min], and calculate the median of these occupancies to yield the imputation value;
6) Find valid occupancy records in the past 10 weeks at the same station with the same time, and calculate the median of these occupancies to yield the imputation value;

The whole data imputation process is described as follows:

1) Query invalid data records by station, slot and invalid flag
2) Impute missing or erroneous data records according to invalid flags
3) Check if the imputed data is valid, if not, assign a default value
4) Update imputed data in batch mode

Figure 6 shows an example of data imputation for CASE 001.

	userdatetime	det_stn_id	slot_number	vol	speed_avg_mph	occ_200pct
1	2015-01-02 04:18:48	679	33	0	0	0
2	2015-01-02 04:18:48	679	37	1	69	2
3	2015-01-02 04:18:48	679	39	1	77	0

	record_time	det_stn_id	slot_number	volume	speed	occupancy	invalid_flag	is_processed
1	2015-01-02 04:18:40	679	39	1	77	1	1	1
2	2015-01-02 12:33:40	679	39	1	73	3	1	1
3	2015-01-02 13:24:00	679	37	1	83	7	1	1
4	2015-01-02 20:53:20	679	37	1	75	1	1	1
5	2015-01-02 16:59:00	679	33	1	73	3	1	1
6	2015-01-02 17:18:00	679	33	1	68	1	1	1

Figure 6 Data Imputation for CASE 001

e) Traffic Flow Data Temporal Aggregation for 20-second data

The final goal of SHRP project is to estimate the 5-minute travel time and travel time reliability, thus aggregating 20-second data to 5-minute date becomes a natural step after the data imputation step. Original raw 20-second data were recorded lane by lane so temporal aggregation must also be built on the lane-by-lane 20-second data. The temporal aggregation process is described as follows:

1) Query slots for each station of the given route
2) For each slot
    a. Query lane type
    b. Query 20-second traffic flow data
    c. Calculate time interval for each 5 minute
    d. Aggregate volume, speed, occupancy according to aggregation equations

$$Q_T = \sum_{i=1}^{n} q_i \times \frac{N}{n} \tag{1}$$

$$V_T = \frac{\sum_{i=1}^{n} q_i \times v_i}{\sum_{i=1}^{n} q_i} \tag{2}$$

$$O_T = median(o_i | i = 1, \dots, n) \tag{3}$$

    Where Q_T, V_T, O_T are the temporal aggregated volume, speed and occupancy, respectively. q_i, v_i, o_i are the valid volume, speed and occupancy over a 20-second time interval. N is the number of total observations in a 20-second time interval. n is the number of valid observations in a 20-second time interval.

    e. Update aggregated variables with aggregated values in batch mode

Figure 7 shows the temporal aggregation results over 20-second time intervals.

	record_time	det_stn_id	slot_number	volume	speed	occupancy	invalid_flag	is_processed
1	2015-01-01 00:00:00	675	39	0	65	0	0	0
2	2015-01-01 00:00:20	675	39	0	65	0	0	0
3	2015-01-01 00:00:40	675	39	0	65	0	0	0
4	2015-01-01 00:01:00	675	39	0	65	0	0	0
5	2015-01-01 00:01:20	675	39	0	65	0	0	0
6	2015-01-01 00:01:40	675	39	0	65	0	0	0
7	2015-01-01 00:02:00	675	39	0	65	0	0	0
8	2015-01-01 00:02:20	675	39	1	70	2	0	0
9	2015-01-01 00:02:40	675	39	0	65	0	0	0
10	2015-01-01 00:03:00	675	39	0	65	0	0	0
11	2015-01-01 00:03:20	675	39	0	65	0	0	0
12	2015-01-01 00:03:40	675	39	0	65	0	0	0
13	2015-01-01 00:04:00	675	39	0	65	0	0	0
14	2015-01-01 00:04:20	675	39	0	65	0	0	0
15	2015-01-01 00:04:40	675	39	0	65	0	0	0

	record_time	det_stn_id	slot_number	volume	speed	occupancy	valid_percent	invalid_flag	is_processed
1	2015-01-01 00:00:00	679	39	1	80	1	100	0	0
2	2015-01-01 00:05:00	679	39	2	71	1	100	0	0
3	2015-01-01 00:10:00	679	39	0	65	0	100	0	0
4	2015-01-01 00:15:00	679	39	2	74	1	100	0	0
5	2015-01-01 00:20:00	679	39	2	72	2	100	0	0

Figure 7 Temporal Aggregation over 20-second time intervals

f) Traffic Flow Data Spatial Aggregation for 20-second data

The purpose of spatial aggregation is to generate the aggregated results over the different lanes for a given detector station. In this project, 5-minute spatial aggregated traffic flow data were computed based on the prior 5-minute temporal aggregated data. It should be noted that the HOV lane should not be included, as this has different traffic flow characteristics from the other lanes. Hence, two groups were considered when conducting spatial aggregation, i.e. the HOV lane and the other aggregated lanes. The spatial aggregation process is described as follows:

1) Query stations for each route
2) For each station
    a. Query lane type
    b. Query 5-minute temporal aggregated traffic flow data
    c. Aggregate flow, speed, occupancy according to aggregation equations

Where Q_S, V_S, O_S are the spatial aggregated volume, speed and occupancy, respectively. q_j, v_j, o_j are the valid volume, speed and occupancy over 5-minute time interval. M is the number of total observations in 5-minute time interval. m is the number of valid observations in 5-minute time interval.

d. Update aggregated variables with aggregated values in batch mode

Figure 8 shows the spatial aggregation results over 5-minute time interval.

	record_time	det_stn_id	slot_number	volume	speed	occupancy	valid_percent	invalid_flag	is_processed
1	2015-01-01 00:00:00	679	39	1	80	1	100	0	0
2	2015-01-01 00:00:00	679	37	6	74	1	100	0	0
3	2015-01-01 00:00:00	679	35	12	68	3	100	0	0
4	2015-01-01 00:00:00	679	33	1	68	1	100	0	0

	record_time	det_stn_id	lane_type	volume	speed	occupancy	valid_percent	invalid_flag	is_processed
1	2015-01-01 00:00:00	679	1	1	80	1	100	0	0
2	2015-01-01 00:00:00	679	2	19	70	5	100	0	0

Figure 8 Spatial Aggregation over 5-minute data

g) Traffic Flow Data Checking for 5-minute data

Through the aggregation rule mentioned above, 5-minute temporal aggregated data and 5-minute spatial aggregated data can be obtained. However, since only valid observations were used to produce the aggregated data, there are still missing data and erroneous data in the new aggregated data. It is thus necessary to conduct traffic flow data checking over the 5-minute aggregated data. The basic checking rules for 5-minute temporal and spatial aggregated data are described as follows:

1) Whether the volume of the given data record drops into [0, 250 * lane_count];
2) Whether the speed of the given data record drops into [3, 90];
3) Whether the occupancy of the given data record drops into [0, 0.85];

Figure 9 and 10 present the checked results over 5-minute temporal and spatial aggregated data, respectively.

	record_time	det_stn_id	slot_number	volume	speed	occupancy	valid_percent	invalid_flag
1	2015-01-01 02:20:00	672	39	1	76	1	100	10
2	2015-01-01 07:05:00	85	41	1	56	1	100	10
3	2015-01-01 04:30:00	87	39	2	68	3	93.3	1

Figure 9 Traffic Flow Data Checking over 5-minute temporal aggregated data

	record_time	det_stn_id	lane_type	volume	speed	occupancy	valid_percent	invalid_flag
1	2015-01-01 02:20:00	672	1	1	76	1	0	10
2	2015-01-01 07:05:00	85	1	1	56	1	0	10
3	2015-01-01 04:30:00	87	1	2	68	3	0	1

Figure 10 Traffic Flow Data Checking over 5-minute spatial aggregated data

h) Traffic Flow Data Imputation for 5-minute data

As for the 20-second data checking, there are nine cases related to 9 encoding flags that must be treated in data imputation over 5-minute time interval. We again implemented six solutions for each case. The six imputation solutions for the case 001, i.e. with the occupancy variable invalid or missing, are shown below; other cases can be conducted in a similar way.

Six imputation solutions for CASE 001 are described in descending order of priority:

1) Find valid occupancy records in the past 10 weeks at the same station and slot (lane type) with the same time (time of day) and with the same volume and speed values, and calculate the median of these occupancies to yield the imputation value;
2) Find valid occupancy records in the past 10 weeks at the same station and slot (lane type) with the same time (time of day), and calculate the median of these occupancies to yield the imputation value;
3) Find valid occupancy records in the past 10 weeks at the same station and slot (lane type) with the time between [current time - 10min, current time + 10min] and with the same volume and speed values, and calculate the median of these occupancies to yield the imputation value;
4) Find valid occupancy records in the past 10 weeks at the same station and slot (lane type) with the time between [current time - 10min, current time + 10min], and calculate the median of these occupancies to yield the imputation value;
5) Find valid occupancy records in the past 10 weeks at the same station and slot (lane type) with the time between [current time - 30min, current time + 30min], and calculate the median of these occupancies to set it as the imputation value;
6) Find valid occupancy records in the past 10 weeks at the same station with the same time, and calculate the median of these occupancies to yield the imputation value;

The whole imputation process was implemented as for the 20-second imputation process.

i) Segment Travel Time Estimation for lane-by-lane 5-minute data

Travel time is the key index for SHRP2 project. It is also the input parameter for the method of travel time reliability estimation. Thus, estimating travel time index become the next step following the 5-minute data imputation. Generally, the process is composed of two steps, described as follows:

1) Calculate lane-by-lane road segments for each route
2) For each segment
    a. Query traffic flow data for start node
    b. Query traffic flow data for end node
    c. Estimate lane-by-lane average speed over 5-minute time interval according to the queried traffic flow data mentioned above
    d. Estimate travel time and travel rate by the computed average speed

$$travel\_time = \frac{segment\_length}{average\ speed} \quad (7)$$

    e. Estimate valid percent of data records using the following equation

$$valid\_percent = min(valid\_startflow, valid\_endflow) \quad (8)$$

Figure 11 shows the computed lane-by-lane segment information.

id	route_id	direction	start_stn_id	end_stn_id	start_slot_number	end_slot_number	segment_length	lane_descr
1	I-10_	E	684	683	39	39	0.93	HOV
2	I-10_	E	684	683	37	37	0.93	lane1
3	I-10_	E	684	683	35	35	0.93	lane2
4	I-10_	E	684	683	33	33	0.93	lane3
5	I-10_	E	683	685	39	39	0.99	HOV
6	I-10_	E	683	685	37	37	0.99	lane1
7	I-10_	E	683	685	35	35	0.99	lane2
8	I-10_	E	683	685	33	33	0.99	lane3
9	I-10_	E	685	686	39	39	1.06	HOV
10	I-10_	E	685	686	37	37	1.06	lane1

Figure 11 Segment Information for lane-by-lane 5-minute data

Figure 12 shows the computed travel time index for lane-by-lane 5-minute data.

record_time	segment_id	speed	travel_time	travel_rate	regime	valid_percent
2015-01-01 00:00:00	9	79.5	0.6	0.8	0	100
2015-01-01 00:05:00	9	66	0.8	0.9	0	100
2015-01-01 00:10:00	9	72	0.7	0.8	0	100
2015-01-01 00:15:00	9	73	0.7	0.8	0	100
2015-01-01 00:20:00	9	72	0.7	0.8	0	100
2015-01-01 00:25:00	9	71.5	0.7	0.8	0	100
2015-01-01 00:30:00	9	73	0.7	0.8	0	100
2015-01-01 00:35:00	9	73.5	0.7	0.8	0	100
2015-01-01 00:40:00	9	74.5	0.7	0.8	0	100

Figure 12 Travel Time Information for lane-by-lane 5-minute data

j) Segment Travel Time Estimation for aggregated 5-minute data

The process of computing segment travel time index for aggregated 5-minute data is the same as for the lane-by-lane 5-minute data. The process is described as follows:

1) Calculate road segments for each route
2) For each segment
    a. Query traffic flow data for start node
    b. Query traffic flow data for end node
    c. Estimate average speed over 5-minute time interval according to the queried traffic flow data mentioned above
    d. Estimate travel time and travel rate by the computed average speed

$$travel\_time = \frac{segment\_length}{average\ speed} \tag{9}$$

    e. Estimate valid percent of data records using the following equation

$$valid\_percent = min(valid\_startflow, valid\_endflow) \tag{10}$$

Figure 13 shows the computed segment information for aggregated 5-minute data.

id	route_id	direction	start_stn_id	end_stn_id	lane_type	segment_length
1	I-10_	E	684	683	1	0.93
2	I-10_	E	684	683	2	0.93
3	I-10_	E	683	685	1	0.99
4	I-10_	E	683	685	2	0.99
5	I-10_	E	685	686	1	1.06
6	I-10_	E	685	686	2	1.06
7	I-10_	E	686	679	1	0.96
8	I-10_	E	686	679	2	0.96
9	I-10_	E	679	672	1	0.83
10	I-10_	E	679	672	2	0.83

Figure 13 Segment Information for aggregated 5-minute data

Figure 14 shows the computed travel time index over aggregated 5-minute data.

record_time	segment_id	speed	travel_time	travel_rate	regime	valid_percent
2015-01-01 00:00:00	10	70.5	0.7	0.9	0	100
2015-01-01 00:05:00	10	68	0.7	0.9	0	100
2015-01-01 00:10:00	10	70.5	0.7	0.9	0	100
2015-01-01 00:15:00	10	69	0.7	0.9	0	100
2015-01-01 00:20:00	10	69.5	0.7	0.9	0	100
2015-01-01 00:25:00	10	69.5	0.7	0.9	0	100
2015-01-01 00:30:00	10	68.5	0.7	0.9	0	100
2015-01-01 00:35:00	10	69.5	0.7	0.9	0	100
2015-01-01 00:40:00	10	70	0.7	0.9	0	100
2015-01-01 00:45:00	10	68.5	0.7	0.9	0	100

Figure 14 Travel Time Information for aggregated 5-minute data