Integrated Environment for Performance Measurements and Assessment of Intelligent Transportation Systems Operations

Draft Final Report

Contract: BDK80 977-11

FIU Project: 800000446

Prepared for

The Florida Department of Transportation

by the

Florida International University Lehman Center for Transportation Research

July 2012

Disclaimer

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

Metric Conversion Chart

APPROXIMATE CONVERSIONS TO SI UNITS

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
NOTE: volumes greater than 1000 L shall be shown in m^3				
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
^{o}F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	^{o}C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Integrated Environment for Performance Measurements and Assessment of Intelligent Transportation Systems Operations		5. Report Date July 2012 6. Performing Organization Code
7. Author(s) Mohammed Hadi, Ph.D., PE; Yan Xiao, Ph.D.; Chengjun Zhan, Ph.D.; Patricio Alvarez, Ph.D.		8. Performing Organization Report No.
9. Performing Organization Name and Address Lehman Center for Transportation Research Florida International University 10555 W. Flagler Street, EC 3680, Miami, FL 33174		10. Work Unit No. (TRAIS) 11. Contract or Grant No. BDK80 977-11
12. Sponsoring Agency Name and Address Office of Research and Development State of Florida Department of Transportation 605 Suwannee Street, MS 30, Tallahassee, FL 32399-0450		13. Type of Report and Period Covered Draft Final Report June 2010 – September 2012 14. Sponsoring Agency Code
15. Supplementary Notes Mr. Mark Plass of the Florida Department of Transportation (FDOT) District 4 served as the project manager for this project. Mr. Vidya Mysore and Mr. Arun Krishnamurthy served as the co-project managers.		
16. Abstract This project has developed and implemented a software environment to utilize data collected by Traffic Management Centers (TMC) in Florida, in combination with data from other sources to support various applications. The environment allows capturing and fusing the data from multiple sources. The combined data can support the performance measurements of transportation system, transportation system modeling, assessment of the benefits of Intelligent Transportation System (ITS) applications such as incident management, ramp metering, and so on, and discovery of different relationships and associations of attributes through data mining and visualization methods. The developed modules are demonstrated in a series of use cases.		

17. Key Word	18. Distribution Statement
intelligent transportation systems, performance measurements, data mining, simulation support	Unrestricted

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	139	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Executive Summary

Intelligent transportation systems (ITS) are generating a wealth of data that can be used in combination with traffic analysis, simulation modeling, data fusion/data mining and optimization to support transportation system performance measurement, planning, operation, and management. The Florida Department of Transportation (FDOT) has recognized the importance of ITS data archiving and analysis and has invested in these areas. This has included the development and implementation of the advanced traffic management software, referred to as SunGuide, software that allows the collection, utilization, and archiving of detector and incident data. Another example is the development of a statewide ITS data warehouse.

The goal of this project is to develop and implement tools and methods for the utilization of data collected by Traffic Management Centers (TMCs) in combination with data from other sources to support various applications. The result of this project is an integrated ITS data analysis environment referred to as the ITS Data Capture and Performance Management Tool (ITSDCAP).

The specific objectives of this project are to develop tools and methods to support the following:

- Utilization of data from ITS and other sources to support performance measurements of transportation systems
- Provision of data from ITS and other sources to support transportation system modeling
- Development of assessment capabilities within ITSDCAP to assess the benefits of ITS applications based on ITS data
- Support the discovery of different relationships and associations of attributes, utilizing data mining and visualization methods
- Demonstration of the use of the developed environment.

Below is a description of the modules of the data analysis environment developed to satisfy the project objectives listed above.

Data Capture and Pre-Processing

The developed environment allows capturing and grouping data from multiple sources, including:

- SunGuide Traffic Sensor System(TSS) data at a 20-second aggregation level and lane-by-lane level
- Statewide ITS data warehouse data at 5-minute aggregation levels.
- SunGuide Dynamic Message Sign (DMS) activation data
- SunGuide incident management database
- Florida Highway Patrol (FHP) incident database
- INRIX data
- Work zone (construction) database
- Number of 5-1-1 number calls and Web site hits at an aggregation level of 15 minutes per corridor per county
- Dynamic pricing toll of managed lanes
- FDOT Crash Analysis Reporting (CAR) system
- Weather data from Road Weather Information Systems (RWIS) (provided through Excel file specified by the user)
- Weather data from national agencies (provided through Excel file)
- FDOT statistics database.

ITSDCAP allows the extraction and grouping of data based on different criteria for use in the analysis, including user-specified criteria or similarity in traffic patterns. It also allows fusing data from the above sources based on specific needs for integration using a common spatial and temporal referencing scheme. If the data comes from a source that does not check for data quality, the tool checks, filters, and repair missing data by imputation to ensure data quality.

Estimation of Performance Measures

A module is included in the developed environment to estimate various performance measures as discussed below.

Mobility Measures. ITSDCAP allows the estimation of seven mobility performance measures: speed, density, queue length/location, travel time, delay, vehicle-mile traveled (VMT), and vehicle-hour traveled (VHT). Alternative methods are incorporated to calculate different measures based on point detector data and INRIX data, allowing maximum flexibility in deriving these measures and assessing the accuracies of the estimates. For example, one of four different methods can be used to estimate travel time based on point detector data. Estimates are also possible of both the instantaneous travel time at the period of the estimation and the experienced travel time that accounts for traffic conditions as the vehicle progresses in its route from one link to the next.

Reliability Measures. ITSDCAP allows calculating different travel time reliability metrics including the standard deviation/variance, buffer index, failure/on-time performance, planning time index (PTI) based on the 95th or 80th percentile, skew statistics, and misery index. The investigation of this study found that the 95th percentile PTI, 80th percentile PTI, and misery index showed better continuity and sensitivity in their variations in response to the increase in variability, compared to other measures. The results from the study also confirmed that the use of at least a 30-minute period of analysis is preferable to using longer periods for applications that require fine-grained analysis in order to reasonably represent the reliability pattern during congested periods. In addition, study results confirmed that at least one year's worth of data should be collected to obtain more stable values of reliability metrics.

Safety Measures. ITSDCAP can estimate a number of safety measures, including crash frequency by crash type, crash frequency by severity, total crash frequency, crash rate by type, crash rate by severity, and total crash rate. In addition to the analysis based on the data stored in the Crash Analysis Reporting (CAR) database normally used for crash analysis, an initial exploration was made for potentially using crash data from the incident management databases, possibly in combination with the CAR system database in crash analysis. Although the incident databases collected by the TMCs may not have the details of crash attributes available in the CAR system database, they may be useful to complement the analysis under certain conditions.

Energy and Emission Measures. ITSDCAP currently calculates emission measures using MOBILE6.2 model parameters that are specific to Florida. These rates will be updated when emission rates that are specific to Florida are calculated based on the more recently released Motor Vehicles Emission Simulator (MOVES), which is the latest emission modeling system developed by the U.S. Environmental Protection Agency (USEPA). The results from a comparison conducted in this study show that MOVES estimates a higher emission increase due to lane blockage incidents compared to Mobile6.2, indicating that MOVES is more sensitive to reductions in speed.

Modeling Support

The modeling support module in ITSDCAP provides data to support different types of models, including demand forecasting, dynamic traffic assignment (DTA), macroscopic models, mesoscopic simulation models, and microscopic simulation. The module-provided data are extracted from different sources, processed as needed, and outputted in standard formats for use in various modeling activities. Four databases are used as the sources for the modeling support module: TSS data, detector data from the statewide data warehouse, travel time (TVT) data, and INRIX data.

The modeling support module has several functions that support the development, validation, and calibration of traffic models. These functions include estimating average speed, volume, and density for normal days or user-defined days, as well as calculating diurnal factor for demand modeling, estimating free-flow speed, estimating maximum throughput, estimating travel time, and fitting traffic flow models.

Map Visualization

Even though the estimated performance measurements and traffic flow parameters can be visualized through tabular or chart formats in other modules of ITSDCAP, it is useful to visualize these spatial-variant results on an ArcGIS map. Therefore, a map visualization module is provided in the developed environment.

ITS Benefit Estimation

Evaluating the benefits of ITS implementation is necessary for both planning and operational purposes. An ITS evaluation tool referred to as the Florida ITS Evaluation tool (FITSEVAL) has been developed in previous effort to evaluate ITS at the planning level in Florida. With the availability of rich ITS data and wide implementations of ITS, it becomes feasible to evaluate the impacts of ITS based on real-world data. Furthermore, impact factors based on ITS data can also be derived and applied to enhance the FITSEVAL evaluation accuracy. To achieve this, an ITS evaluation module was developed and incorporated in the developed environment to evaluate five types of freeway-related ITS implementations; including, incident management, ramp metering, managed lanes, smart work zone, and road weather information system.

Data Mining

The data mining module in ITSDCAP explores the relationship between different attributes of the data from one or more data sources. However, more advanced data mining analyses can be conducted by utilizing data outputted by ITSDCAP. This project demonstrated the use of the built-in Data Mining module in ITSDCAP, as well as using external data mining methods conducted outside ITSDCAP based on the data provided by the tool. The demonstration includes the association of incident attributes with DMS and 5-1-1 message attributes, matching of crashes recorded in the incident and crash databases, exploration of the measures used when selecting managed lanes, and the correlation of operational and safety measures.

Recommendations

The current version of the developed environment focuses on freeway facilities. It is recommended that the work be extended in the future to include data from signalized arterials, freight, transit, and connected vehicles. In addition, it is recommended that the tool be ported to a Web-based application to maximize the usability of the software.

Table of Contents

Disclaimer ii

Metric Conversion Chart iii

Executive Summary v

1. Introduction 1
   1.1. Background 1
   1.2. Project Objectives 2
   1.3. Project Activities and Report Organization 2
2. Development Requirements 5
3. Data Capture and Pre-Processing 8
   3.1. Data Requirements 8
   3.2. Data Description 9
      3.2.1. SunGuide Data 9
      3.2.2. FDOT Statewide Data Archive 10
      3.2.3. INRIX Data 10
      3.2.4. FDOT Transportation Statistics Office Data 11
      3.2.5. Express Lane Toll Rate Data 12
      3.2.6. Crash System Data 12
      3.2.7. FHP Incident Data 12
      3.2.8. 511 Data 12
      3.2.9. Weather Data 13
      3.2.10. Construction Data 13
   3.3. Data Acquisition 13
   3.4. Data Filtering and Cleaning 14
   3.5. Data Grouping 15
      3.5.1. Data Grouping Based on Attributes 15
      3.5.2. Time Series Data Grouping Based on Clustering Analysis 16
   3.6. References 19
4. Estimation of Performance Measures 20
   4.1. Mobility Measures 20

4.1.1. Estimation of Queue and Congestion Levels 21
4.1.2. Estimation of Travel Time 22
4.1.3. Estimation of VMT and VHT 26
4.1.4. Estimation of Density 26
4.1.5. Mobility Estimation User Interface 27
4.1.6. Mobility Measurement Test Cases 29
4.2. Travel Time Reliability Measures 30
4.2.1. Literature Review 30
4.2.2. Travel-Variant Travel Time Distributions 32
4.2.3. Travel-Variant Reliability Metrics 34
4.2.4. User Interface for Travel Time Reliability Measures 35
4.2.5. Reliability Measurement Test Cases 36
4.3. Safety Measures 44
4.3.1. Safety Measure Estimation Procedures 44
4.3.2. User Interface for Safety Measures 45
4.3.3. Safety Measurement Test Cases 49
4.4. Energy and Emission Measures 52
4.4.1. Emission Models Comparison 53
4.4.2. User Interface for Energy and Emission Measures 53
4.4.3. Emission Estimation Use Case 54
4.5. References 57
5. Modeling Support 60
5.1. Data Sources Used in Modeling Support Module 60
5.2. Modeling Support Module Functions 60
5.2.1. Average Speed, Volume, and Density 60
5.2.2. Diurnal Factor 62
5.2.3. Free-Flow Speed 62
5.2.4. Throughput 62
5.2.5. Average Travel Time 62
5.2.6. Fitting Traffic Flow Model 63
5.3. Case Studies 66

5.4. References 70
6. Map Visualization 72
6.1. Map Visualization Module Description 72
6.2. Examples of Map Visualization 73
7. ITS Evaluation 82
7.1. Incident Management 82
7.1.1. Incident Statistics 83
7.1.2. Incident Rate 84
7.1.3. Demand, Mobility Impacts, and Secondary Incident Probability for Individual Incidents 84
7.1.4. Benefit/Cost Analysis 85
7.2. Ramp Metering 87
7.2.1. Average Speed, Travel Time, and Maximum Throughput 88
7.2.2. Vehicle-Mile Traveled (VMT) and Vehicle-Hour Traveled (VHT) 89
7.2.3. Travel Time Reliability 89
7.2.4. Safety 89
7.2.5. Fuel Consumption and Emissions 90
7.3. Managed Lanes 90
7.4. Smart Work Zone 92
7.5. Road Weather Information System 93
7.6. Case Studies 95
7.6.1. Determine the Incident Recovery Ending Time 95
7.6.2. Comparison between FITSEVAL and ITSDCAP for Incident Management 97
7.6.3. Ramp Metering and Express Lane Evaluation 101
7.7. References 103
8. Data Mining 106
8.1. Data Mining Procedures in ITSDCAP 106
8.2. Use Cases of the Built-In Data Mining Module 109
8.3. Comparison of Crash Data in the CAR and Incident Database 114
8.4. User Choice of I-95 Express 116

8.4.1. Historical Information Considered by Users 117
8.4.2. Exploration of Performance Measures that Impact HOT Lane Choices 120

List of Figures

Figure 2-1 High Level Design of ITSDCAP 7
Figure 3-1 Data Acquisition Interface in the ITSDCAP Tool 14
Figure 3-2 ITSDCAP Interface for Data Grouping Based on Attributes 16
Figure 3-3 ITSDCAP Interface for Time Series Data Grouping 18
Figure 3-4 Example of Clustering Results 19
Figure 4-1 Clustering Results for a Detection Station on SR-826 22
Figure 4-2 Schematic Diagram of Detector Configuration 22
Figure 4-3 Schematic Diagram for Off-Line Travel Time Estimation 25
Figure 4-4 Interface for Mobility Performance Estimation 27
Figure 4-5 Starting and Ending Location Selection for Mobility Performance Estimation 28
Figure 4-6 Mobility Performance Estimation Visualization 29
Figure 4-7 User Interface for Travel Time Reliability Measure Calculation 36
Figure 4-8 Reliability Metrics Variation with Time-of-Day GP Lanes 40
Figure 4-9 Reliability Metrics Variation with Time-of-Day HOT Lanes 40
Figure 4-10 Impact of Aggregation Levels on the 80th Percentile PTI Metric 42
Figure 4-11 Effect of Data Collection Period Length 43
Figure 4-12 User Interface for Safety Measure Calculation 46
Figure 4-13 Roadway Segment Selection for Safety Measure Calculation 47
Figure 4-14 Study Period Selection for Safety Measure Calculation 48
Figure 4-15 Example of Segment Information File for Safety Measure Calculation 49
Figure 4-16 Minimum Average Speed for AM Peak Period 50
Figure 4-17 Minimum Average Speed for Each Hour in AM Peak Period 51
Figure 4-18 Travel Time Variation during the AM Peak Period 51
Figure 4-19 VMT Variation during the AM Peak Period 52
Figure 4-20 VHT Variation during the AM Peak Period 52
Figure 4-21 User Interface for Energy and Emission Measure Calculation 54
Figure 4-22 Queue Formation and Dissipation for Incident on 1/18/2011 55

Figure 4-23 Additional Emissions due to the Incident....56
Figure 5-1 Modeling Support Module Interface in ITSDCAP....61
Figure 5-2 Traffic Flow Model Selection....66
Figure 5-3 Average Normal Day Speed and Volume....67
Figure 5-4 Diurnal Factor....68
Figure 5-5 Free-Flow Speed....69
Figure 5-6 Average Normal Day Travel Time Based on INRIX Data....69
Figure 5-7 Traffic Flow Model Fitting Results....70
Figure 6-1 Input Interface for Map Visualization Case Study 1....74
Figure 6-2 Display Interface for Map Visualization Case Study 1....75
Figure 6-3 Color Settings for Map Visualization Case Study 1....76
Figure 6-4 Chart Display for Map Visualization Case Study 1....77
Figure 6-5 Input Interface for Map Visualization Case Study 2....78
Figure 6-6 Display Interface for Map Visualization Case Study 2....79
Figure 6-7 Input Interface for Map Visualization Case Study 3....80
Figure 6-8 Display Interface for Map Visualization Case Study 3....81
Figure 7-1 Incident Management Evaluation Interface....83
Figure 7-2 Ramp Metering Evaluation Interface....88
Figure 7-3 Managed Lanes Evaluation Interface....91
Figure 7-4 Smart Work Zone Evaluation Interface....93
Figure 7-5 Example of Work Zone Input Information....93
Figure 7-6 Example of Work Zone Input Information....95
Figure 7-7 Example of Incident Recovery Ending Time....96
Figure 7-8 Study Corridor Utilized in Incident Mobility Impacts Assessments....98
Figure 7-9 Estimated Parameters from ITSDCAP (Continued on Next Page)....98
Figure 7-9 Estimated Parameters from ITSDCAP....99
Figure 7-10 Comparison of Incident Delay in Dollar Values....101
Figure 7-11 Map of Study Corridor for RM and HOT Evaluation....102
Figure 7-12 Maximum Throughput Results for I95SB between Golden Glades Interchange and SR 836....103
Figure 8-1 Data Mining Interface in ITSDCAP....107

Figure 8-2 Data Mining Case Study Results for Crash Data 109
Figure 8-3 Data Mining Case Study Results for Incident Data 110
Figure 8-4 Data Mining Case Study Results for Incident and DMS Data 113
Figure 8-5 Relationship between Average 511 Calls and Number of Travel Lanes Blocked 114
Figure 8-6 R^2 Considering Different Dataset Lengths in AM and PM Periods 118
Figure 8-7 R^2 Considering Different Fading Memory Rules 119
Figure 8-8 R^2 Considering Different Real-Time Information Parameter a 120

List of Tables

Table 4-1 Travel Time Reliability Operational Definitions 34
Table 5-1 Dual-Regime Modified Greenshields Model Fitting Results in Case Study 5 . 70
Table 7-1 Comparison of IM Evaluation Results between FITSEVAL and ITSDCAP . 100
Table 8-1 Data Source and Associated Attributes Considered in the Built-In Data Mining Module 108
Table 8-2 Total Possible Matches 115
Table 8-3 Typical Correlation Factors 121
Table 8-4 Results of Logistic Regression 121

List of Selected Acronyms and Abbreviations

Acronyms and abbreviations used in the report are listed below.

ATIS	Advanced Traveler Information System
CAR	Crash Analysis Reporting System
CSV	Comma-Separated Value
DLL	Dynamic Link Library
DMS	Dynamic Message Sign
DTA	Dynamic Traffic Assignment
FDOT	Florida Department of Transportation
EPA	Environmental Protection Agency
FHP	Florida Highway Patrol
FHWA	Federal Highway Administration
GIS	Geographic Information System
GUI	Graphic User Interface
HCM	Highway Capacity Manual
HOV	High Occupancy Vehicle
ITS	Intelligent Transportation System
Matlab	A numerical computing environment developed by MathWorks
RWIS	Road Weather Information System
SIRV	Severe Incident Response Vehicle
STEWARD	Statewide Transportation Engineering Warehouse for Archived Regional Data
TMC	Traffic Management Center
TSS	Traffic Sensor System
TVT	Travel Time

1. Introduction

1.1. Background

Intelligent Transportation Systems (ITS) are generating a wealth of data that can be used in combination with traffic analysis, simulation modeling, data fusion/data mining, and optimization to support transportation system performance measurement, planning, operation, and management. The archiving and analysis of data using advanced computational techniques will provide the information necessary to measure the performance of the systems, assess the impacts of alternative strategies and technologies, and provide decision support for transportation agencies. These developments are expected to result in significant improvements in transportation system performance.

The Florida Department of Transportation (FDOT) has recognized the importance of ITS data archiving and analyses and has invested in these areas. This has included the development and implementation of an advanced traffic management software, referred to as the SunGuide software, that allows the collection, utilization, and archiving of detector and incident data. Another example is the development of a statewide ITS data warehouse. In addition, FDOT recognizes the importance of assessing system performance and estimating the benefits of advanced strategies. A tool, referred to as the Florida ITS evaluation tool (FITSEVAL), has been recently developed to evaluate ITS deployments in Florida. The tool evaluates ITS at the planning level and is suitable for assessing ITS as part of short-term and long-term transportation plans.

There is a need to develop and implement tools and methods to utilize the data collected by Traffic Management Centers (TMCs) in combination with data from other sources for various purposes. This document is the final report for a project funded by FDOT for the development and use of an integrated ITS data analysis tool referred to as ITS Data Capture and Performance Management Tool (ITSDCAP). The development of this tool was built on the FITSEVAL development effort and a number of other research efforts conducted for the archiving and utilization of ITS data.

1.2. Project Objectives

The goal of this project is to develop and implement tools and methods to utilize data collected by TMCs in combination with data from other sources to support various applications. The specific objectives of this project are to develop tools and methods to support the following functionalities:

- Utilization of data from ITS and other sources to support performance measurements of transportation system
- Provision of data from ITS and other sources to support transportation system modeling
- Development of an environment that can be used to assess the benefits of ITS applications based on ITS data
- Support the discovery of different relationships and associations of attributes utilizing data mining and visualization methods
- Demonstration of the use of the developed environment.

It should be noted that the current version of the developed environment focuses on freeway facilities. It is recommended that the work is extended in the future to include data from signalized arterials, freight, transit, and connected vehicles.

1.3. Project Activities and Report Organization

To satisfy the project objectives listed above, a tool referred to as ITS Data Capture and Performance Management (ITSDCAP) was developed in this project. This section presents an overview of the project activities and associates these activities with the chapters of this report.

- Development Requirements. This task was conducted at the beginning of this project to identify high level requirements that are traceable to the project objectives listed earlier. The requirements were discussed in a stakeholder workshop and further refined based on the results of the workshop. A high level design was produced based on these requirements. The requirements and the high level design are presented in Chapter 2.

- Data Capture and Pre-Processing. The first module of the developed ITSDCAP tool allows the capture, cleaning, aggregation, and grouping of data from multiple sources and integrating them into the ITSDCAP tool. Chapter 3 provides a description of this module.
- Estimation of Performance Measures. To facilitate the estimation of the performance of transportation systems, a module was developed in the ITSDCAP tool to estimate various mobility, reliability, safety, and environmental measures of the transportation system. Chapter 4 provides a detailed description of this module.
- Modeling Support. The Modeling Support module in ITSDCAP was designed to provide data to support different types of models; including, demand forecasting, dynamic traffic assignment (DTA), macroscopic models, mesoscopic simulation models, and microscopic simulation. This module is discussed in Chapter 5.
- Visualization. Even though the estimated performance measurements and traffic flow parameters can be visualized through table or chart formats in other modules of ITSDCAP, it is useful to visualize these spatial-related results on an ArcGIS map. Therefore, a "Map Visualization" module is provided in ITSDCAP and is described in Chapter 6 of this document.
- ITS Benefits. Evaluating the benefits of ITS implementation is necessary for both planning and operational purposes. With the availability of rich ITS data and wide implementations of ITS, it becomes feasible to evaluate the impacts of ITS based on real-world data. Furthermore, impact factors based on ITS data can also be derived and applied to enhance the FITSEVAL evaluation accuracy. To achieve this, an ITS Evaluation module was developed and incorporated in the developed environment to evaluate five types of freeway-related ITS implementations; including, incident management, ramp metering, managed lanes, smart work zone, and road weather information system. This development is discussed in Chapter 7.
- Data Mining. The data mining module in ITSDCAP explores the relationship between different attributes within the same data source or among the attributes from different data sources. Chapter 8 discusses the Data Mining module in ITSDCAP is discussed. In addition, more advanced data mining analysis conducted outside ITSDCAP in this project

based on the data provided by the tool is presented to illustrate the benefit of combining data from multiple sources in the analysis.

2. Development Requirements

This task was conducted at the beginning of this project to identify high level requirements that are traceable to the project objectives listed in Chapter 1. The requirements were discussed in a stakeholder workshop and refined based on the results of the workshop. Follow-up meetings were conducted with stakeholders that were not able to attend the workshop.

The core project stakeholders that provided inputs to this effort included representatives from the following FDOT offices:

- FDOT District 4 and 6 traffic operations/ITS engineers
- FDOT District 4 traffic safety engineer
- FDOT traffic demand forecasting staff
- FDOT Central Office System Planning
- FDOT central office ITS section.

A list of needs and issues that are related to the project objectives have been identified based on the project workshop and interviews with project stakeholders. These needs are listed below.

1. Data are becoming available from multiple sources including ITS. There is a need for the capture and provision of integrated data from these sources to support transportation agency planning, operation, and research.
2. Data from different sources has different location referencing, quality, and aggregation levels. The data from these sources need to be cleaned, aggregated, and grouped to allow their use in various needed applications.
3. The increasing emphasis on performance measurement and management has increased the need for the development of procedures to utilize data to assess the performance of transportation systems; including, mobility, reliability, safety, and environmental measures.
4. The transportation planning and traffic operation community have used various types of models such as traffic demand forecasting, dynamic traffic assignment (DTA), analytical

traffic analysis procedures, and simulation models. These models require significant amounts of data that are expensive to collect, reduce, and maintain. There is a need for tools and procedures to allow the provision of required inputs to these models based on data collected from ITS devices and other data sources.

5. With the increasing competition for resources, transportation agencies have increasingly realized the importance of the assessment of the benefits and costs of operation strategies and technologies. The availability of integrated data from multiple sources combined with advanced modeling techniques will support this assessment. FITSEVAL can be used to assess the benefits of ITS at the planning level. Although it includes default parameters, the recommendation is to update these parameters based on real-world data to allow more accurate analysis of the benefits. In addition to the need to provide input parameters to FITSEVAL based on data, as mentioned above, the utilization of ITS data will allow calculating the benefits outside FITSEVAL based on real-world data at the operational level rather than based on modeling as is the case with FITSEVAL.
6. Advanced mining techniques need to be developed and used to allow the discovery of patterns in the data, allowing the development of advanced decision support tools for use by transportation agencies.
7. Visualization of performance measures is needed to better understand the available data and calculated performance measure values.

The ITSDCAP tool was developed in this study to satisfy the above need. Figure 2-1 shows the high level design of the tool.

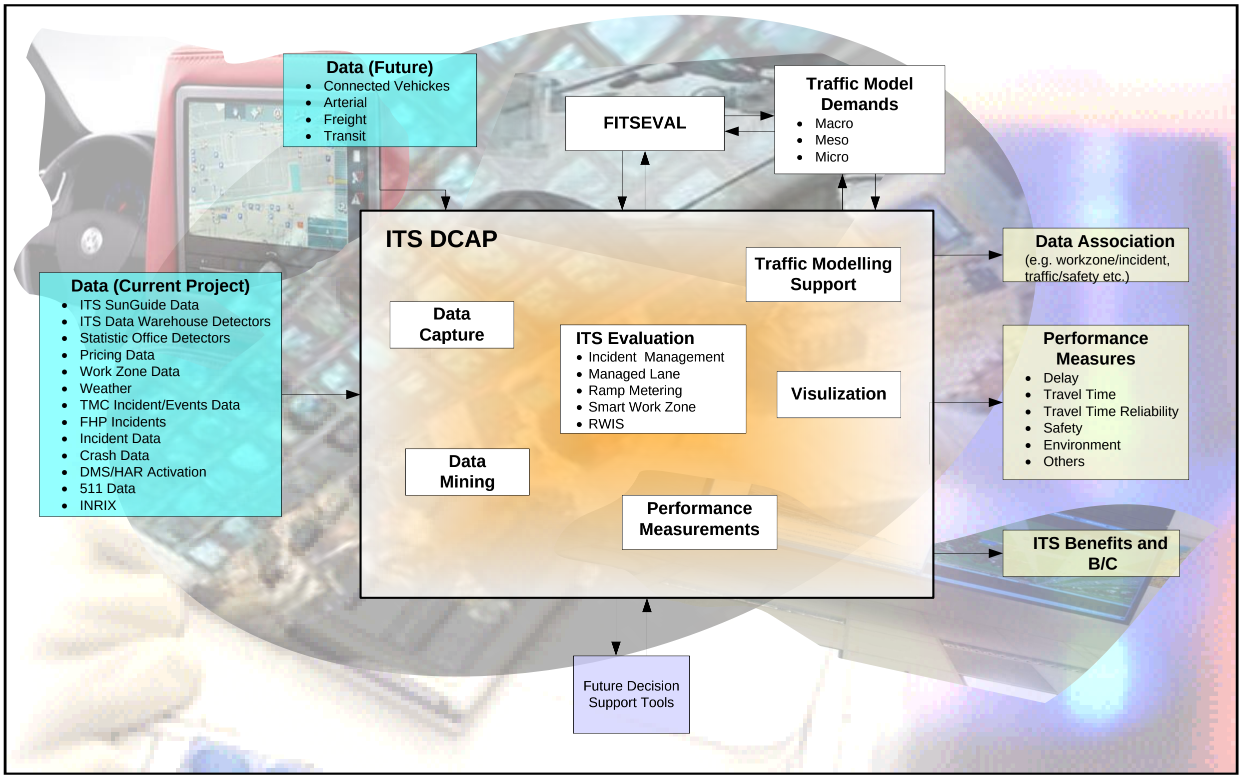


Figure 2-1 High Level Design of ITSDCAP

3. Data Capture and Pre-Processing

The first module of the developed ITSDCAP tool allows the capture, cleaning, aggregation, and grouping of data from multiple sources and integrating them into the ITSDCAP tool. This chapter provides a description of this module.

3.1. Data Requirements

Based on the identified needs, it was determined that ITSDCAP shall be able acquire data from different sources to use in the developed modules. Data from the following sources that can be integrated and used in the current version of the tool are:

- SunGuide TSS traffic detector data at a 20 second aggregation level and lane-by-lane level
- Statewide ITS data warehouse data at 5 minute aggregation levels
- SunGuide DMS activation data
- SunGuide incident management database
- FHP incident database
- INRIX data
- Work zone (construction) database
- 511 number of calls and web site hits at an aggregation level of 15 minutes per corridor per county
- Dynamic pricing toll of managed lanes
- FDOT Crash Analysis Reporting (CAR) system
- Weather data from Road Weather Information Systems (RWIS) (provided through Excel file specified by the user)
- Weather data from national agencies (provided through Excel file)
- FDOT statistics database.

ITSDCAP shall allow fusing data from the above sources based on specific needs for integration using a common spatial and temporal referencing scheme. If the data comes from a source that does not check for data quality, ITSDCAP shall check, filter and impute data, as needed. In addition, ITSDCAP shall allow the extraction and grouping of data based on different criteria for use in the analysis including user-specified criteria or similarity in traffic patterns.

3.2. Data Description

Data from various sources can be captured utilizing the ITSDCAP tool developed in this project. This section includes a description of the data.

3.2.1. SunGuide Data

The SunGuide TMC software is being used by all FDOT districts and the central office for the control of roadway devices and information exchange across transportation agencies. Although the SunGuide system architecture is “data bus” centric instead of “database” centric, a lot of information is stored in a back-end Oracle database and/or text format data archive files that can be used for visualization and decision support purposes, either in “online” (real-time) or “offline” formats. Several data items are captured and utilized ITSDCAP as listed below.

- **TSS Data.** TSS data is traffic measurements collected by the Traffic Sensor System (TSS) of SunGuide. In the SunGuide system, the aggregated TSS data are stored in Oracle database for report generation, while the raw data are saved in the TSS text file in comma separated file (csv file) format. The TSS file contains one record per lane for each detection station at a 20-second polling interval. Each TSS data record includes the following information: timestamp, detection station name, lane number, speed, occupancy, and raw count. Each TSS file contains data for a 24-hour period (midnight to midnight) for each day.
- **TVT Data.** The traffic data collected by TSS are used by SunGuide to compute the travel times for predefined travel time links (TVT links). The travel time information is then displayed on Dynamic Message Signs (DMS), used for operator displays, and saved in TVT text files as data archives. Similar to the TSS file, the information in TVT files is comma separated. This information includes the timestamp that the travel time is posted, travel link ID, the timestamp that the travel time is calculated, the travel time and delay after rounding, and raw travel time and delay data.

- **Incident Data.** Detailed incident information is stored in the SunGuide Oracle database, including incident timestamps (detection, notification, arrivals, and departures), incident ID, responding agencies, event details, chronicle of the event, and environmental information. The detection timestamp is the time when an incident is reported to the TMC and inputted in the SunGuide system. The notification timestamps are recorded per responding agency and refer to the time when such responding agencies are notified. The arrival and departure timestamps are also recorded per responding agency and refer to the time when responding agencies arrive and depart from the incident site.
- **DMS Data.** When incidents occur and corresponding incident alerts posted on the upstream DMS, information is recorded in this data archive regarding DMS activation timestamp, termination timestamp, and text messages.

3.2.2. FDOT Statewide Data Archive

The Statewide Transportation Engineering Warehouse for Archived Regional Data (STEWARD) has been developed as a proof of concept prototype for the collection and use of ITS data in Florida (Courage and Lee, 2008). The STEWARD data warehouse retrieves point traffic detector data from district TMCs on a daily basis, and then processes and archived these data in the data warehouse. It contains summaries of traffic volumes, speeds, and occupancies at the aggregation levels of 5, 15, 60 minutes. The FDOT is currently making arrangements for an FDOT ITS data warehouse maintained by the University of Maryland.

3.2.3. INRIX Data

INRIX utilizes what is sometimes referred to as crowd sourcing of GPS probe data, in which data from multiple sources are collected and fused to provide travel time estimates that are then distributed to public and private sector users. INRIX technology is based on research and development originally by Microsoft Research. It utilizes statistical analysis to estimate travel time based on data collected from multiple sources including commercial fleet, delivery and taxi vehicles across the U.S., as well as consumer cellular GPS-based devices including smartphones and onboard driver assistance systems, sometimes combined with traditional real-time traffic flow information. INRIX also utilizes information about events that affect traffic, such as construction and road closures, real-time incidents, sporting and entertainment events, weather forecasts, and school schedules. The data from the above sources are fused using the INRIX fusion engine to produce travel time estimates. The density of GPS-based data used as the basis of the INRIX GPS feed continues to grow, and thus it is expected that the quality of the data will improve over time.

In this project, INRIX data are acquired through the I-95 Corridor Coalition system and include information about roadway segment number, timestamp of measurement, measured speed, average historical speed, free-flow speed, travel time, and confidence level and scores associated with data fidelity.

3.2.4. FDOT Transportation Statistics Office Data

The FDOT Transportation Statistics Office (TranStats) collects traffic data through various telemetered and portable traffic monitoring sites along all Florida state highways. It reports the data of annual average daily traffic (AADT), peak hour factors, directional distribution factors, and truck factors. Depending on locations, it may also report the daily traffic count, speed, and vehicle classification at the 15-minute aggregation level.

The FDOT Statistics Office traffic monitoring sites (TMS) are classified into permanent (telemetered) and portable (temporary) categories. Telemetered traffic monitoring sites (TTMS) refer to traffic counters that are permanently placed at specific locations throughout Florida to record the distribution and variation of traffic flow by hour of the day, day of the week, and month of the year, from year-to-year, and transmit the data to the TranStat Office in Tallahassee at the end of each day via telephone lines. These sites record traffic volumes 24 hours a day and seven days a week. They provide counts classified by vehicle type (13 FHWA classification category scheme) per lane and per direction of travel. They also provide the average speed per lane. The data is usually collected at one-hour intervals. In Jacksonville, most of the stations provide only hourly volumes. Overall, there are about 320 statewide TTMS count locations in Florida.

The second type of monitoring sites maintained by the FDOT central office is the portable traffic monitoring site (PTMS). A PTMS site is a traffic monitoring site that has loops and/or axle sensors in the roadway, with leads running back into a cabinet located on the shoulder. When a traffic count is desired, a portable counter is connected to the sensor leads and placed in the cabinet. After the count has been collected, the counter is removed and placed at another count site. Counts are normally taken for a minimum of 48 hours per year between Monday 6:00 AM and Friday 2:00 PM. PTMS provide only 15-minute volume data. No speed or classification data are provided.

3.2.5. Express Lane Toll Rate Data

The dynamic toll rate information for I-95 Express Lane as archived by FDOT District 6 was also integrated with ITSDCAP. The toll rates for both the northbound and southbound directions of I-95 Express between Dec. 16, 2008 and Dec. 12, 2010 were acquired and used in this project.

3.2.6. Crash System Data

One source of the crash data used in this project is originated from the Crash Analysis Reporting (CAR) System maintained by the Florida Department of Transportation. In Florida, all crashes that occur on state roads and result in a fatality, an injury, or a property-damage-only (PDO) higher than $1,000 are included in the CAR System, and the data in this system is updated yearly. Based on the police report, 38 data elements for crashes are recorded in the CAR System, including crash location, time stamp, property damage dollar value, injury, fatality, pavement conditions, weather and lighting conditions, and crash cause.

3.2.7. FHP Incident Data

The Florida Highway Patrol (FHP) data for Miami-Dade County and Broward County are accessed through the Signal Four Analytics, a traffic crash database environment developed for the FHP. Currently, this program gathers information from FHP reports on a daily basis. The information of crash occurrence time and the FHP response timeline are archived in the database.

3.2.8. 511 Data

The 511 data for the I-95 corridor in Miami-Dade County and Broward County are acquired from the contractor that operates the Florida Advanced Traveler Information System (FLATIS). It contains the number of 511 calls for this corridor that aggregated at a 15 minute interval.

3.2.9. Weather Data

Seven Road Weather Information System (RWIS) monitoring sites have been deployed along the I-95 corridor by the FDOT District 4. These sites measure the road surface status, air and surface temperature, relative humidity, precipitation type, intensity and rate, visibility distance, wind speed and direction, and barometric pressure. The reported RWIS data were downloaded from the FDOT District 4 TMC SCAN Web webpage.

3.2.10. Construction Data

It is anticipated that this data will be imported into ITSDCAP when the Lane Closure Information database of FDOT District 4 becomes available.

3.3. Data Acquisition

Data from the different data sources mentioned earlier can be integrated to the ITSDCAP tool through the data acquisition tab page, as shown in Figure 2-1. Note that in order to associate the data from different sources, all the data should have the same location reference system. In this project, the roadway milepost is used as the reference system, and if a database has another reference system (such as longitude and latitude coordinators), the locations were converted first to mileposts. The users can upload data from several data sources to the ITSDCAP Oracle database. However, the TSS and Incident data, because they are very large data files, are accessed directly by ITSDCAP without importing to the Oracle database. Data from some sources are currently not imported to the database and accessed directly utilizing their original file formats. The developed ITSDCAP tool also allows the users to check the data availability. Any missing data will be reported in the data availability check results.

Figure 3-1 shows how the user can upload the different data items. As can be seen, the user can upload data from ITS data warehouse, CAR system, weather database, and construction lane blockage database to the ITSDCAP in the lower half of the user interface shown in Figure 3-1. In the upper half, the user can check whether each type of data has already been uploaded and available for use for a specific period and corridor. ITSDCAP provides a list of dates with missing information between the beginning and ending date to alert the user that they may need to upload these additional days.

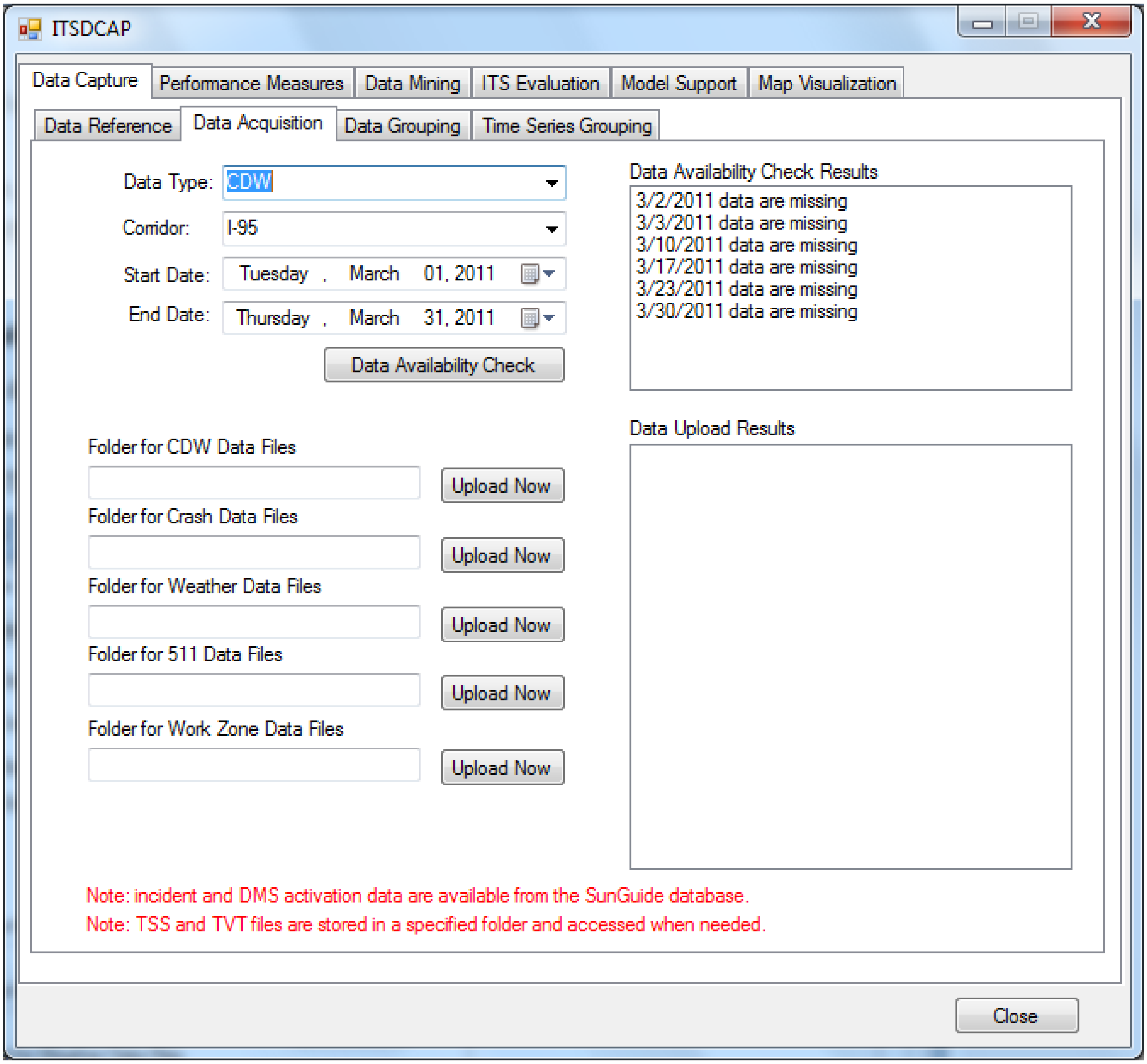


Figure 3-1 Data Acquisition Interface in the ITSDCAP Tool

3.4. Data Filtering and Cleaning

Most of the collected data such as those from the statewide data warehouse and INRIX data are already filtered and cleaned. However, the data in the TSS file (the raw detector data at the 20-sec. aggregation level) usually includes erroneous or missing measurements. Thus, data pre-processing procedures are first applied in the ITSDCAP tool before utilizing TSS data. These procedures include data filtering, temporal and spatial data aggregation, and data imputation. In the data filtering step, rule-based tests are applied to two levels of detector data: raw 20-second data and temporal aggregated data. The 20-second level tests include elimination of duplicate measurements, univariate tests for individual traffic parameters, multivariate tests for unreasonable combinations of traffic parameters, and also a temporal variability check for constant values of measurements. The checks for average effective vehicle length and maximum density are conducted at the temporal aggregated level.

The filtered 20-second lane-by-lane detector data are then aggregated to the station level for use in later steps. Since detector data may be missing due to detector failures not reporting any data or erroneous measurements identified in the previous data filtering step, spatial and temporal data imputation procedures are applied in ITSDCAP to substitute these missing values. The spatial imputation employs neighboring detector information to impute the missing values while the temporal imputation uses the past information from the same detector to replace the missing values.

3.5. Data Grouping

In most analyses, the users are expected to require performing the analysis on a subset of the data rather than analyzing all the data in the database. The purpose of the data grouping sub-model is to classify the data into groups based on certain criteria for later applications. In ITSDCAP, two types of grouping are possible. The first is based on ranges of attributes specified by the user and the second is based on the similarity in traffic patterns automatically identified utilizing clustering analysis.

3.5.1. Data Grouping Based on Attributes

The first data grouping method developed in this project is to group the data based on user specified criteria since most of the applications require the extraction of data for a specific time period and corridor segment. As shown in Figure 3-2, the selection criteria can be based on combinations of parameters such as a specific corridor for both directions or one direction, starting and ending date and time, day of week, day type (normal or incident days), and data sources for grouping. When the button “Download Data” is clicked, the user-requested data will be downloaded to a local file with a data group name as input, based on the criteria specified by users. The grouping results will indicate that the data downloading is finished. Clicking the "Export Group" button can save users' grouping criteria input for later use. The option of "Grouping using Time Series Clustering Results" allows users to further classify the time series clustering results into subgroups using database attributes.

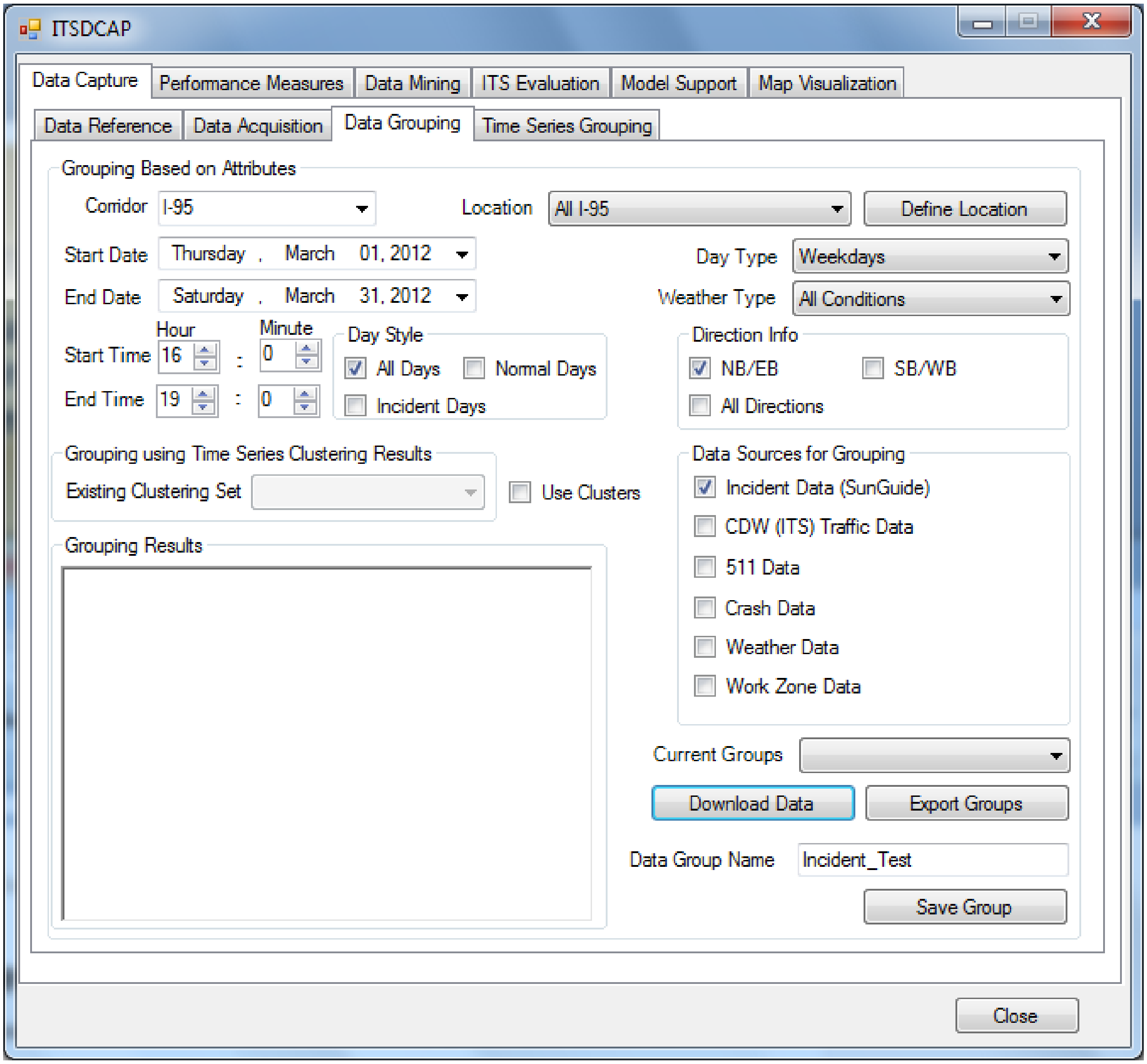


Figure 3-2 ITSDCAP Interface for Data Grouping Based on Attributes

3.5.2. Time Series Data Grouping Based on Clustering Analysis

The second data grouping method is to classify the traffic data into different patterns based on the similarity between the time series of volume counts on multiple days. This method was originally developed by Hadi et al. (2011) and can be used to isolate normal day traffic patterns from days with incidents and/or special events. The basic algorithm behind this method is the *k*-means clustering analysis, which minimizes the sum of time series distances to cluster centroids while maximizing the distance between clusters. The Euclidian distance is used to quantify the time series distance, as shown in Equation 3-1 and 3-2.

$$dist((v_j, c_k) = \sum_i (v_j(t_i) - c_k(t_i))^2 \qquad \forall j \in k \tag{3-1}$$

$$c_k(t_i) = \frac{1}{n_k} \sum_j v_j(t_i), \forall j \in k \tag{3-2}$$

where $v_j(t_i)$ is the j time series measurement at time interval i, and $c_k(t_i)$ is the centroid of cluster k at the same time interval. The term n_k denotes the total number of time series in cluster k. Depending on the starting point (selected stochastically in each run), the optimization routine used in the clustering analysis reaches a local optimal that may vary with the runs. Thus, the algorithm is replicated for ten times to associate the daily traffic counts with the clusters.

Figure 3-3 illustrates the ITSDCAP interface for the time series data grouping. As shown in this figure, either single or multiple detector stations can be selected. The volume data for the selected stations during the time period specified by the user will be used in the clustering analysis. It is also shown in this figure that the number of patterns used in analysis can be changed using the interface. As an example, the traffic data for station 600641 located at the south of NW 103rd St. along the I-95 Corridor in Miami-Dade County between Jan. 1, 2011 and Mar. 31, 2011, were clustered into seven patterns, as shown in Figure 3-4. It is seen that Patterns 1 and 2 are for days with detector malfunction. Patterns 3 and 6 are clearly for incident days. Patterns 5 and 7 are weekend traffic. The remaining Pattern 4 represents the normal day traffic at this detection station.

Note that two clustering modes are listed in the interface: “Initial Clustering” and “Re-Clustering.” The “Initial Clustering” starts the clustering analysis using all the volume count data, while the “Re-Clustering” mode allows the analysis to start with previous clustering results for additional clustering as needed. The clustering results can be visualized through the list box in the interface.

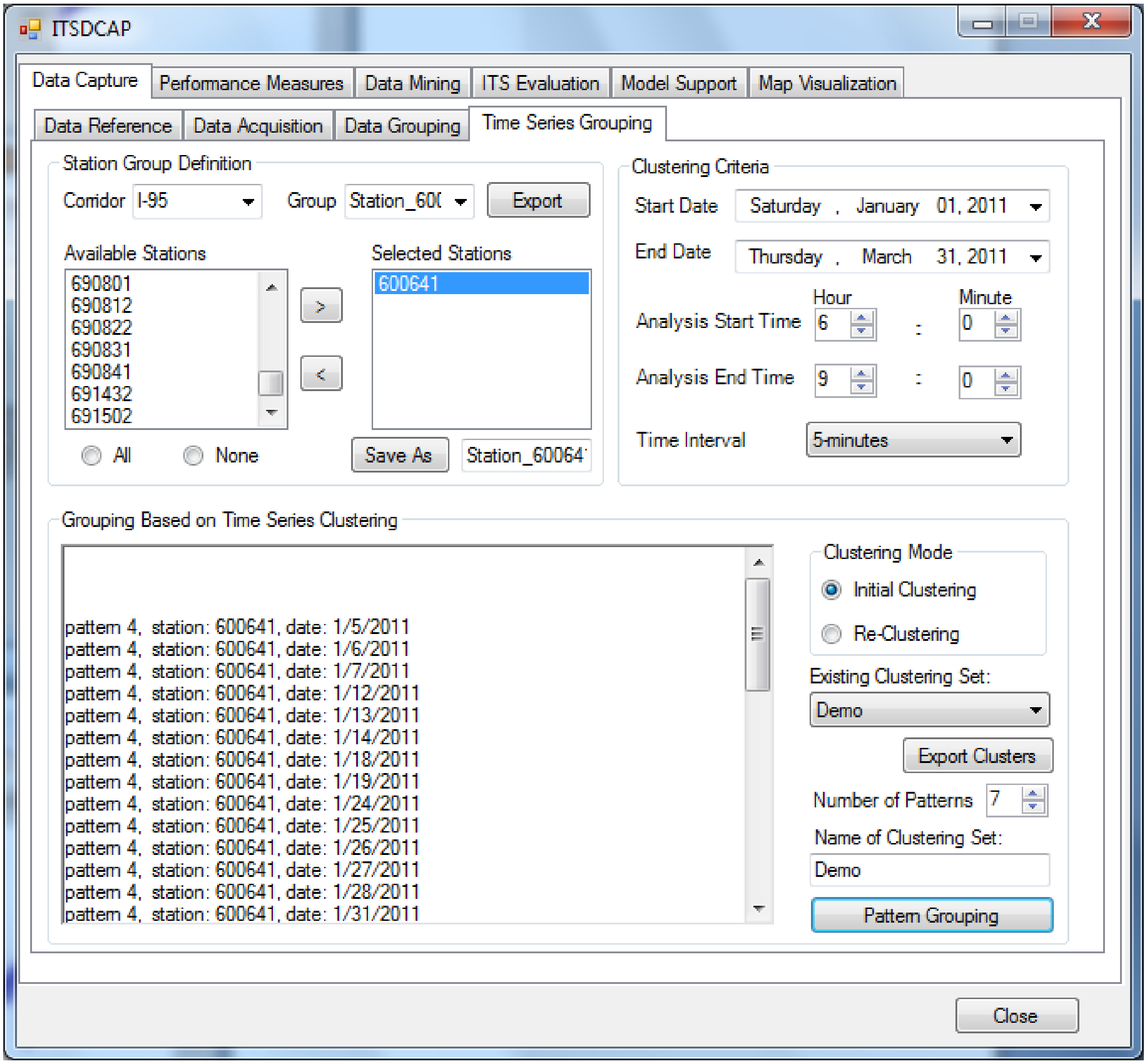


Figure 3-3 ITSDCAP Interface for Time Series Data Grouping

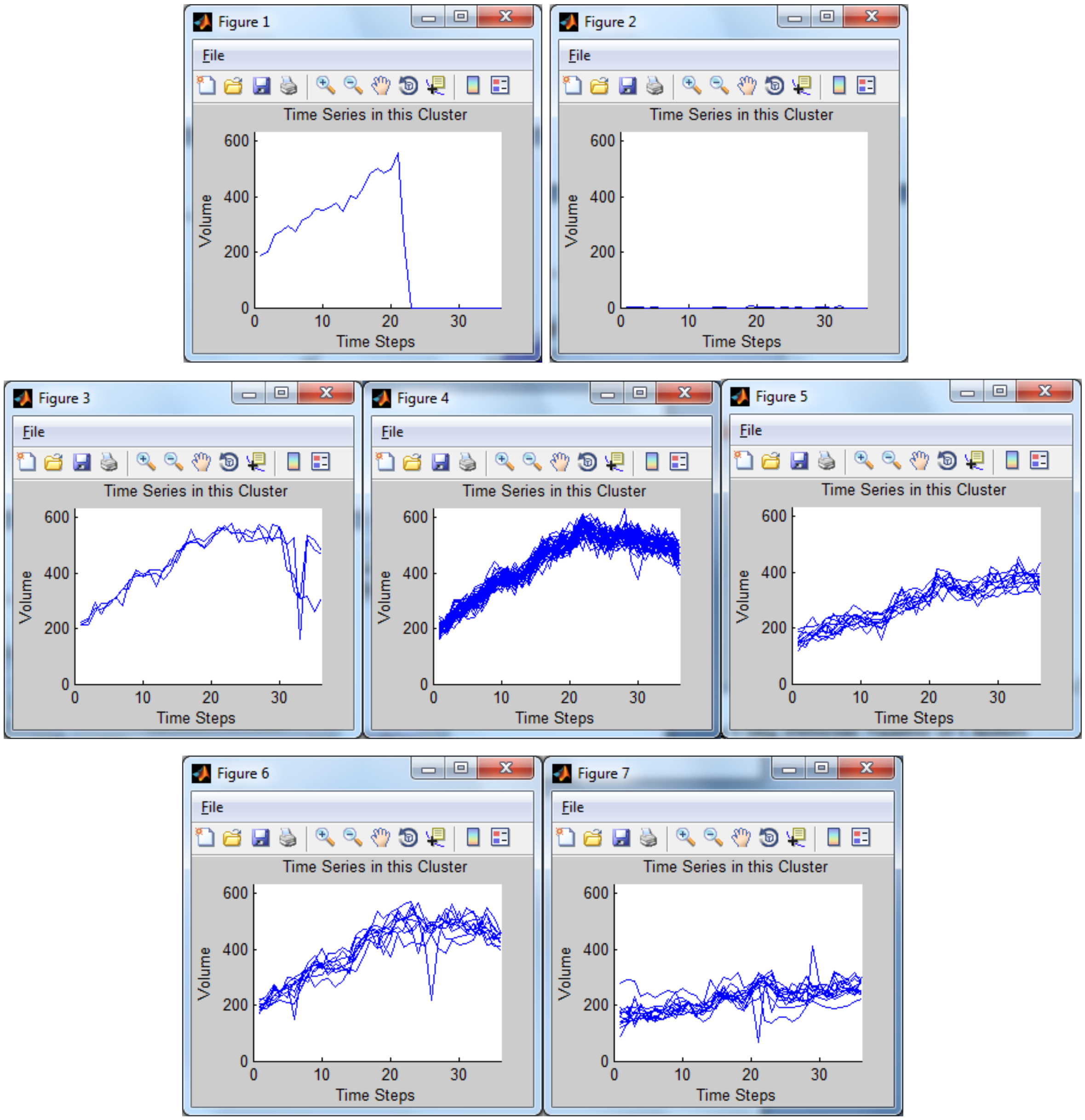


Figure 3-4 Example of Clustering Results

3.6. References

Courage, K.G. and S. Lee. Development of a Central Data Warehouse for Statewide ITS and Transportation Data in Florida: Phase II Proof of Concept. A Report Developed for the Florida Department of Transportation by the University of Florida, Tallahassee, FL, 2008.

Hadi, M., C. Zhan, and P. Alvarez. Traffic Management Simulation Development. Final Report, BDK80 977-03, Jan. 2011.

4. Estimation of Performance Measures

To facilitate the estimation of the performance of transportation systems, a module was developed in the ITSDCAP tool to estimate various mobility, reliability, safety, and environmental measures of the transportation system. This chapter provides a description of the performance.

4.1. Mobility Measures

To fully describe the mobility of roadway system in Florida, the FDOT developed 15 primary mobility performance measures, which consider the quantity of travel, quality of travel, accessibility, and roadway capacity utilization (McLeod and Morgan, 2012). These mobility performance measures have been reported and tracked by the FDOT since 2000. However, the measures have been mainly calculated based on the FDOT Transportation Statistics Office data, which is limited by the measurement locations and durations. The high resolution of ITS data in space and time provide a good alternative data source for estimating mobility performance measures. Therefore, ITSDCAP provides the analyst the option of selecting either the central data warehouse data (currently the STEWARD data), TSS data, or INRIX data as the data source for mobility performance measure calculations.

In the ITSDCAP tool, seven key mobility performance measures can be estimated, as listed below:

- Speed
- Density
- Queue length/location
- Travel Time
- Delay
- Vehicle-Mile Traveled (VMT)
- Vehicle-Hour Traveled (VHT).

INRIX data does not include volume and occupancy measurements. Thus, density, vehicle-mile traveled, and vehicle-hour traveled cannot be estimated based on INRIX data.

4.1.1. Estimation of Queue and Congestion Levels

In the ITSDCAP tool, three methods are provided to determine if the detector station is congested or not, and in turn to estimate the queue length. The first method is based on a predefined speed threshold. If the measured speed at the detector station is less than the speed threshold, the station is considered congested; otherwise, it is considered uncongested. The second method uses occupancy threshold to determine the congestion level. The third method is based on cluster centroids identified from the k-mean clustering analysis. In this method, the fundamental diagram of traffic flow is subdivided into four clusters representing different congestion levels. Cluster I corresponds to nearly free-flow conditions, and the average speed is almost constant at free-flow speed regardless of the demand. Traffic Cluster II is still uncongested but with a reduced speed. Cluster III is a more congested region where the speed drops, but to a lesser degree than the points in Cluster IV. Cluster IV corresponds to extremely congested conditions, with low speed and low constrained flows. Depending on the Euclidean distances from each cluster centroid, the traffic measures at each detection station are associated with one of these clusters (that is, congestion levels). Figure 4-1 presents an example of clustering results for a detector station, DS-1507E, located on SR-826 eastbound in Miami-Dade County, Florida, based on traffic detector data from December 1, 2008, to December 31, 2008. The four cross symbols in these figures denote the locations of the cluster centroids. Once the congestion level is identified for each detector station, the spatial distribution of congestion levels is used to determine the queue length.

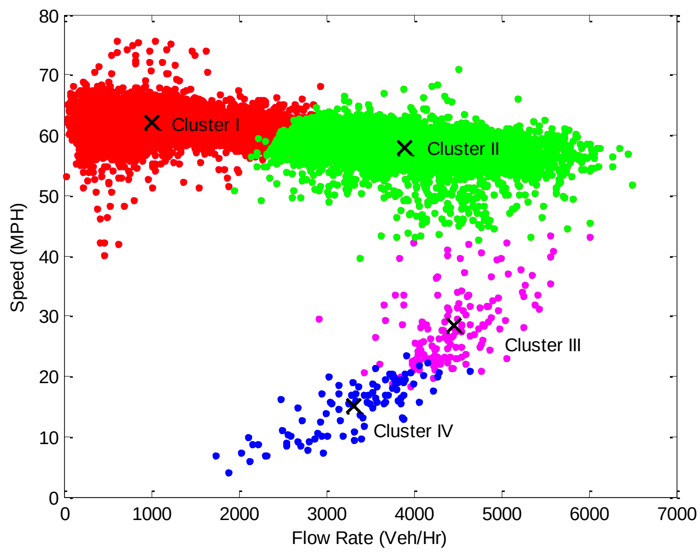


Figure 4-1 Clustering Results for a Detection Station on SR-826

4.1.2. Estimation of Travel Time

Various on-line and off-line travel time estimation methods based on point detector data are provided in the ITSDCAP tool. In addition, travel time can be calculated based on INRIX data. In the future, if data becomes available, data from Automatic Vehicle Identification (AVI) technologies (such as Bluetooth or electronic tag readers) or Automatic Vehicle Location (AVL) technologies (such as GPS devices on different vehicle types) will be used. This allows comparing the accuracy of different estimation methods. Below are descriptions of the methods that are based on point detector data. Figure 4-2 presents a schematic diagram of detector configuration to facilitate the explanation.

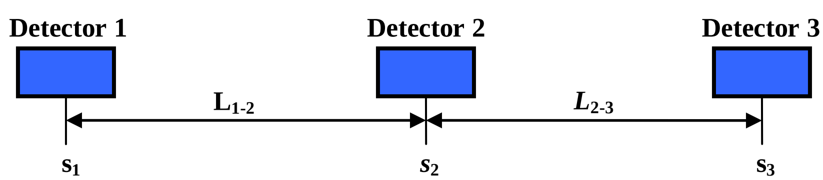


Figure 4-2 Schematic Diagram of Detector Configuration

Mid-Point Method

A travel time link is divided into several segments. The summation of roadway segment travel time yields the total travel time along the link. For each segment, the speed detected at the upstream detector is used to represent the average speed of the first half segment, and the speed at downstream detector is used for the second half segment. Thus, the travel time $TT_{1\text{-}2}$ along the segment $L_{1\text{-}2}$ is

$$TT_{1-2} = \frac{L_{1-2}/2}{s_1} + \frac{L_{1-2}/2}{s_2} \tag{4-1}$$

where s_1 and s_2 are the measured speeds at the upstream and downstream detector locations, respectively.

Minimum Speed Method

Similar to the Mid-Point method, the Minimum Speed method also estimates the travel time along a travel time link by summing up the segment travel time. However, the difference is that the Minimum Speed method uses the lower value of detector speeds at either end of the roadway segment in travel time estimation; that is,

$$TT_{1-2} = \frac{L_{1-2}}{\min(s_1, s_2)} \tag{4-2}$$

Hybrid Method 1

Two hybrid models for travel time estimation were developed in a previous research project (Xiao, 2011). Below is a brief description of Hybrid Model 1, and a description of Hybrid Model 2 will follow. Hybrid Model 1 uses a speed-based method to estimate travel time for uncongested roadway segments and a traffic flow-based method for congested segments, since previous studies have shown that speed-based travel time estimation methods work well under low congestion levels, while traffic flow theory-based methods work well under congested conditions, and the combination of these two methods may achieve better estimation performance than either of these methods alone.

The widely used Mid-Point method is selected as the speed-based method for uncongested segments, which has the expression in Equation 4-3. For a congested segment, the travel time is estimated using a traffic flow theory-based method. This approach estimates travel time by comparing cumulative traffic entering and exiting the segment for a given time interval. The expression is as follows:

$$TT_{1-2} = \frac{L_{1-2}k_{1-2}}{q_2} = \frac{L_{1-2}(k_1 + k_2)}{2q_2} \tag{4-3}$$

where q_2 is the flow rate at downstream detector location. The $k_{1\text{-}2}$ is the segment density, which can be estimated as the average of densities at both immediate upstream and downstream locations. When a segment is partially queued, (for example, when the head of the queue or end of queue is located within the segment), the travel time for this segment is divided into two parts, as shown in Equation 4-4.

$$TT_{1-2} = \frac{L_{i,1}k_{i,1}}{q_{i,1}} + \frac{L_{i,2}}{S_{i,2}} \tag{4-4}$$

In Equation 4-4, the first part is for the queued portion and the second part is for the unqueued portion. The variable $L_{i,1}$ is the length of queue section within the segment and $L_{i,2}$ is the uncongested section length.

Hybrid Model 2

Since the speed-based methods are more straightforward, Hybrid Model 2 combines the Mid-Point method with the Minimum Speed method, instead of combining the speed-based method with traffic flow theory-based method, as in Hybrid Model 1. The rationale for selecting the Minimum Speed method is that this method has been found by the research team to work better for congested conditions when the queue is building.

Similar to Hybrid Model 1, if the segment is free of queue, the Mid-Point method is used to estimate the segment travel time. When the segment includes a queue and currently the queue is propagating backwards, the Minimum Speed method as shown in Equation 4-2 is applied, to implicitly capture the growth of queue. If the queue is dissipating, the Mid-Point method is applied to account for the fast moving backward recovery shock wave.

Travel Time Based on INRIX Data

Another option available to ITSDCAP users is to employ INRIX data to estimate travel time. As described in Section 3, INRIX collects and fuses data from multiple sources (mainly GPS data) to provide travel time estimates that are then distributed to public and private sector users.

Off-Line versus On-Line Travel Time Estimation Methods

Travel time estimation can be divided into two categories: on-line estimation and off-line estimation. The difference between these two categories is that for on-line travel time estimations, short-term future traffic conditions along the paths of the vehicle are not available and only the instantaneous travel time (based on the traffic conditions at the time of the estimation) can be used in the estimation. For off-line estimation, the traffic conditions at later time periods can be determined based on historical data, allowing more accurate estimation of experienced travel time by travelers accounting for traffic conditions, as the vehicle progresses in its route from one link to the next. In ITSDCAP, both off-line and on-line types of estimation can be requested and compared.

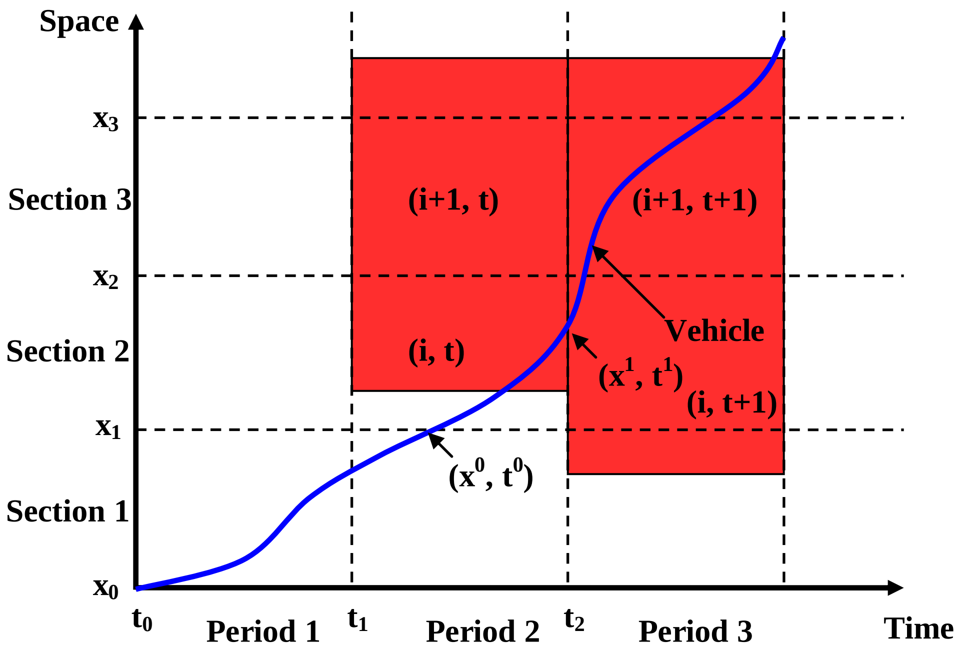


Figure 4-3 Schematic Diagram for Off-Line Travel Time Estimation

The off-line estimation methods used in the ITSDCAP tool divide the whole time duration into small time periods that are critical to the temporal aggregation level of the detector data, as shown in Figure 4-3. As shown in this figure, a vehicle enters cell i at location x^0 and time t^0. The remaining time in this time period is compared to the time that is required to reach the downstream station, and the minimum value of these two values is used in the travel time estimation for this cell. Depending on the location of the exit point (x^1, t^1), the vehicle can either enter the next link during the same period, which is cell (i+1, t), stay on the same link but experience different traffic conditions at time t+1, which is cell (i, t+1), or enter the downstream link at the next period of time, which is cell (i+1, t+1). The resulting route travel time of a vehicle is the time that the vehicle arrives at the destination (last detection station on the path for which the travel time is estimated) minus the time that the vehicle departs from the origin (first detection station on the path for which the travel time is estimated). Different off-line estimation methods use different approaches to estimate the travel time within each cell, which are the same as their on-line equivalent methods.

4.1.3. Estimation of VMT and VHT

The calculation of vehicle-mile traveled and vehicle-hour traveled in ITSDCAP is straightforward and achieved by multiplying the volume count with the corresponding distance and travel time, respectively.

4.1.4. Estimation of Density

Depending on whether the measured speeds by traffic detectors are time-mean speed or space mean speed, the density is calculated in two different ways. When space-mean speed is reported, density can be estimated based on the fundamental relationship of traffic flow, as follows:

$$k_{i,j} = v_{i,j} / s_{i,j} \tag{4-5}$$

where the variable $k_{i,j}$ denotes the density at the detector station j at the time interval i, and $v_{i,j}$, $s_{i,j}$ are the corresponding traffic volume and speed, respectively. However, if the time-mean speed is measured, the density is estimated from the occupancy, as shown below:

$$k_{i,j} = 52.8 O_{i,j} / L_{eff} \tag{4-6}$$

where the variable $O_{i,j}$ represents the measured occupancy at the detector station j and time i. The variable L_{eff} is the effective vehicle length, which is the summation of the detector length and the average vehicle length. In ITSDCAP, the default value for the effective vehicle length is set at 22.9 ft, based on the assumption of 6.5 ft for detector length and 16.4 ft for average vehicle length. This default value can be changed based on local information.

4.1.5. Mobility Estimation User Interface

Figures 4-4, 4-5, and 4-6 present the user interface in ITSDCAP for the mobility performance measure calculations. It is seen in Figure 4-4 that the mobility measures can be estimated for the periods of time and locations, specified by the users. The time period can be for a continuous period of time or for days that meet specific criteria such as days with incidents or specific patterns based on clustering results. The selection of the segment for travel time estimation can be accomplished through GIS maps, as shown in Figure 4-5. The estimation results can be visualized in either a table or graph formats. For the graphical display, a chart plot is provided for most of the measures, except that a contour plot is also provided for speed and density visualization to show the variations in space and time.

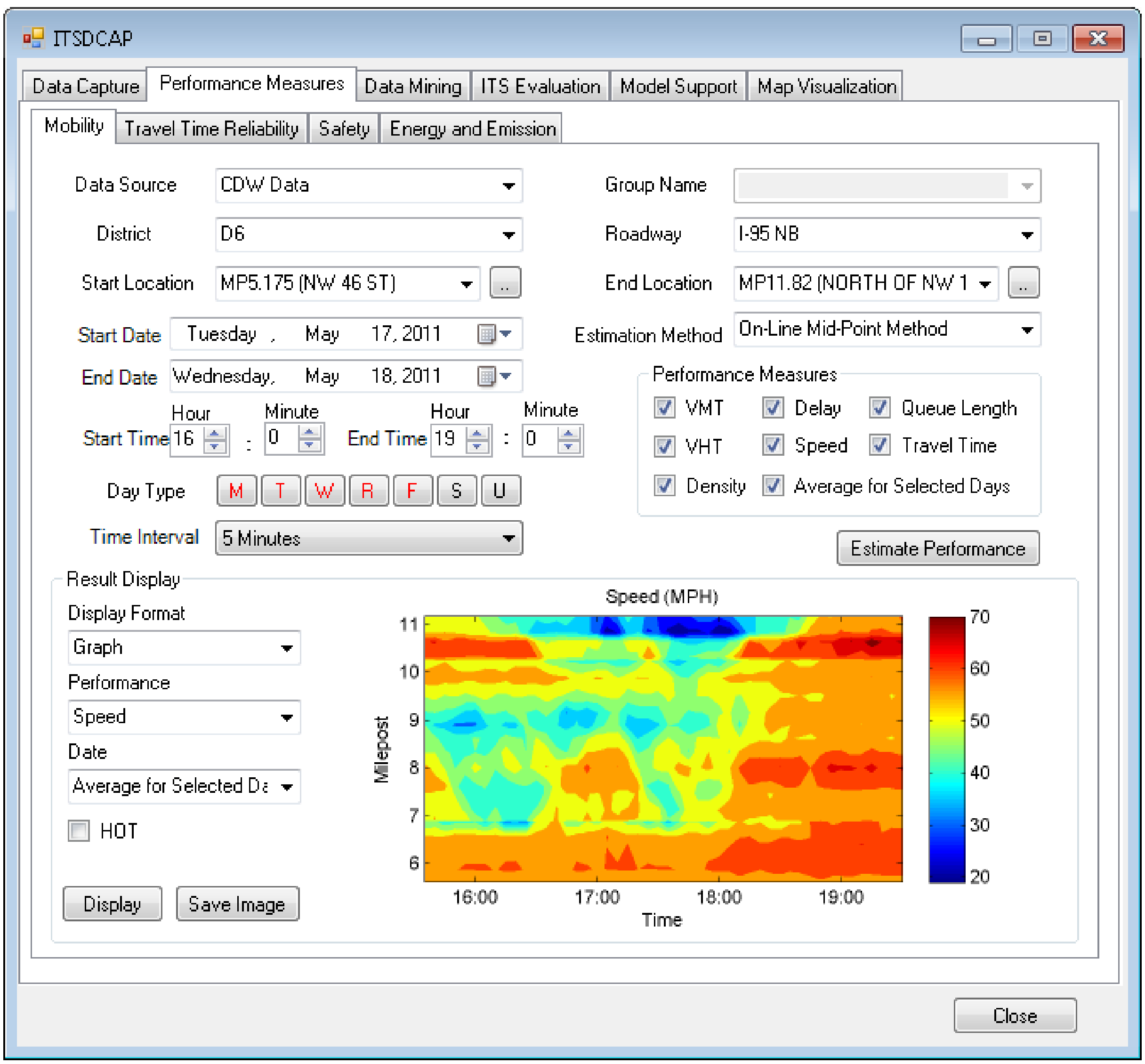


Figure 4-4 Interface for Mobility Performance Estimation

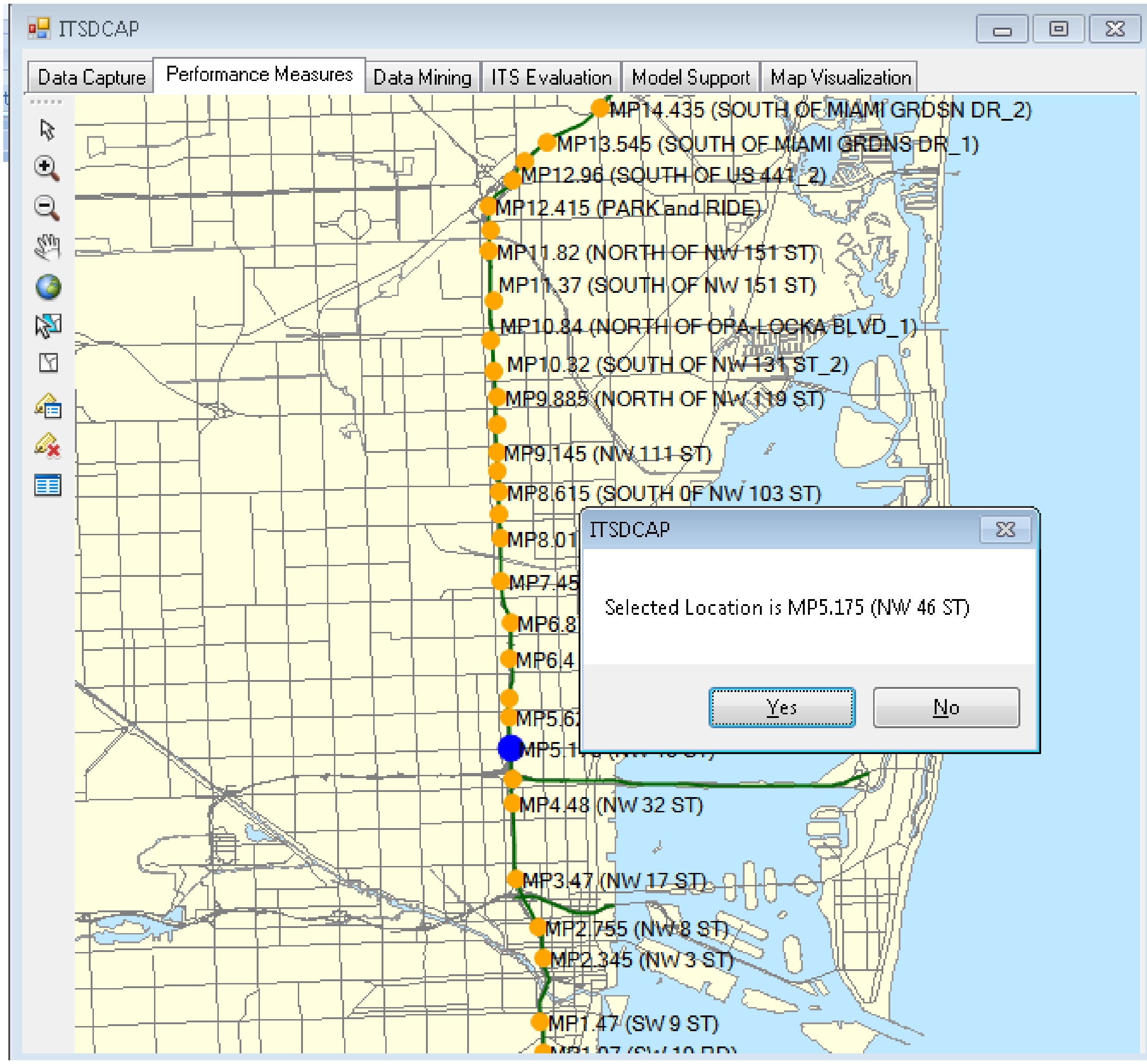


Figure 4-5 Starting and Ending Location Selection for Mobility Performance Estimation

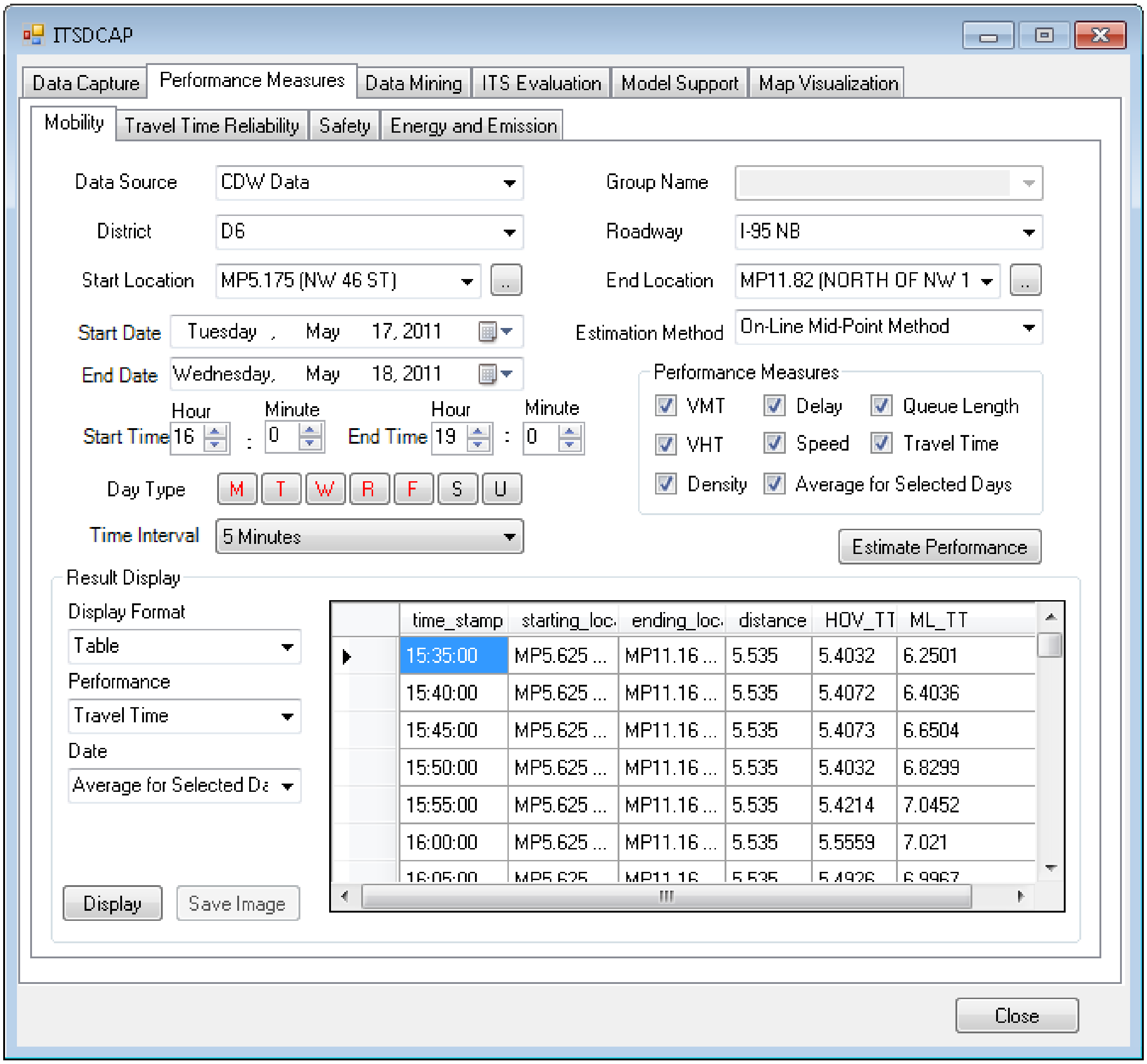


Figure 4-6 Mobility Performance Estimation Visualization

4.1.6. Mobility Measurement Test Cases

A Professional Engineer trainee from the FDOT District 4 tested the ability of the ITSDCAP tool to analyze the mobility and safety of a segment of I-95 SB in Broward County and identified the locations where mobility bottlenecks and safety problems are likely to occur. The study time period was 6:00 AM-9:00 AM on weekdays between March 1, 2010 and November 1, 2010. The detector data is obtained from the STEWARD data warehouse, and the crash data is retrieved from the Crash Analysis Reporting (CAR) System. Since the case study involved both mobility and safety analysis, the results will be presented after discussing the estimation of safety measures in Section 4.3.

4.2. Travel Time Reliability Measures

4.2.1. Literature Review

Increasingly, travel time reliability is considered as an important component of the performance of transportation systems and of travelers' perceptions of this performance. The increased recognition of the importance of travel time reliability is reflected by changes to traditional monitoring programs. For example, in a report published in July 2005, the National Transportation Operations Coalition (NTOC) initiative selected the buffer index (BI), a travel time reliability measure, as one of a few good measures for transportation operations agencies to use for internal management, external communications, and comparative assessments (NOTC, 2005). Public agencies from around the United States are now using travel time reliability metrics as key performance measures to monitor their system operations (FDOT, 2008; GDOT, 2011; and WSDOT, 2011). Recognizing the critical need for researching travel time reliability, the Second Strategic Highway Research Program (SHRP2) has specified travel time reliability as one of the four main research areas of the program and has funded extensive research activities in this area.

Travel time reliability measures the level of consistency of travel conditions over time. A unifying reliability definition can be found in the final report of the SHRP2 LO3 project (Cambridge Systematics et al., 2010) where reliability is defined as "the level of consistency in travel conditions over time, and is measured by describing the distribution of travel times that occur over a substantial period of time. . . ." The different approaches to defining reliability have led to recommending several metrics for use. These metrics may not necessarily produce consistent assessments of reliability among themselves, since they define travel time consistency in different manners.

Tu et al. (2008) classified reliability metrics as statistical range methods, buffer time methods, tardy trip measures, probabilistic measures, and skew-width methods. Lomax et al. (2003) discussed the development of reliability measures and the factors to consider before selecting a measure. They concluded that the metrics that are most promising are the percent of variation, misery index, and the BI. This conclusion was based on five factors, including the compatibility with multimodal analyses, ability to measure urban and rural travel conditions, consideration of the effects of trip length and time, ability to serve several audiences, and applicability to different area sizes. In a later work, Van Lint and Zulen (2005) noted that the BI and the misery index may not be appropriate because of the underlying skewed travel time distribution. They concluded that most currently utilized reliability metrics should be used and interpreted with some reservations.

The common assumption about the normality in travel time distribution was investigated by Rakha et al. (2007). The study concluded that the normality assumption is not supported by the observed data. Instead they proposed that a log normal distribution can better describe the travel time during uncongested conditions. Similarly during the congested hours, a mixed or a bimodal distribution fits better the observed travel time distribution. Additionally, they pointed out that the rate at which the mean changes during the congested hours can be faster than the rate of change of the standard deviation, thus the coefficient of variation may actually decrease when trips are becoming more unreliable.

Pu et al. (2010) examined a number of reliability metrics assuming a log normal distribution of travel time with a constant median, while varying the variability and skewness of the travel time distribution. He concluded that the coefficient of variation is a good proxy for a range of reliability metrics and suggested that the use of the BI is not always appropriate unless BI is computed based on the median rather than the mean. The latter is because in heavily skewed travel time distributions, the use of the mean may underestimate the travel time reliability. The same author in a different study (2010) pointed out that some reliability metrics may be inconsistent in their depictions of reliability, such as is the case of the BI that remains constant for different values of the coefficient of variation. As discussed later in this paper, fixing the median while varying the standard deviation and skewness as is done in the Pu study may not reflect real-world conditions in which the parameters of travel time distributions vary by time of day and may be correlated with each other.

The SHRP2 LO3 project (Cambridge Systematics et al., 2010) examined a set of six reliability metrics to determine their sensitivities to different types of freeway improvements. The utilized metrics were the BI, on-time performance, 95th planning time index, 80th percentile planning time index, skew statistics and misery index. Based on empirical tests, it was found that all metrics were sensitive to the effects of improvements. However, it was noticed that the 95th percentile travel time or travel time index (TTI) may be too extreme a value to be influenced significantly by operation strategies and that the 80th percentile was more sensitive to these improvements. Another aspect, related to the amount of data required to assess system reliability was tested concluding that an absolute minimum of six months of data is required to establish reliability within a small error rate, in areas where winter weather is not a major factor. However, a full year of data is preferred.

Even though reliability metrics have been explored and compared as discussed above, most of the work in the literature has focused on using these measures for relatively coarse levels of aggregation of travel time data; for example, for the whole peak periods. Such uses imply that the parameters of the travel time distributions are assumed to remain the same for the whole period of analysis. This may not be sufficient for advanced management strategies such as for setting managed lane pricing and for capturing travelers' behaviors for use in dynamic traffic assignment (DTA)/simulation modeling. These models are normally used to simulate travelers' route selection behaviors at 15 to 30 minute intervals and are being extended to include reliability in their generalized cost functions. In addition, the analyses of the Highway Capacity Manual (TRB 2010) are also conducted at the 15 minute analysis level, and although the procedures of the current version of the manual do not consider reliability metrics, discussion has already started for the potential inclusion in future versions. This above discussion supports the argument that the reliability metrics need to be assessed at fine grained levels of aggregation for an increasing number of applications.

4.2.2. Travel-Variant Travel Time Distributions

Travel time reliability metrics are calculated to reflect the variability of travel time from day to day. Thus, they are supposed to represent the attributes of travel time distributions that are most relevant to the perception of system reliability. Since this section focuses on the examination of the attributes of time-variant reliability metrics, it is useful to examine first the time variant parameters of travel time distributions. In addition, as stated earlier, previous reliability studies have calculated reliability measures for the whole peak periods, implicitly assuming that the travel time distribution parameters remain constant during these periods of the analysis. Investigating the variations in travel time distribution parameters by time interval allows the examination of the validity of this implicit assumption.

Comparing the parameters of travel time distributions under different conditions can also provide important information regarding the sources of the unreliability of these facilities. These sources can be classified as recurrent and non-recurrent events. Recurrent events include demand fluctuations, variations in traveler behaviors, and stochastic variations in flow breakdown and capacity. Non-recurrent events include incidents, work zones, special events, and weather events. It is logical to expect travel time variability mainly due to recurrent events to have distributions that are much flatter and more symmetrical compared to conditions with variability significantly influenced by non-recurrent events, which are expected to have skewed distributions. However, the shapes of the distributions could also be influenced by facility types and their operational characteristics. Previous studies have produced limited information regarding the characteristics of travel time distributions during incident and no-incident conditions and their influences on travel time variability, as explained next.

Li et al. (2006) found that, under the free-flow conditions of a transportation system, travel time distribution has the shortest right tail whereas the afternoon peak has the most skewed distribution. They also found that the variation in demand in the morning peak explains most of the travel time variability while incidents play a major role explaining the variability during the PM peak. The authors warned that these results are likely to be site specific. Boyles et al. (2010) examined four operational conditions on two corridors: no incident-good weather (NIGW), poor weather (PW) and incident present (IP). They found that the normal distribution is best suited for describing speed in NIGW conditions, while the beta distribution is best suited for fitting PW and IP conditions. The authors also warned analysts about the site specific character of the results. Tu et al. (2008) pointed out that the effect of traffic incidents on travel time distribution is a function of the volume using the facility. Below a certain threshold, 1400 vphl, incidents slightly increase the 10*th* percentile of travel time by 3% and the median travel time by an average of 6%. Above this threshold, incidents result in significantly higher median travel time (38% increment) while the 90^{th} percentile travel time is increased in average by 75%.

4.2.3. Travel-Variant Reliability Metrics

As described previously, it is important to examine how the reliability metrics vary by time-of-day and particularly how they react to the increase in demand. These metrics have to be sensitive to the increase in congestion and they need to show a consistent trend as the congestion increases in order to allow their use in highway capacity/traffic analysis applications, or when included as parameters in the objective functions of dynamic traffic assignment tools, or of the optimization of the pricing of managed lanes and other strategies.

There are a number of metrics that have been used to quantify reliability. The variation in the parameters of the travel time distributions with the increase in congestion by time of day as described in the previous section is expected to have significant impacts on the values of these metrics. In this project, an investigation was made of the sensitivity of the reliability metrics to the changes in the congestion levels and the ability of these metrics to reflect the reliability differences between facilities in a time-variant context. Table 4-1 shows the definitions of travel time reliability measures used in this study. These measures are mainly selected based on the recommendations given in SHRP2 L03 project (Cambridge Systematics et al, 2010) and constitute a common set of metrics used for reliability assessments.

Table 4-1 Travel Time Reliability Operational Definitions

Reliability Performance Metric	**Definition**
Buffer Index (BI)	The difference between the 95th percentile travel time and the average travel time, normalized by the average travel time.
Failure/On-Time Performance	Percent of trips with travel times less than: • 1.1* median travel time • 1.25* median travel time
95th Planning Time Index	95th percentile of the travel time index distribution
80th Percentile Travel Time Index	80th percentile of the travel time index distribution

Table 4-1 Travel Time Reliability Operational Definitions (Continued)

Reliability Performance Metric	**Definition**
Skew Statistics	The ratio of 90th percentile travel time minus the median travel time divided by the median travel time minus the 10th travel time percentile
Misery Index	The average of the highest five percent of travel times divided by the free-flow travel time.

4.2.4. User Interface for Travel Time Reliability Measures

Figure 4-7 presents a snapshot of user interface for travel time reliability measures in the ITSDCAP tool. The following travel time reliability metrics are calculated, including:

- Standard deviation/variance
- Buffer index based on mean or median free-flow travel time
- Failure/on-time performance based on the threshold of 1.1 or 1.25 times of median travel time
- Planning time index based on 95th or 80th percentile
- Skew statistics
- Misery index.

After a user specifies the study location and time period, the user can click the "Estimate Performance" button to run the analysis. Once the analysis is finished, the corresponding display window will indicate which index is estimated. The user can then click the "Save Now" button to save the estimation results to the user-specified filename.

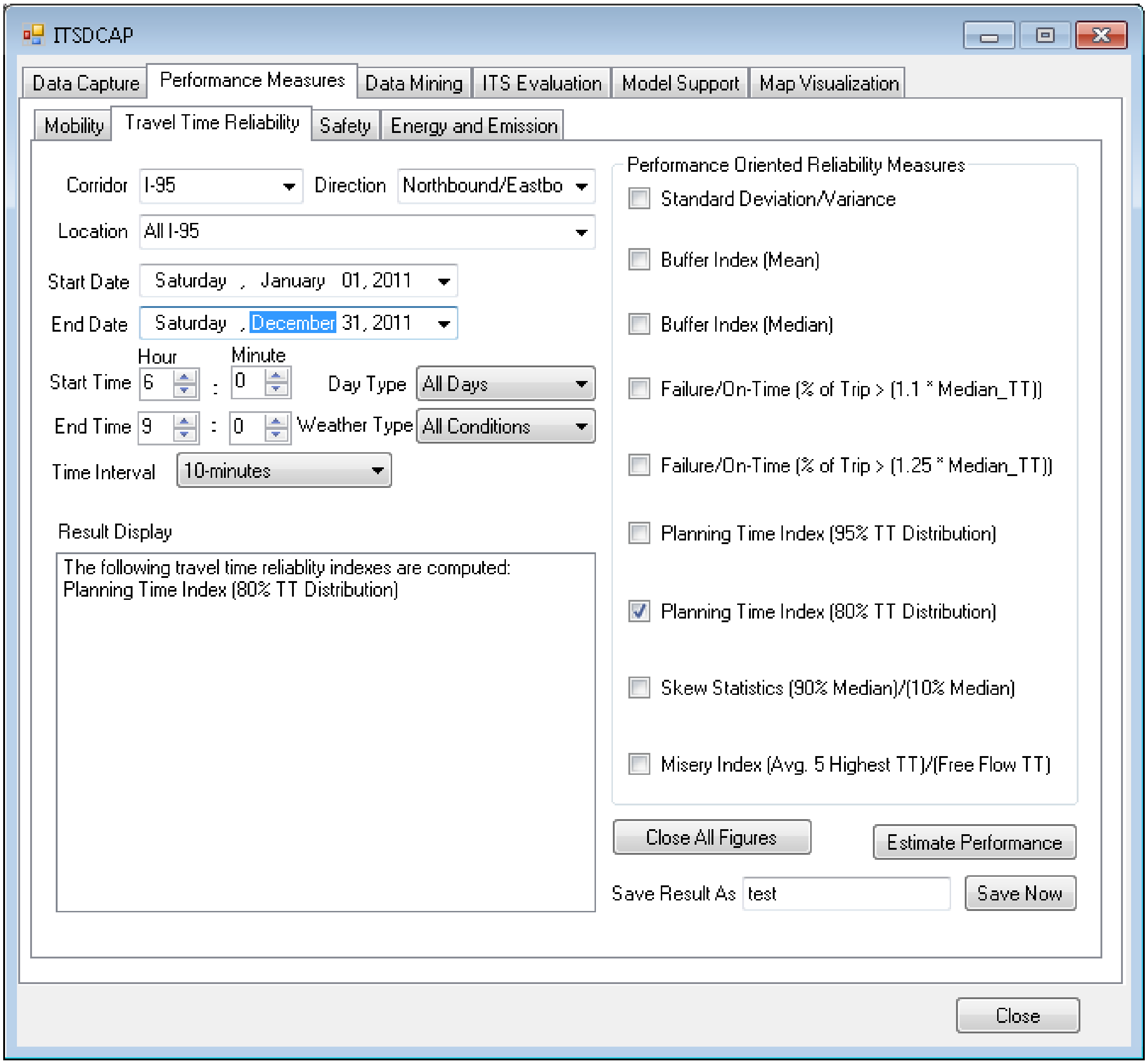


Figure 4-7 User Interface for Travel Time Reliability Measure Calculation

4.2.5. Reliability Measurement Test Cases

This project investigated the variation of the parameters of travel time distributions by time of day, the sensitivity of various reliability metrics to these variations, the effect of time of day analysis interval on the calculated metrics, and the amount of data required to estimate stable values for the reliability metrics. The investigation was made for the I-95 corridor in Miami-Dade County that has general purpose (GP) lanes and the I-95 express lanes, which are high-occupancy toll (HOT) lanes. The investigation allowed the comparison of how travel time reliability attributes vary between the two facilities by time-of-day. The project also explored the trends of the variations of various metrics as the congestion increases during the peak period, which is important when selecting reliability metrics for various applications. These include the use of the metrics as part of the generalized cost functions of assignment models and in optimization of strategies such as congestion pricing. Below are the detailed steps.

The data used in the analysis of this study was obtained from a 6.5 mile segment of the northbound direction of the I-95 limited access facility in Miami, Florida. This segment has two HOT lanes and four lanes operating free of charge as GP lanes. The two HOT lanes have been in operation since December 2008, utilizing a dynamic congestion pricing scheme. Registered vehicles with high occupancy can use the HOT lanes without paying tolls. The HOT lanes have a single entry point and a single exit point and are fully segregated from the GP lanes by plastic poles. This section of I-95 is equipped with point traffic detectors located every 0.3-0.5 miles that collect volume, speed, and occupancy measurements every 20 seconds for both the HOT and GP lanes.

This study utilized travel time estimates for the HOT and GP lanes that are archived for every minute of the day by the SunGuide software. The travel time is estimated by the system based on the speed measurements collected by the microwave detectors installed on the corridor utilizing the mid-point speed estimation method. One year’s worth of data was used in the analysis. However, the project also investigated using data collected for shorter periods of time, as described later in this section. Only weekday travel time data was used in the analysis.

Estimated Reliability Metrics

In this study, the reliability metrics in Table 3-1 were estimated for the GP and HOT lanes, utilizing different time interval lengths in the analysis. The results presented in Figures 4-8 and 4-9 are based on 15 minute intervals. These two figures show different trends of different reliability metrics with time of day for the GP and HOT lanes respectively, as explained below.

The metrics that exhibit continuity and sensitivity in their variations in response to the increase in variability as the congestion in the peak hour is approached are the 95th percentile planning time index (PTI), 80th percentile PTI, and the misery index. Figures 4-8 and 4-9 show that the

95th Percentile PTI metric and even more the 80th percentile PTI metric clearly indicate that reliability decreases as the peak demand period (at 5:00 PM) is approached. The reliability is lowest at this peak reflecting the highest variability observed when examining the parameters of the travel time distributions. These two figures also show that the 95% PTI of the GP lanes is higher than that of the HOT lanes except at the peak hour, at which the PTI is equal for both facilities, reflecting the comparable travel time variability of the GP and HOT lanes at the peak hour. It is interesting to note that the 80th percentile PTI is lower for the HOT lanes for all of the investigated hours, including the peak hour at 5:00 PM, and that the difference in reliability between the GP and HOT lanes is higher when the comparison is based on the 80th percentile PTI compared to when measured based on the 95th percentile PTI.

A similar trend to those observed with the PTI measurements discussed above was observed with the misery index, as indicated in Figures 4-8 and 4-9. For the GP and HOT lanes, the misery index values are lower in the uncongested periods compared to the congested periods, as expected. The misery index of the HOT lane is lower than that of the GP lanes prior to the peak hour but exceeds it at the peak hour.

The results of examining the other measures do not show the consistent trend of increase with the increase in congestion, observed when examining the PTI and the misery index in the discussion above. Figure 4-9 shows that the BI metric for the GP lane is not sensitive to the increase in congestion. BI is computed as the difference between the 95th percentile and the mean divided by the mean travel time. The BI value of the GP lanes remains almost constant during the peak period because, although the variability of travel time increases significantly as the peak demand approaches the values of the numerator in the BI calculations, the travel time mean (or median) also increases as the congestion increases. For the GP lanes, the rate of change of the travel time mean/median parameter is close to the rate of change of the difference between the 95th percentile and the mean/median travel time resulting in BI values that remain almost constant. In the case of the HOT lanes, the rate of increase in the variability of travel time with the increase in congestion is higher than the rate of the increase in the mean/median resulting in an increase in the BI. The inconsistency and lack of sensitivity of the BI to the increase in congestion in some cases may limit its use, at least for some reliability assessment and analysis tasks.

The failure/on-time metric (FOT) represents the number of occurrences of travel time that are lower than the median travel time multiplied by a factor (1.1 or 1.25). The implication is that the higher this metric, the more reliable the system, because more of the travel time instances are close to the median. The value of this metric is affected by both the median and standard deviation values. As discussed previously, as the congestion increases both the median and standard deviation increases. The results shown in Figures 4-8 and 4-9 indicate that there is no clear trend of the time-variant trend of FOT metric with the increased congestion. This is because the change in the metric value from one step to the next depends on how the rate of change in the median value is compared with the rate of the change in the standard deviation for the facility under consideration.

The trend of the skew statistics with the increase in congestion is also not very clear. It appears to be a function of the relative contributions of different sources of unreliability to the total travel time variability. For the GP lanes, it is clear that the value of this metric is lower at higher congestion levels, indicating more symmetrical distributions due to the higher contribution of the recurrent congestion to the variability. The variations due to recurrent congestion have been characterized by a more symmetrical distribution (i.e., less skewed distributions), as described earlier. For HOT lanes, the Skewness Statistics initially increases as the shoulder of the peak hour is reached. As explained earlier in this section, the travel time on the HOT lanes has a high sensitivity to the variations in demand due to the lower capacity of the managed lanes. This causes the variation in demand at the shoulder of the peak and thus results in a high increase in the variability and Skewness statistics. However, at the peak hour itself, the managed lanes operate consistently at capacity causing a lower variability between days and significantly lower Skewness statistics.

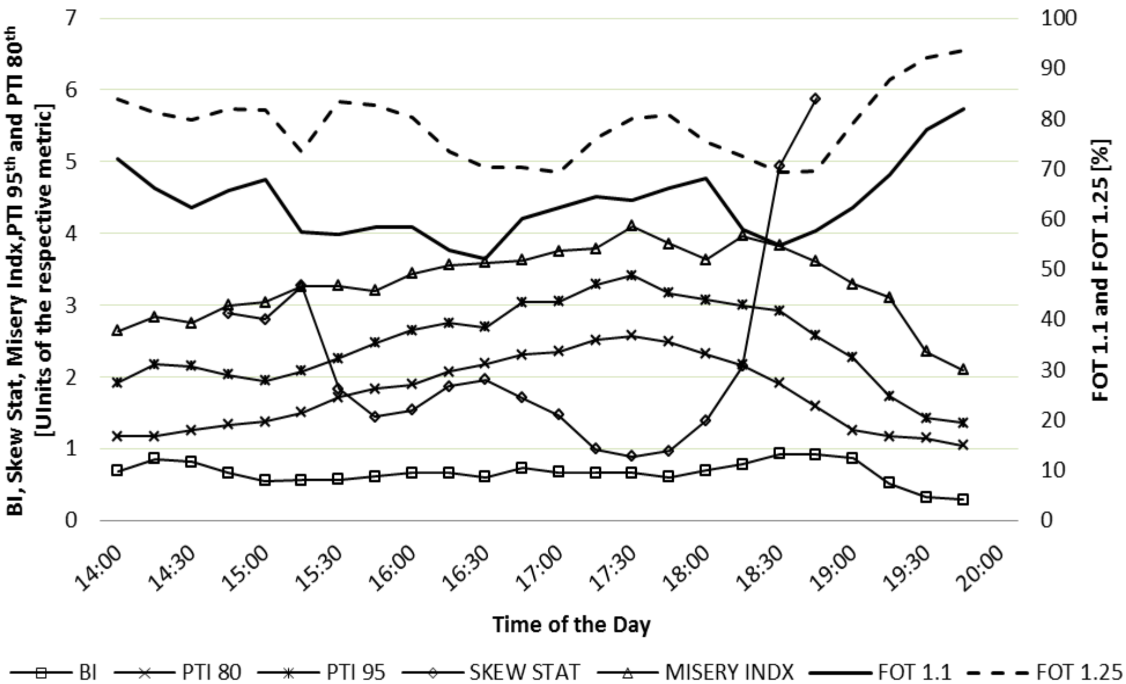


Figure 4-8 Reliability Metrics Variation with Time-of-Day GP Lanes

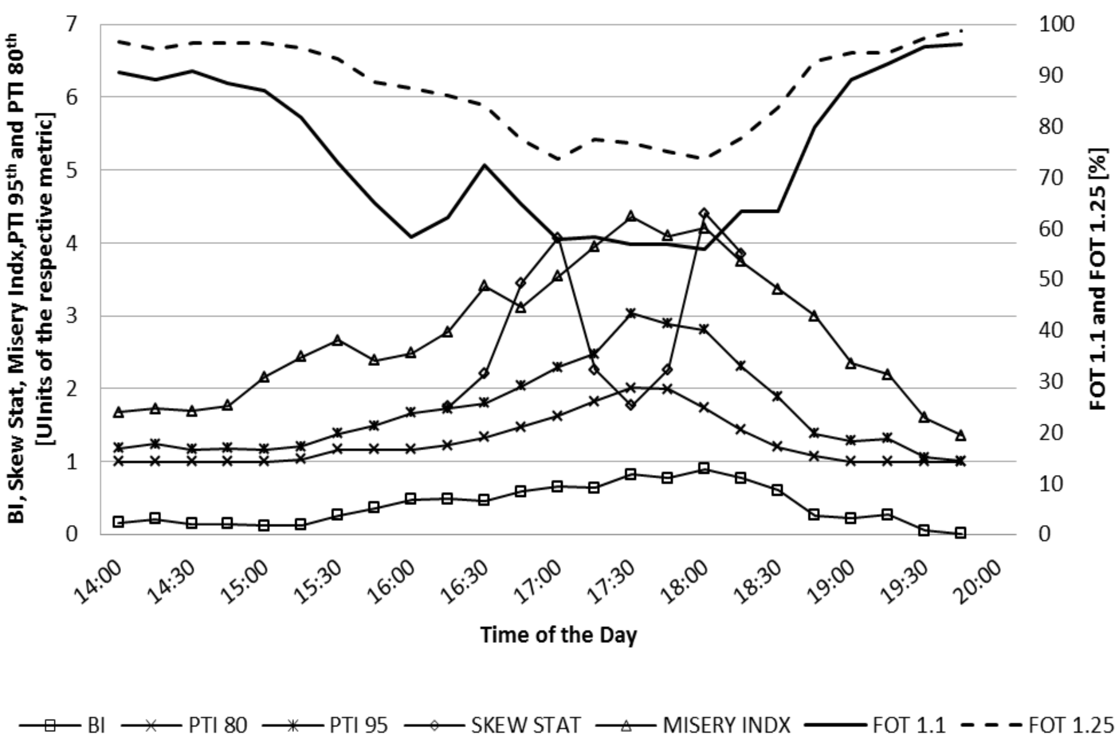


Figure 4-9 Reliability Metrics Variation with Time-of-Day HOT Lanes

Analysis Period Length

The discussion presented in the previous sections indicates that there is a considerable variation in the travel time distribution parameters and travel time reliability metrics when these parameters and metrics are estimated for different time intervals of the peak periods. This section addresses the effect of the choice of the time of day analysis interval (segmentation of the peak period into subintervals) on the calculated metrics.

The available data allow estimating travel time for every minute of the day. However, it may not be useful to produce the reliability metrics for every minute. At the same time, as can be concluded from the earlier discussion, it may not be appropriate to compute the metrics to represent long periods of time. As stated earlier, existing analysis and modeling approaches utilize 15- to 30-minute intervals suggesting that segmenting the peak period into 15- to 30-minute intervals may be needed for estimating the reliability metrics for such applications.

In this project, the 80th Percentile PTI metric was computed for analysis periods of 15 minutes, 30 minutes, 1 hour, 2 hour, and 4 hour, and the results are presented in Figure 4-10. The results show that depending on the congestion levels, different aggregation periods may lead to significantly different assessments of the reliability. One effect of using more aggregated periods of time would be the dilution of the travel time variability during the period of interest, particularly during periods of varying congestion levels. Another potential undesirable effect of such aggregation is the dilution of the relative differences between the facilities being compared, such as the comparison of the GP and HOT lanes conducted in this study.

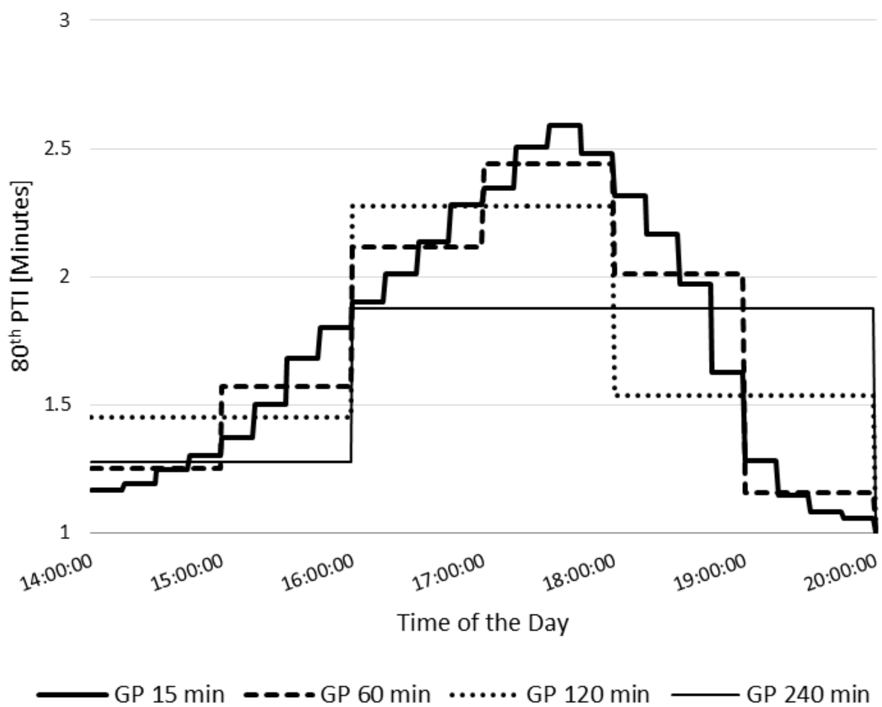


Figure 4-10 Impact of Aggregation Levels on the 80th Percentile PTI Metric

Length of Data Collection Period

Another aspect of the travel time reliability researched is the length of the period for which the data is collected as it impacts the accuracy of the computed reliability metrics. This accuracy can be assessed compared to the accuracy achieved if the data is collected for a long period of time (a whole year in this study).

The analysis investigated the root mean square error (RMSE) between the values of the calculated reliability metrics based on data collected for each data collection period length and the metrics values calculated using a year's worth of data. The four measures are the 95th percentile and 80th percentile PTI for GP and HOT lanes. Figure 4-11 shows how the RMSE of the computed values of reliability measures varies with the number of weeks considered for reliability estimation. The figure shows that as the data collection period increases, the error relative to the estimates based on one year's data decreases. However, Figure 4-11 shows that even when collecting data for 40 weeks, the RMSE is still at least 10% and as high as 18% for the investigated measures. The SHRP2 LO3 project recommended using at least 6 months and preferably a year's worth of data to estimate the reliability metrics. The results in Figure 4-11 confirm that at least one year's worth of data is needed to estimate stable values of the investigated reliability measures. Based on the results of the study, at least one year of data is recommended for obtaining stable values of reliability metrics.

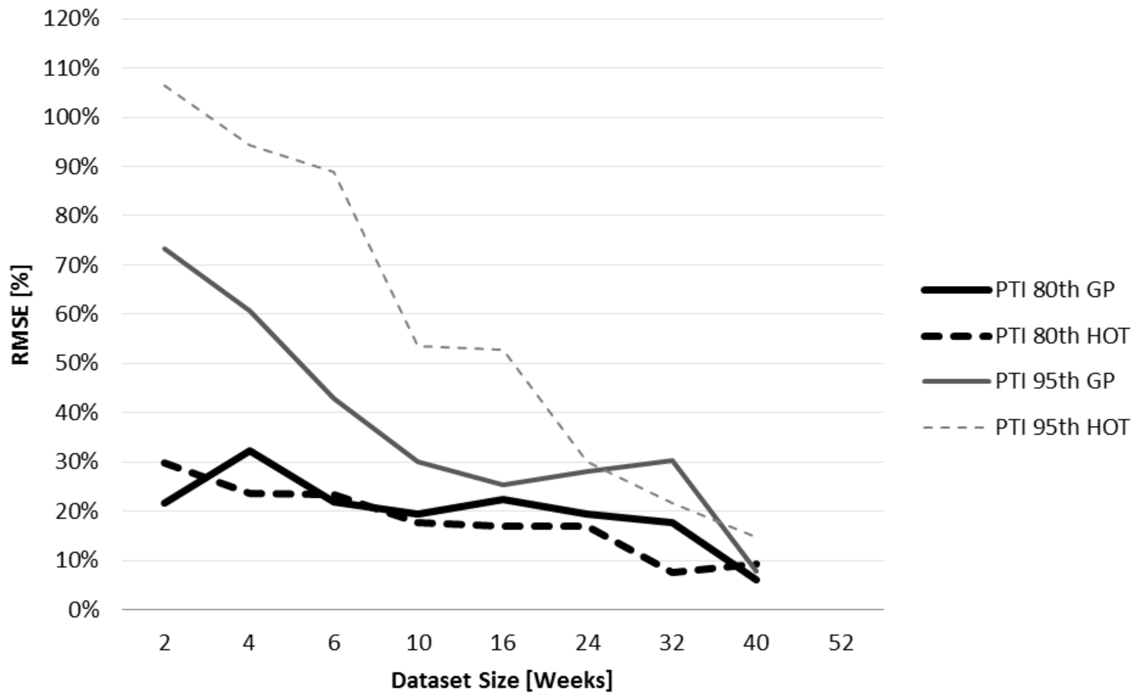


Figure 4-11 Effect of Data Collection Period Length

Summary

There are several metrics of reliability that have been recommended for use. These metrics need to be calculated at relatively fine levels of aggregation of travel time data when used for advanced management strategies and analysis methods. The results show that the parameters of travel time distributions vary during the peak period reflecting the effects of the traffic congestion, traffic flow dynamics, and the proportion of the contribution of non-recurrent factors, such as incidents to the unreliability of travel time on the investigated facility. Different trends of travel time variations are observed when using different reliability metrics to assess reliability as the congestion level changes during the peak period.

The results also show that examining combinations of time-variant static distribution parameters and reliability metrics of the GP and HOT lanes can provide valuable information that cannot be obtained when performing the analysis at higher aggregation levels. The 95th Percentile PTI, 80thPercentile PTI, and misery index showed continuity and sensitivity in their variations in response to the increase in variability as the congestion in the peak hour. The BI measure was insensitive to the increase in congestion on the GP lane but showed sensitivity to the congestion for the HOT lanes due to the difference in the rates of changes of standard deviations and medians of travel time with the increase of congestion on these facilities, indicating that the interpretation of the results based on this metric should be done with caution. The FOT and skew statistics showed inconsistent patterns of travel time variation with the increase in congestion on GP and HOT lanes reflecting the differences in the rates of the increase in the median and standard deviation rates of travel time with the increased congestion on the two facilities and the relative contributions of non-recurrent events to the unreliability.

The results also show that examining combinations of time-variant static distribution parameters and reliability metrics of the GP and HOT lanes can provide valuable information on the variations of the reliability of these facilities during the analysis periods and the levels of contribution of nonrecurring events to the unreliability at different times of the day.

The results from the study also confirmed that the use of at least 30 minute periods of analysis is preferable to using longer periods for applications that require fine-grained analysis in order to reasonably represent the reliability pattern during congested periods. In addition, the results from the study confirmed that at least one year’s worth of data should be collected to obtain a more stable value of reliability metrics.

4.3. Safety Measures

4.3.1. Safety Measure Estimation Procedures

Safety performance measures are important indicators of system performance. Starting from fiscal year 2010, each state is required to track a set of safety performance measures that were developed by the National Highway Traffic Safety Administration (NHTSA) and the Governors Highway Safety Association (GHSA) (Herbel et al., 2009). These safety performance measures can be classified into three categories: core measures, behavioral measures, and activity measures. The core measures, also referred to as the outcome measures, include the frequency and rate of crashes, injuries, fatalities. The behavior measures focus on the linkage between safety activities and associated behaviors; for example, observed seat belt usage for passengers. The activity measures track the crash-reduction actions taken by various agencies.

The ITSDCAP tool developed in this project mainly focuses on estimating the core safety measures. Below is a list of the safety measures that can be estimated utilizing the ITSDCAP tool:

- Crash frequency by crash type
- Crash frequency by severity
- Total crash frequency
- Crash rate by type
- Crash rate by severity
- Total crash rate.

The crash frequency by crash type refers to the number of crashes for each crash type such as rear-end, head-on, angle, sideswipe, and so on. The crash frequency by severity is the number of crashes for each severity level; that is, the Property Damage Only (PDO), injury, and fatality. The total crash frequency corresponds to the total number of crashes and does not differentiate by crash type and crash severity. The crash rate by type, crash rate by severity, and total crash rate are defined in the same way, except that these are calculated as the number of crashes per million vehicle miles traveled (MVMT).

4.3.2. User Interface for Safety Measures

Figures 4-12 and 4-13 present the user interface for safety measures in ITSDCAP. As shown in these two figures, the user can select the roadway through the GIS interface. Figure 4-14 shows that the safety performance measures can be either calculated for a whole time period or for each hour within the study period, depending on the user selection of hour type. The input segment file is used to define the segments. The format of this input file is shown in Figure 4-15. As shown in this figure, the segment file includes the following columns: segment ID, segment starting and ending node numbers, corresponding segment starting milepost and ending mileposts, starting hours, volume count for this hour, and the ratio of current hourly volume to the AADT. For some studies that focus on crash data analysis for specific time periods of the day, the hourly volume variation with time of day is important, and therefore either the volume or the hourly volume factor (proportion of daily volume) should be input. After the safety analysis is done, the results are automatically saved to an output folder and also can be visualized through the bottom "Result Display" area. Similar to the mobility performance measures, the safety measure results can be visualized either in a table or a chart format. Depending on the performance measures selected, the users can further specify the crash type and severity type for visualization.

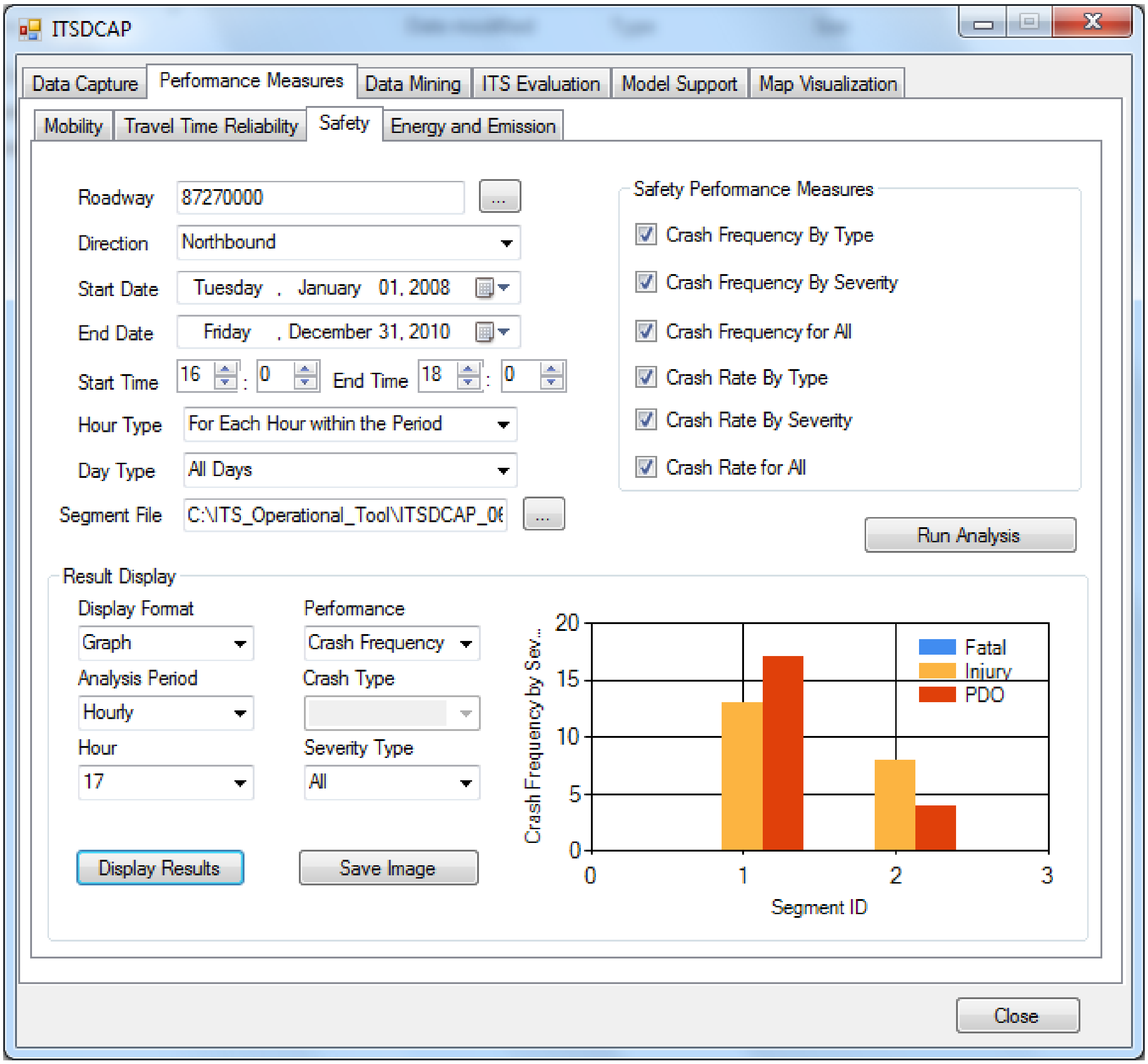


Figure 4-12 User Interface for Safety Measure Calculation

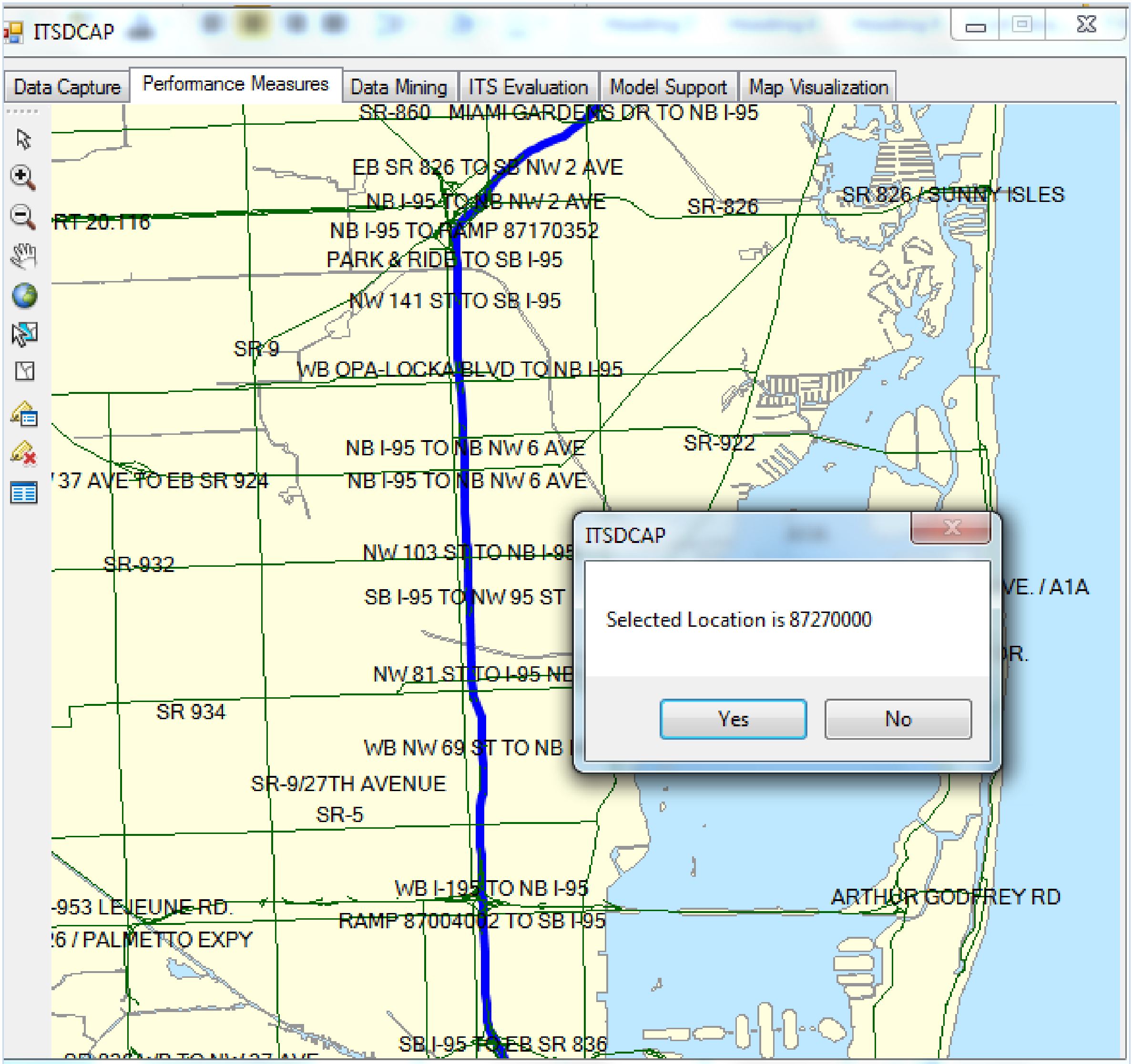


Figure 4-13 Roadway Segment Selection for Safety Measure Calculation

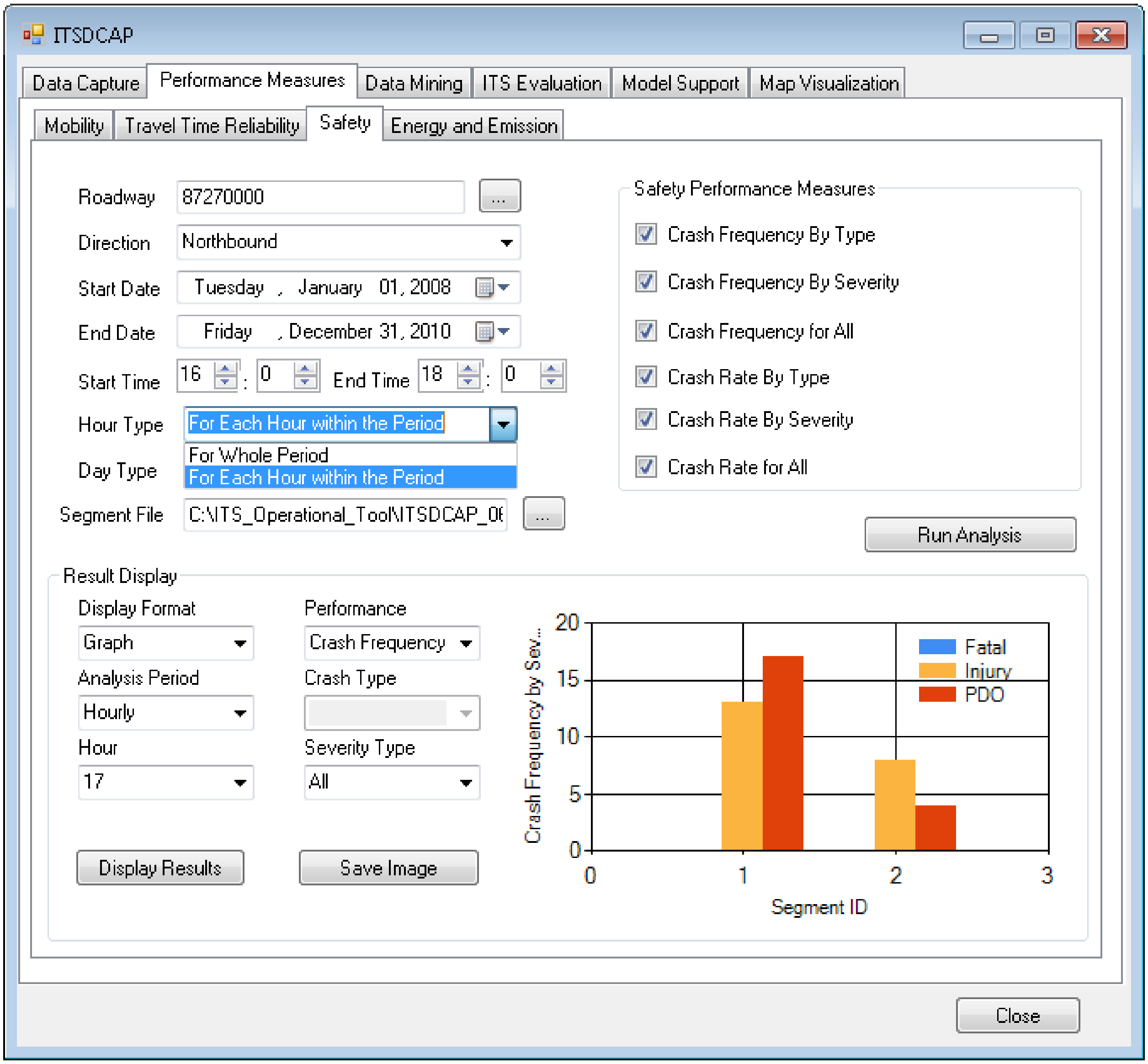


Figure 4-14 Study Period Selection for Safety Measure Calculation

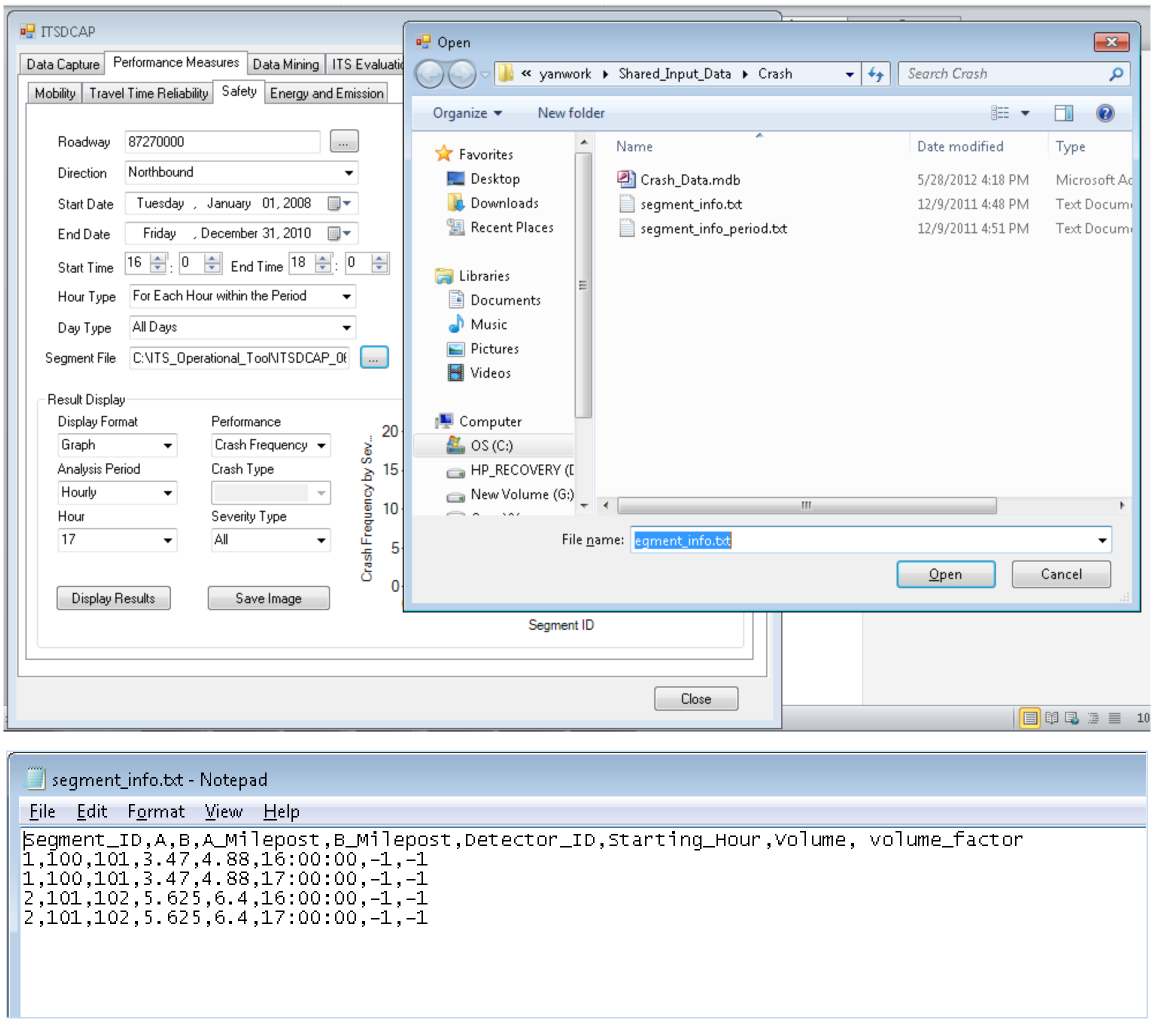


Figure 4-15 Example of Segment Information File for Safety Measure Calculation

4.3.3. Safety Measurement Test Cases

As mentioned in Section 4.1.6, a case study was conducted in this study to test the ITSDCAP tool ability to analyze the mobility and safety of a segment of I-95 SB in Broward County and identified the locations where mobility bottlenecks and safety problems are likely to occur. The results from this analysis are presented in this section.

Figure 4-16 presents the minimum speed within the AM peak period for the normal days as well as the crash frequency for the roadway segments. Note that the links in this figure refers to the predefined segment for crash analysis and that the nodes indicate the detector locations. It is clear from this figure that the segment with both high crash frequencies as well as high congestion in the AM peak is between Hillsborough Blvd. and Commercial Blvd. A short segment close to Miami-Dade County line, in the vicinity of Hallandale Beach Blvd., appears to have a high crash frequency also in the AM peak. These performance measures can be further examined in Figure 4-17 where the time period is divided into one-hour aggregation level. Figure 4-17 shows that the traffic is not congested in the first hour (6:00 AM-7:00 AM), and the number of crashes is also relatively small for this hour; however, in the second and third hours of the AM peak period, the traffic becomes congested and the number of crashes increases. The variation of travel time, vehicle-mile traveled, and vehicle-hour traveled with the time during the AM peak period are presented in Figure 4-18, 4-19, and 4-20, respectively. As shown in these three figures, the travel time, the VMT, as well as the VHT are the highest between 7:40 AM and 8:30 AM.

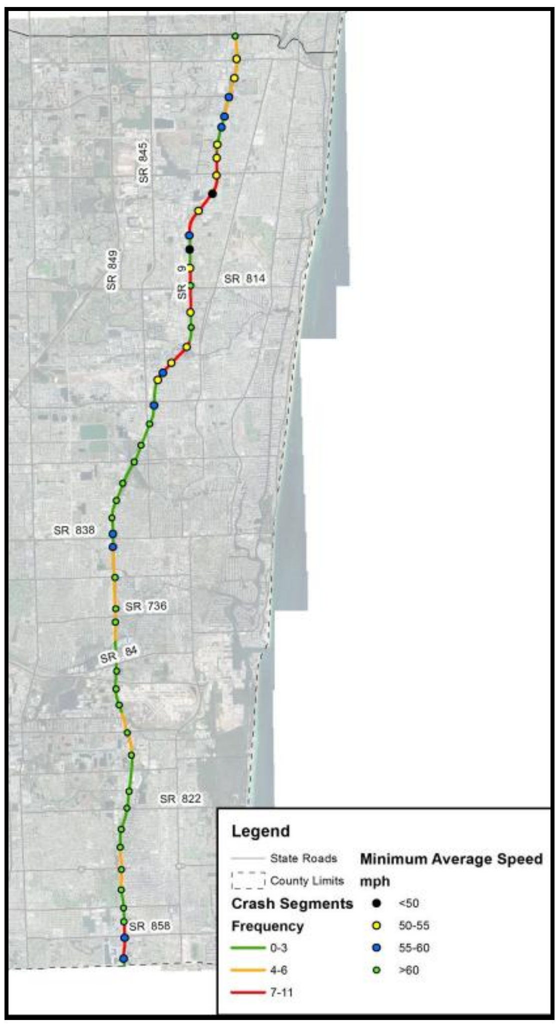
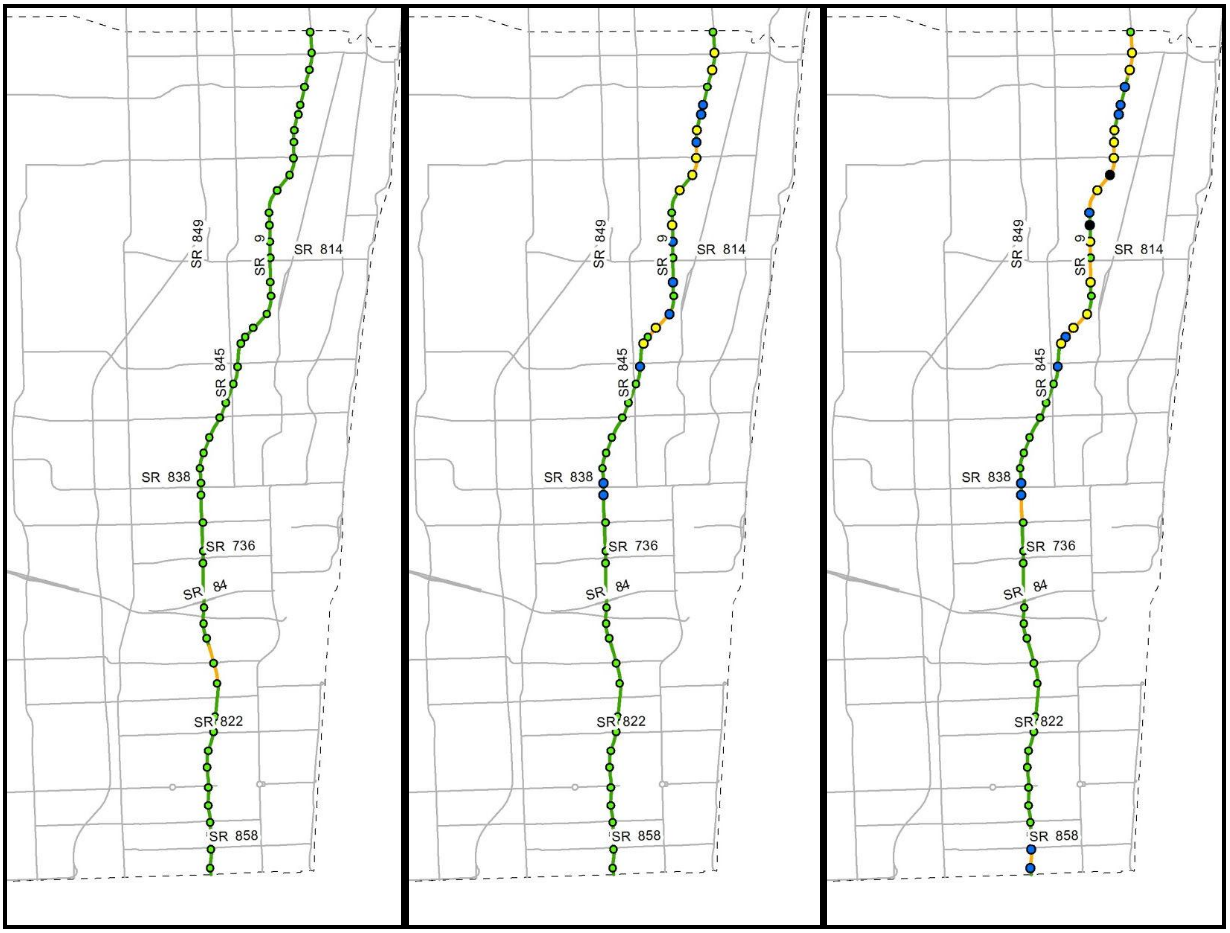


Figure 4-16 Minimum Average Speed for AM Peak Period



6:00am-7:00am 7:00am-8:00am 8:00am-9:00am

Figure 4-17 Minimum Average Speed for Each Hour in AM Peak Period

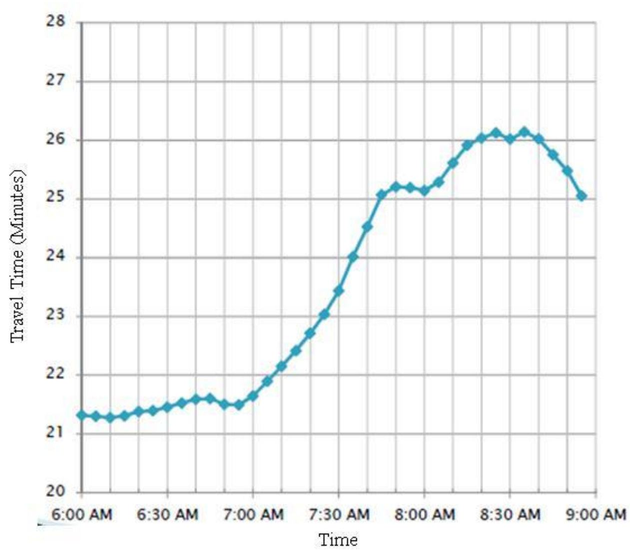


Figure 4-18 Travel Time Variation during the AM Peak Period

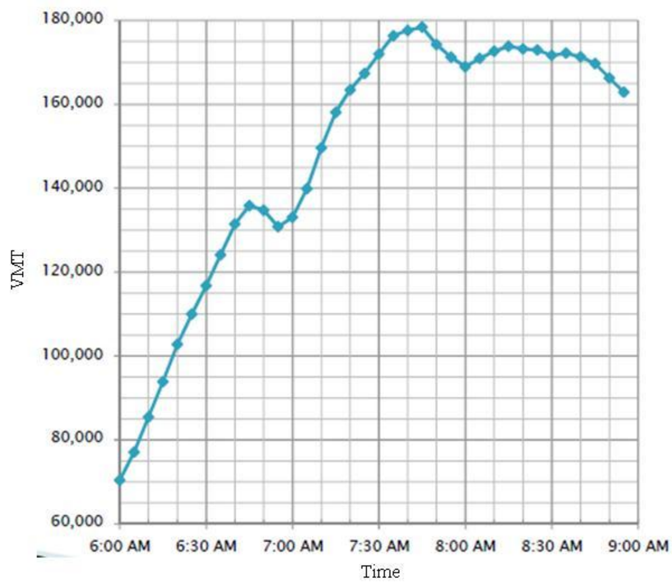


Figure 4-19 VMT Variation during the AM Peak Period

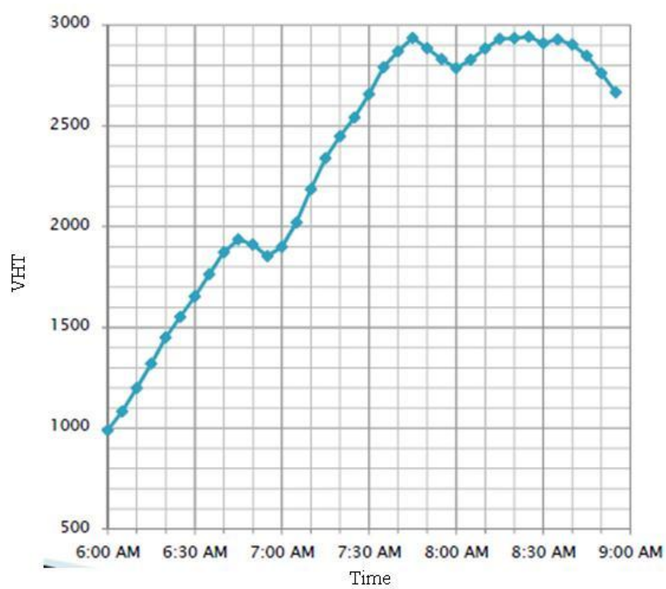


Figure 4-20 VHT Variation during the AM Peak Period

4.4. Energy and Emission Measures

Energy and emission measures are also important elements to be considered in transportation planning and operation processes. The MOBILE6 emission model has been the model used in Florida to estimate emission. The Motor Vehicles Emission Simulator (MOVES) is the latest emission modeling system developed by the U.S. Environmental Protection Agency (USEPA) and it is set to replace the previous emission model, MOBILE6, around the U.S. by 2013.

4.4.1. Emission Models Comparison

MOBILE is an emission factor model developed by EPA for estimating pollution from highway vehicles (USEPA, 2003). The first model, called MOBILE1, was released in 1978 and latest version is MOBILE6.2 released in 2003. MOBILE estimates the amount of emission of pollutants such as hydrocarbon (HC), oxides of nitrogen (NO_x), and carbon monoxide (CO) and other pollutants from passenger cars, motorcycles, and light-and heavy-duty trucks. The emission factors in gram per veh-mile are estimated depending on ambient temperatures, travel speeds, operating modes, fuel volatility and mileage accurate rates for 28 individual vehicle types in low- and high-altitude.

MOVES2010a, released in August 2010, is the latest EPA emission model. Compared to MOBILE6.2, MOVES incorporates substantial new emissions test data and accounts for changes in vehicle technology and regulations and improved in-use emission levels and the factors that influence them. MOVES allows the estimation of emission based on vehicle trajectories (microscopic level of analysis) or based on the VMT (macroscopic or mesoscopic level of analysis).

4.4.2. User Interface for Energy and Emission Measures

Figure 4-21 illustrates the user interface for energy and emission measure calculation in ITSDCAP. Since the vehicle trajectory data is not available in this project, the fuel consumption and emissions have to be estimated based on the vehicle-miles traveled and fuel consumption rates/emission rates. Two types of fuel consumptions are considered: gas and diesel. The default emission rates are currently based on the MOBILE6.2 that are specific to Florida. These rates will be updated when emission rates that are specific to Florida are calculated based on MOVES (the FDOT System Planning Office is currently working on developing such factors). The energy and emission measure results can be visualized either in the table or graphical format, as shown in Figure 4-21. The plotted chart can be further saved to a file by clicking the “Save Image” button for future reporting purposes.

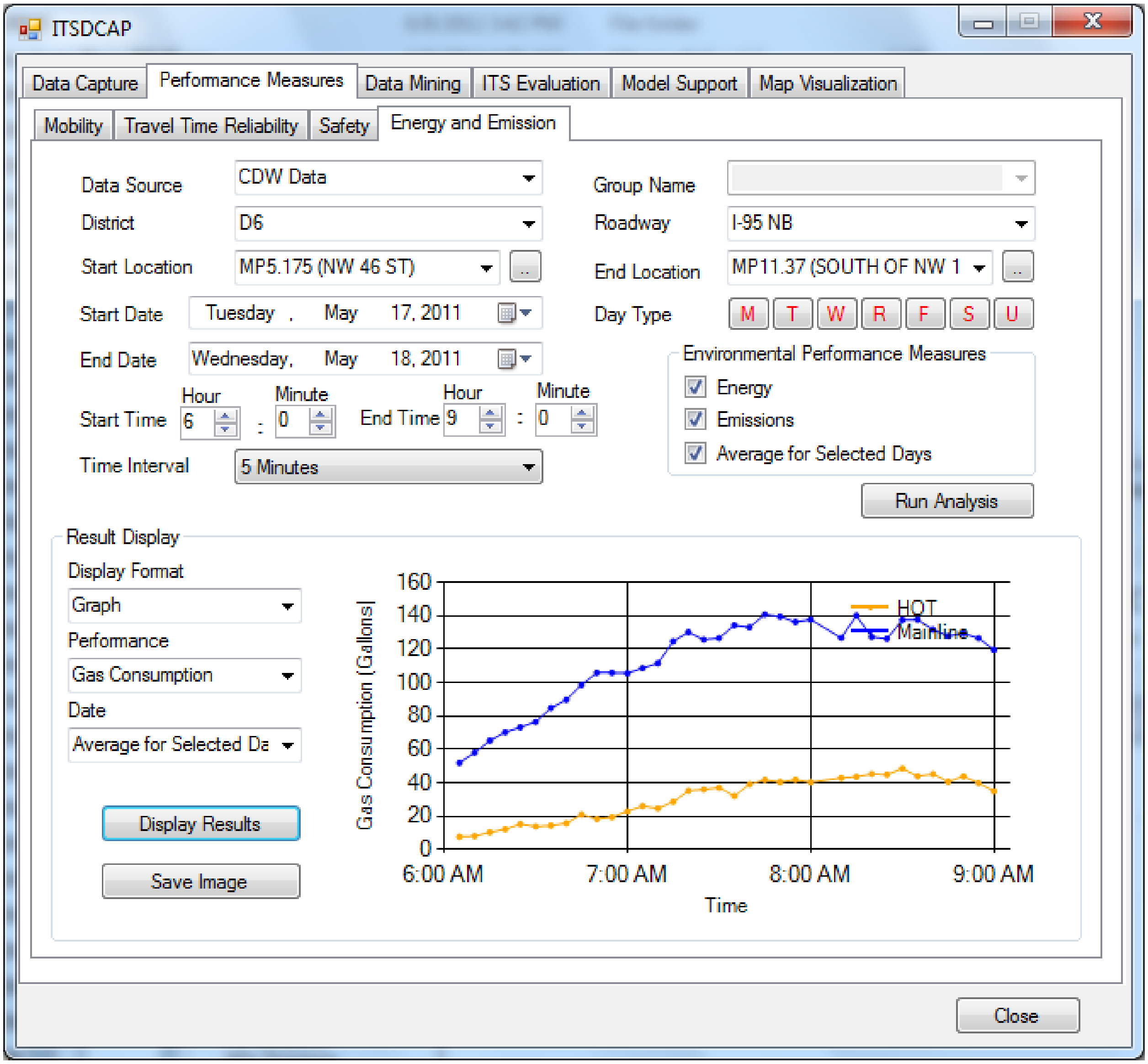


Figure 4-21 User Interface for Energy and Emission Measure Calculation

4.4.3. Emission Estimation Use Case

In this use case, the emission estimated based on Mobile6.2 and MOVES is compared. In this project, an incident occurring along the I-95 Corridor in the Miami-Dade County on January 18, 2011, was used as an example to compare the emissions estimated using these two models. During this incident, two lanes out of four lanes were blocked at the NW 62nd Street interchange for 17 minutes starting from 3:50 PM.

The queue length caused by the incident is verified by the average speed obtained from the upstream detectors as shown in Figure 4-22. This figure is based on traffic data extracted from the STEWARD data warehouse. As shown in Figure 4-22, the incident occurred between detectors D600462 and D600482. The average speed started decreasing and the traffic started queuing at 15:40 upstream of detector D600462. By 4:45 PM, the average speed for all the detectors went back to normal indicating that the queue has dissipated.

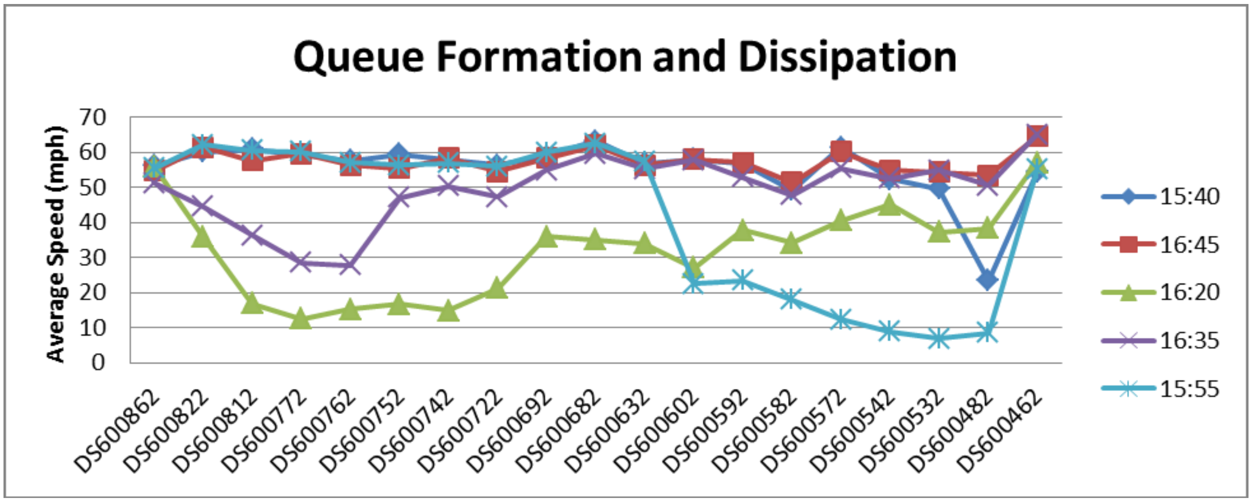
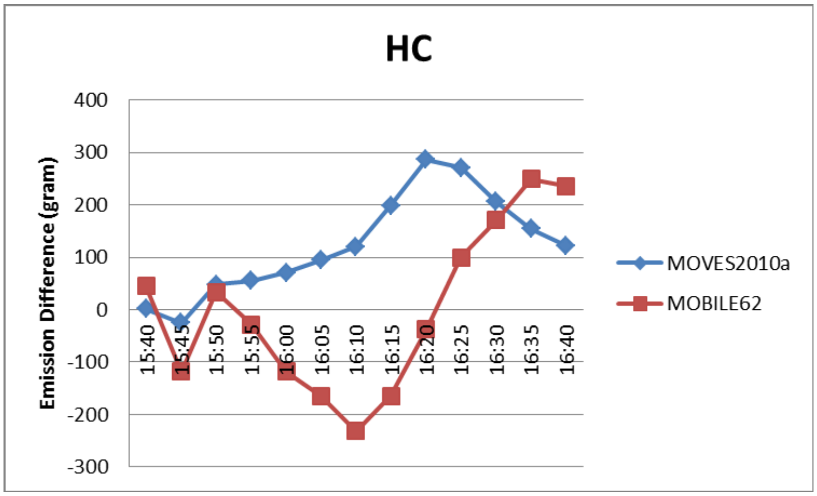
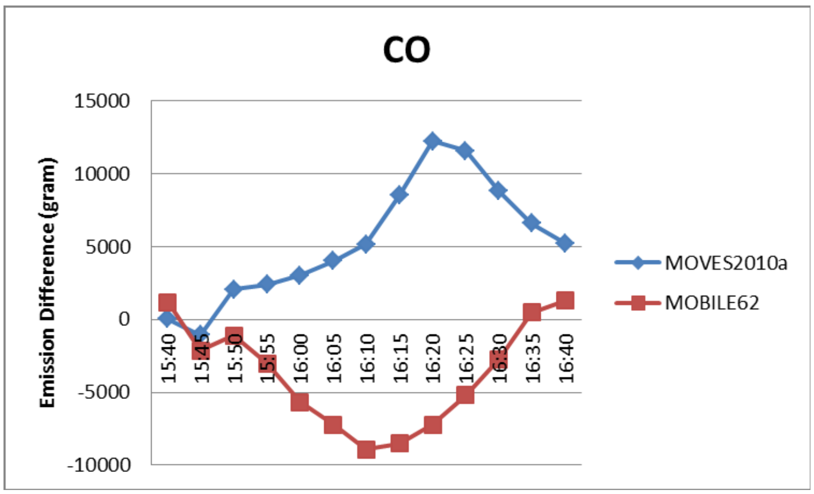


Figure 4-22 Queue Formation and Dissipation for Incident on 1/18/2011

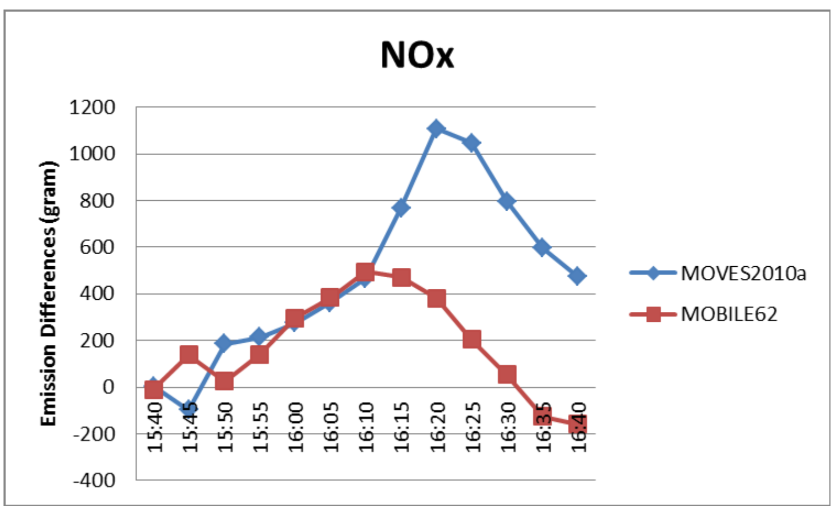
In this case study, a comparison was made of the additional emissions caused by the incident mentioned earlier as calculated using MOVES2010a and MOBILE6.2. The results are presented in Figure 4-23. Note that the emission difference in these figures refers to the difference in emissions between the incident day and normal conditions. The comparisons show that MOVES2010a estimates higher emission due to the incident compared to Mobile6.2 indicating that MOVE is more sensitive to the reduction in speed. It can be concluded that these two models may produce quite different emission results when assessing incident impacts.



(a)



(b)



(c)

Figure 4-23 Additional Emissions due to the Incident

4.5. References

Boyles S., Voruganti A., and Waller T. Quantifying Travel Time Variability in Transportation Networks. Research Report SWUTC/10/167275-1, Austin, Texas, 2010.

Cambridge Systematics, Texas A&M University, Dowling Associates, Street Smarts, H. Levinson and H. Rakha. Analytical Procedures for Determining the Impacts of Reliability Mitigation Strategies, 2010.

Federal Highway Transportation. Archived Data Management Systems - A Cross-Cutting Study. Publication FHWA-JPO-05-044. FHWA, U.S. Department of Transportation, 2004.

Florida Department of Transportation Traffic Engineering and Operations Office (FDOT). Florida Statewide Operations Performance Measures and Data Collection, 2008.

Georgia Department of Transportation (GDOT), Strategic Planning, 2011.

Herbel, S., M.D. Meyer, B. Kleiner, and D. Gaines. A Primer on Safety Performance Measures for the Transportation Planning Process. Final Report, FHWA-HEP-09-043, September, 2009.

Li, R., G. Rose, and M. Sarvi. Using Automatic Vehicle Identification Data to Gain Insight into Travel Time Variability and its Causes. 85th Annual Meeting of the Transportation Research Board, Washington D.C., 2006.

Lomax T., D. Schrank, S. Turner, and R. Margiotta. Report for Selecting Travel Reliability Measures, 2003.

McLeod, D.S. and G. Morgan. Florida's Mobility Performance Measures and Experience. In *Transportation Research Board 91th Annual Meeting*. CD-ROM. National Academies Press, Washington D.C., 2012.

National Transportation Operations Coalition (NTOC), Performance Measurement Initiative, Final Report, 2005.

Pu, W. Analytic Relationships between Travel Time Reliability Measures. In *Transportation Research Board 89th Annual Meeting*. CD-ROM. National Academies Press, Washington D.C., 2010.

Pu W., A. Messe, J. Andrew and L. Long. Exploring Travel Time Reliability under Different Circumstances: A Case Study. In *Transportation Research Board 89th Annual Meeting*. CD-ROM. National Academies Press, Washington D.C., 2010.

Rakha H., I. El-Shawarby, and M. Arafeh. Trip Travel Time Reliability: Issues and Solutions. Intelligent Transportation Systems Conference, ITSC, 2007.

Second Strategic Highway Research Program (SHRP2), http://www.trb.org/StrategicHighwayResearchProgram2SHRP2/Blank2.aspx, Accessed on July 29th, 2011.

Southwest Research Institute. http://sunguide.datasys.swri.edu, Accessed on July 29th, 2011.

Transportation Research Board (TRB). *Highway Capacity Manual*, Washington, D.C., 2010.

Tu H., J. van Lint, and H. van Zulen. Travel Time Reliability Model on Freeways, In *Transportation Research Board 87th Annual Meeting*. CD-ROM. National Academies Press, Washington D.C., 2008.

Tu, H., J. van Lint, and H. van Zulen. The Effects of Traffic Accidents on Travel Time Reliability. Proceedings of the 11th International IEEE Conference on ITS, Beijing, China, 2008.

United States Environmental Protection Agency (USEPA). User's Guide to MOBILE6.1 and MOBILE6.2. EPA420-r-03-010, August 2003.

United States Environmental Protection Agency (USEPA). Motor Vehicle Emission Simulator (MOVES)- User Guide Version, MOVES2010a, Assessment and Standards Division Office of Transportation and Air Quality, 2010.

Van Lint J., and H. van Zulen. Monitoring and Predicting Freeway Travel Time Reliability, In *Transportation Research Board 87th Annual Meeting*. CD-ROM. National Academies Press, Washington, D.C., 2005.

Washington State Department of Transportation (WSDOT), Accountability and Performance Information, 2011.

Xiao, Y. Hybrid Approaches to Estimate Freeway Travel Times Using Point Traffic Detector Data. Ph.D. Dissertation, Florida International University, Miami, FL, 2011.

5. Modeling Support

The Modeling Support module in ITSDCAP provides data to support different types of models, including demand forecasting, dynamic traffic assignment (DTA), macroscopic models, mesoscopic simulation models, and microscopic simulation. The module-provided data are extracted from different sources, processed as needed, and outputted in standard formats for use in various modeling activities.

5.1. Data Sources Used in Modeling Support Module

Four databases are used as the sources for the Modeling Support module in ITSDCAP: TSS data, detector data from STEWARD data warehouse, TVT data, and INRIX data. As mentioned in previous section, TSS data and STEWARD data are based on point traffic detectors in the current version, and include speed, volume, and occupancy measurements. Therefore, they can be used to calculate several traffic parameters for use in modeling support. However, the TVT and INRIX data only consist of travel time for some predefined paths and can only be used for average travel time estimation for calibration purposes.

5.2. Modeling Support Module Functions

Several functions are provided by the Modeling Support module to support the development, validation, and calibration of traffic models. These functions include estimating average speed, volume, and density for normal days or user-defined days, calculating diurnal factor for demand modeling, estimating free-flow speed, estimating maximum throughput, estimating travel time, and fitting traffic flow models. A detailed description of each function is provided in sections 5.2.1 through 5.2.6.

5.2.1. Average Speed, Volume, and Density

Travel demands on normal days are required for all types of models. The average speed and density of normal days and/or incident days are also required for model calibrations. Therefore, the first function in the Modeling Support module in ITSDCAP is to estimate the average speed, volume, and density for different types of days. As shown in Figure 5-1, three types of days can be defined in the interface: all days, normal days, and user-defined days. The first type of day includes all the days within the range of the specified starting and ending dates. The second type, normal days, is obtained by eliminating the days with incidents, construction, or special events, and the days that are holidays/weekends. However, the users can also identify the normal days or incident days by using the Data Grouping module described in Chapter 3 to obtain a set of specific days and input them to the Modeling Support module as user-defined days. These types of days can be further combined with the selection of day type in the interface to limit the selection further to certain days of the week, as shown in Figure 5-1. Once the specific days are defined; the average speed, volume, and density for those days can be estimated using the same procedures as described in mobility performance measure calculation module, discussed in Chapter 4.

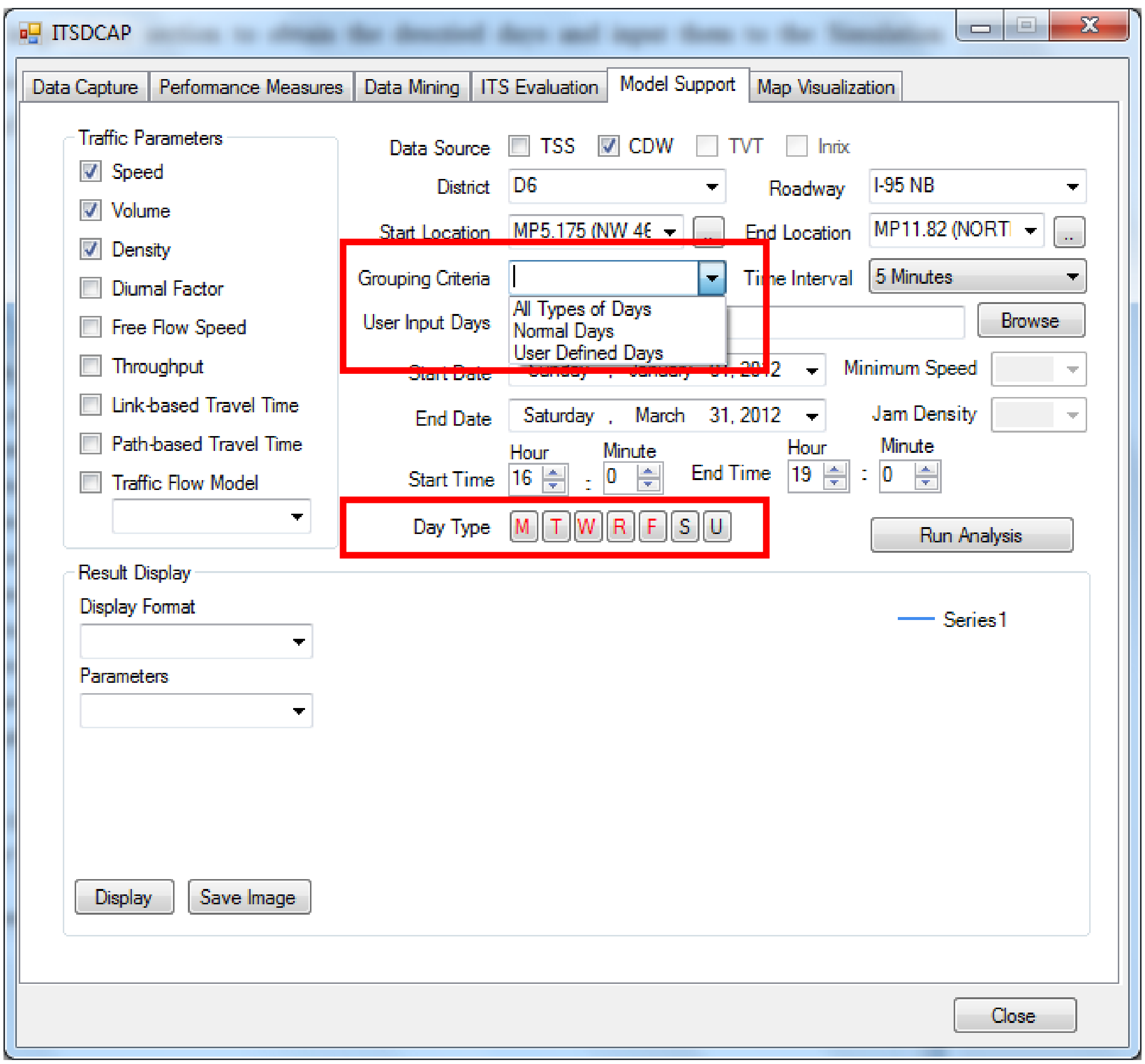


Figure 5-1 Modeling Support Module Interface in ITSDCAP

5.2.2. Diurnal Factor

One of the simple and straightforward ways to convert a daily trip matrix to a period, hourly, or sub-hourly matrix in assignment models is to apply diurnal factors. In this project, the average volume for specified days as mentioned above are first calculated, and then the diurnal factors are obtained by dividing the volume count in each time interval by total volume for the whole study period.

5.2.3. Free-Flow Speed

Free-flow speed is one common input required for transportation modeling. In the INRIX data, the free-flow speed is directly reported. However, free-flow speed has to be estimated when using detector data as the data source. In the HCM 2010, the free-flow speed is defined as the mean speed of passenger cars when the traffic flow is low to moderate (up to 1,000 pc/h/ln). In this project, this definition is used to estimate the free-flow speeds based on speed data from either the TSS data or from the STEWARD data warehouse.

5.2.4. Throughput

Throughput is useful when identifying bottlenecks and also estimating queuing discharge rates. Currently, ITSDCAP produces the maximum throughput by searching the maximum volume counts for a user-specified study period. In the future version of ITSDCAP, this will be modified to allow the user to specify the percentile of volume counts instead of the maximum values in throughput calculations; for example, the 95th percentile or 90th percentile.

5.2.5. Average Travel Time

Travel time is the parameter commonly used in model calibration. Depending on the specific application purpose, either link-based travel time or path-based travel time may be needed. The procedures to estimate the average travel time are similar to those used in mobility performance measure calculation, and therefore the description of these procedures is omitted here for brevity.

5.2.6. Fitting Traffic Flow Model

Traffic flow models are required by macroscopic simulation models as well as mesoscopic simulation models. For example, the Bureau of Public Roads (BPR) curve is used in the Florida Travel Demand Models (FSUTMS), and the modified Greenshields models are used in the mesoscopic dynamic traffic assignment software DynusT (Chiu et al., 2012) and its parent software Dynasmart. In the Modeling Support module of ITSDCAP, four types of traffic flow models can be fitted using detector data: the BPR Curve, the Greenshields Model, the Single-Regime Modified Greenshields Model, and the Double-Regime Modified Greenshields Model as shown in Figure 5-2. In future versions, fitting other types of models may be supported. Descriptions of each traffic flow model follow:

The BPR Curve

The expression of the BPR Curve is expressed as:

$$t = t_0\left[1+\alpha(\frac{q}{C})^{\beta}\right] \tag{5-1}$$

where t represents the travel time and t_0 is the free-flow travel time. Continuing, q and C are the flow rate and capacity, respectively, and α and β are the coefficients in the BPR curve. To fit the BPR Curve using detector data, the expression in Equation 6-1 is transformed as follows:

$$\frac{1}{s} = \frac{1}{s_f} + \frac{\alpha}{s_f}(\frac{q}{C})^{\beta} \tag{5-2}$$

where s is the speed and s_f is the free-flow speed. The relationship between $1/s$ and q/C in Equation 5-2 is fitted using the least-square curve fitting function in the MATLAB program (The MathWorks Inc., 2010). MATLAB was also used to fit the other relationships discussed in this section. This MATLAB code for curve fitting is converted to a dynamic-link library and integrated in the ITSDCAP tool. Note that users have the option to input the capacity value. If the capacity value is not provided, the program will use the maximum flow rate estimated based on detector data, as the capacity. The curve fitting results including the free-flow speed, α and β coefficient values, capacity, and curve fitting errors will be output by the program.

The Greenshields Model:

The Greenshields Model has expressions as follows:

$$s = s_f - \frac{s_f}{k_j} k \tag{5-3}$$

or

$$q = k_j (s - \frac{s^2}{s_f}) \tag{5-4}$$

where k symbolized the density and k_j is the jam density. To fit this model, the Equation 5-4 is converted to the following format:

$$\frac{q}{s} = -\frac{k_j}{s_f} s - k_j \tag{5-5}$$

A linear curve fitting is used in the program to find the jam density and free-flow speed parameters. The R^2 value that quantifies the curve fitting performance is also reported.

Single-Regime Modified Greenshields Model

Equation 5-6 presents the Single-Regime Modified Greenshields model expression:

$$s - s_0 = (v_f - s_0)(1 - -\frac{k}{k_j})^{\alpha} \tag{5-6}$$

where s, k, and k_j have the same meaning as above expressions. The term s_0 represents the minimum speed, v_f denotes the speed-intercept, and α is a coefficient in this model. Similar to the other models, this expression is transformed as follows for a linear curve fit.

$$\ln(s - s_0) = \ln(v_f - s_0) + \alpha \ln(1 - -\frac{k}{k_j}) \tag{5-7}$$

Since the traffic conditions that correspond to the minimum speed and jam density may not be covered in the input detector data, the users have an option to specify these values as input. However, if such information is not input, the program will use the minimum speed and maximum density that are observed in the data set by default.

Dual-Regime Modified Greenshields Model

The expression for the Dual-Regime Modified Greenshields Model is listed below:

$$
\begin{aligned}
&s = s_f && 0 \le k \le k_{bp} \\
&s - s_0 = (v_f - s_0)(1 - -\frac{k}{k_j})^{\alpha} && 0 \le k \le k_{bp}
\end{aligned}
\tag{5-8}
$$

where k_{bp} is the density at the breakpoint for two modeling regimes, and s_f is the free-flow speed. The other variables are as defined above. A similar curve fitting procedure as that reported by Mahmassani et al. (2009) is used in this project, as summarized below:

- Step 1: Calculate the minimum speed and maximum density in the data set, if the value of minimum speed and jam density is not input by users
- Step 2: Select one initial value of breakpoint density, and divide the data set into two subsets based on the breakpoint density
- Step 3: Calculate the average speed using the data in first regime, which is the estimate for free-flow speed
- Step 4: Use linear regression to fit the expression as shown in Equation 6-7 and find the parameters of speed-intercept and α
- Step 5: Calculate the root mean square error based on the fitted value and real-world data
- Step 6: Increase the break point density by 0.5 vpmpl, and repeat the steps 3-5
- Step 7: Compare the fitting errors for each breakpoint density and the one with minimum error will be considered as the best fit

It should be pointed out that traffic flow model fitting should be done only for the bottlenecks in the network due to the capacity constraint of the link itself, as the maximum throughput at a selected location may be limited by a downstream bottleneck or affected by non-recurrent events such as incidents and constructions.

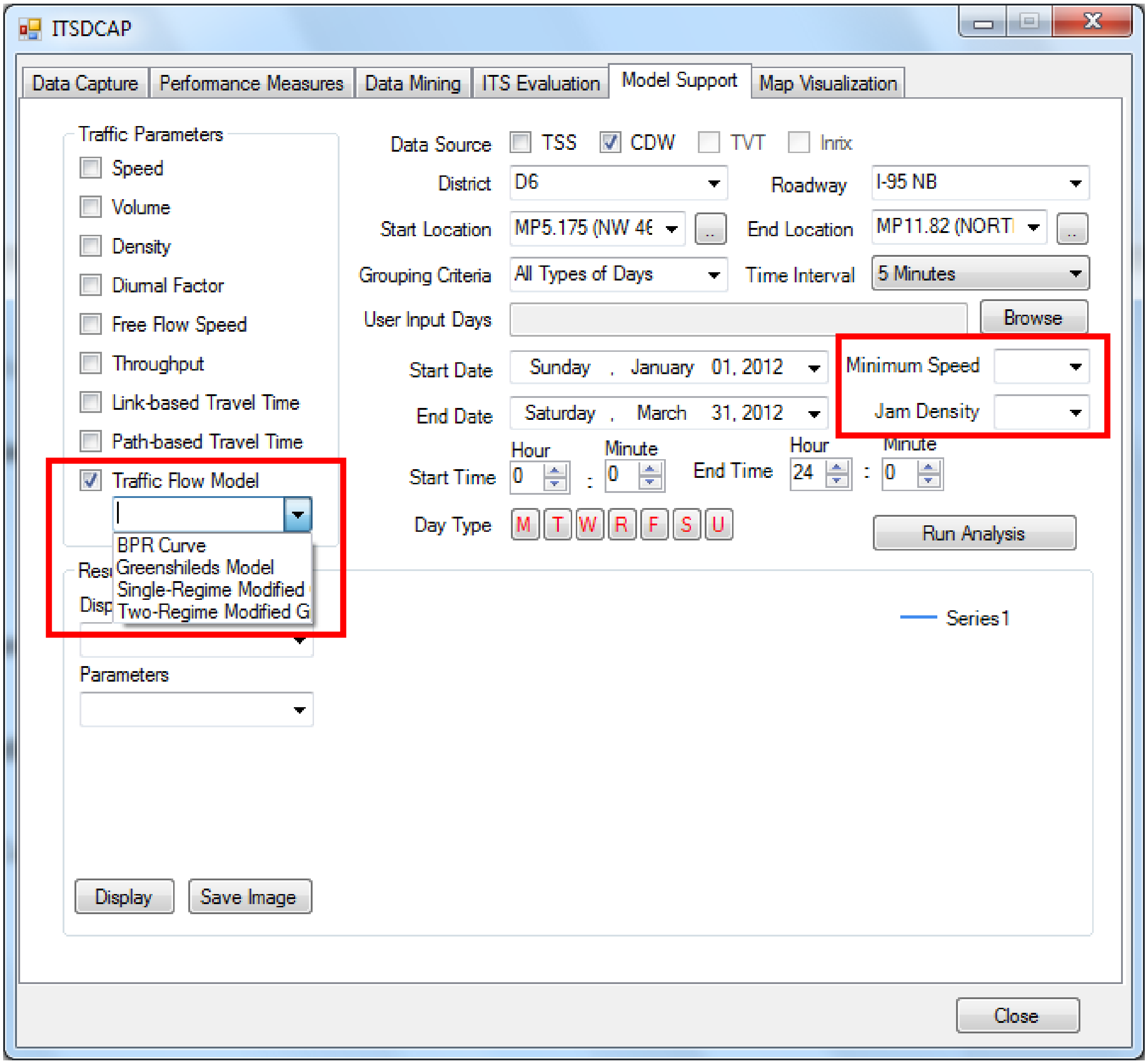


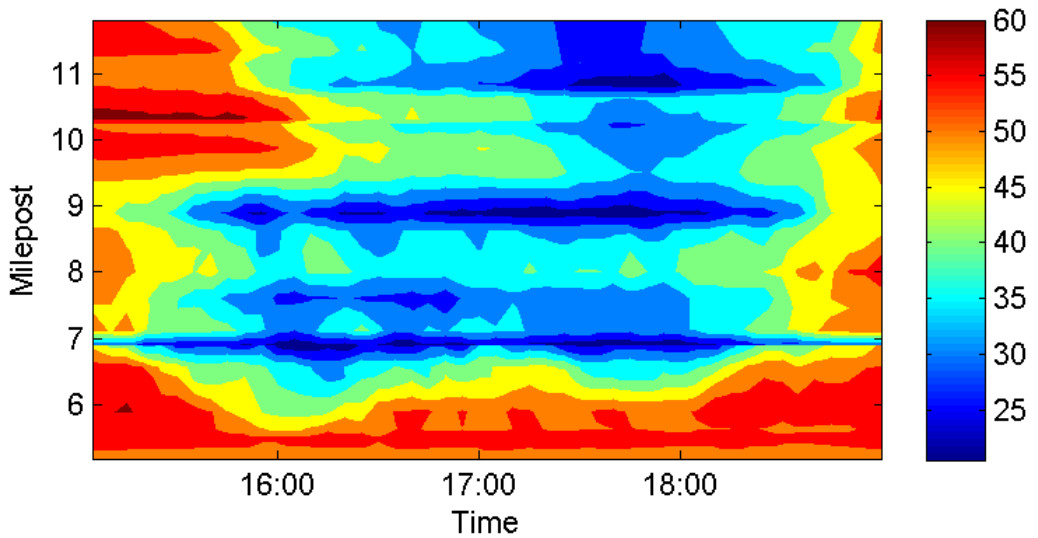
Figure 5-2 Traffic Flow Model Selection

5.3. Case Studies

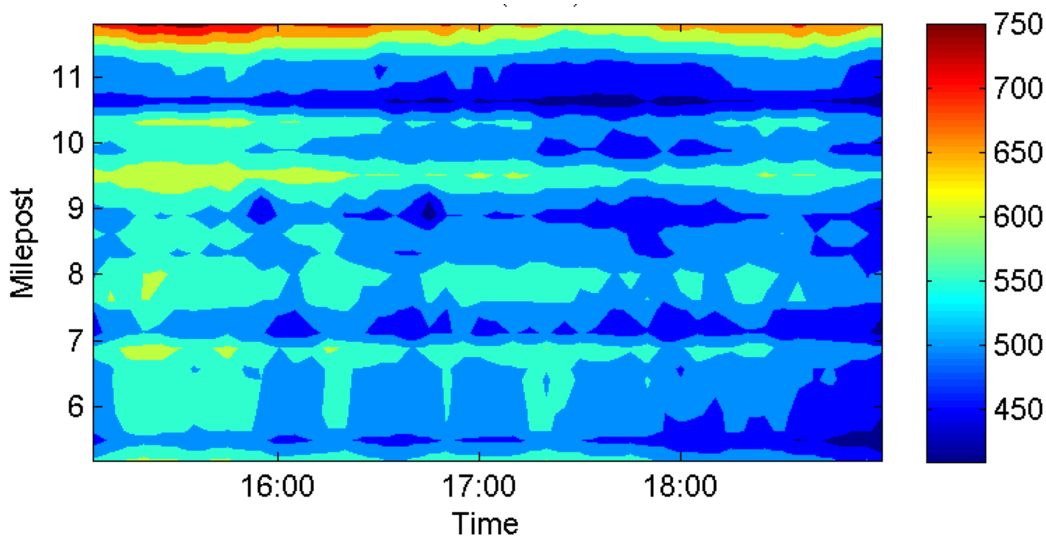
This section presents case studies as examples to illustrate the functions of the Modeling Support Module. In these case studies, the roadway section between NW 46th St. and NW 151st St. along the I-95 NB Corridor in Miami-Dade County was used as the study corridor. This case study was selected to help calibrate a DTA/mesoscopic model for the corridor.

Case study 1: Average Normal Day Speed and Volume

This case study aims at calculating the average normal day speed and volume for the roadway section mentioned above. The time period is between 3:00 PM and 7:00 PM on weekdays from Jan. 1, 2011 and Apr. 30, 2011. As shown in Figure 5-3(a) and (b), the congestion starts from Milepost 7 and 9 around 3:00 PM due to the relative high demand. At about 4:30 PM, the traffic becomes even worse with the propagated congestion resulted from downstream bottleneck. The recurrent congestion dissipates around 7:00 PM. Such a speed contour is very useful in calibrating simulation models to ensure that the real world impacts of bottlenecks are reflected in the calibrated simulation models.



(a) Speed (VPH)



(b) 5-Minute Volume Count

Figure 5-3 Average Normal Day Speed and Volume

Case study 2: Diurnal factor

Figure 5-4 presents the diurnal factor for the detector station at the Milepost 5.89 close to NW 58th St. based on the normal weekday data between Jan. 1, 2011 and Apr. 30, 2011. The time interval for data used in this case study is 15 minutes. As shown in this figure, the lowest volume count occurs around 4:00 AM and the highest volume occurs around 8:00 AM.

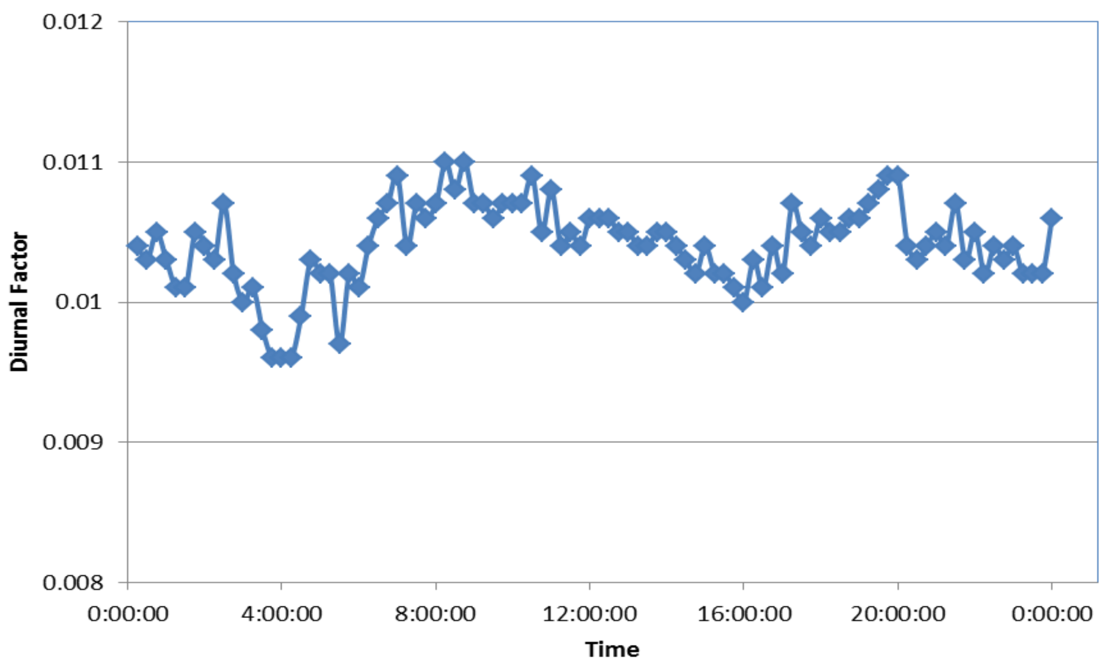


Figure 5-4 Diurnal Factor

Case study 3: Free-Flow Speed

The free-flow speed for the study corridor is estimated using the STEWARD data from Jan.1, 2012 to Mar. 31, 2012. Figure 5-5 presents the estimation results for free-flow speed.

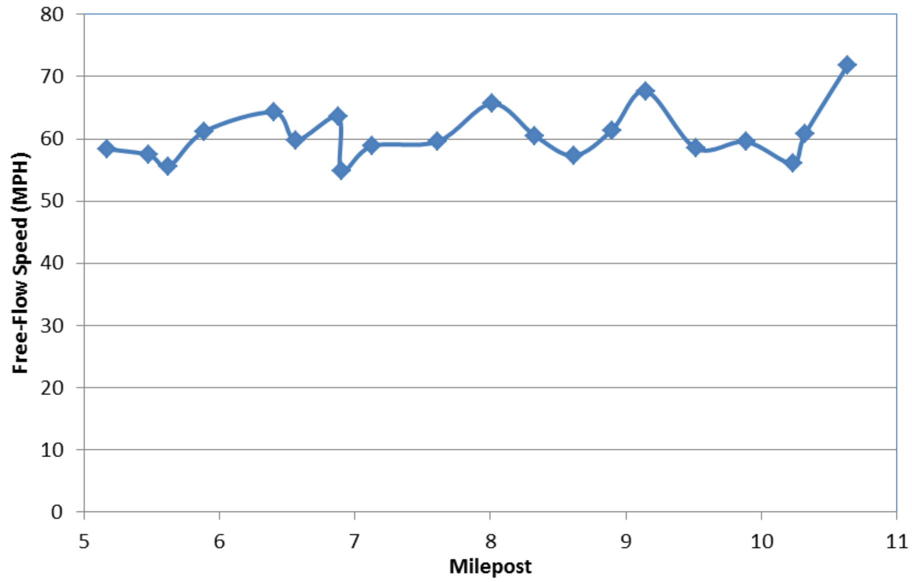


Figure 5-5 Free-Flow Speed

Case study 4: Average Travel Time

In this case study, the average travel time for segments along the study corridor on normal weekdays from Sept. 1, 2011 and Dec. 30, 2011, is calculated using the INRIX Data. Figure 5-6 presents the spatial and temporal distribution of travel time results. This figure shows that the traffic at the segments that are close to Mileposts 7, 9, and 11 are more congested during the PM peak period.

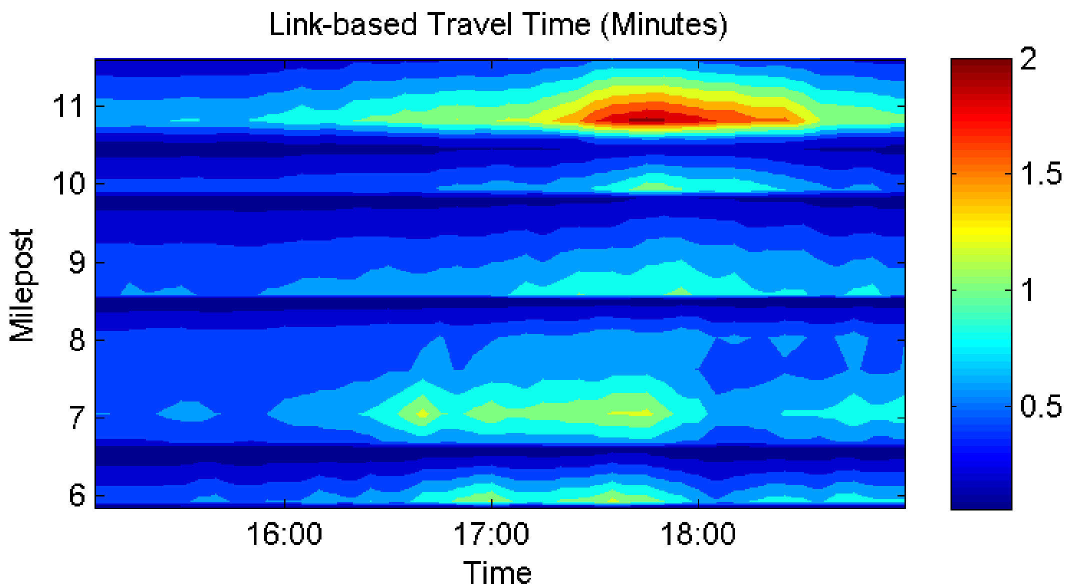


Figure 5-6 Average Normal Day Travel Time Based on INRIX Data

Case study 5: Dual-Regime Modified Greenshields Model

Figure 5-7 presents the fitting results for the Dual-Regime Modified Greenshields Model at Milepost 8.895 located north of NW 103rd St. based on the normal day detector data between Jan. 1, 2011, and June 30, 2011. Table 5-1 lists the fitted parameters for this example.

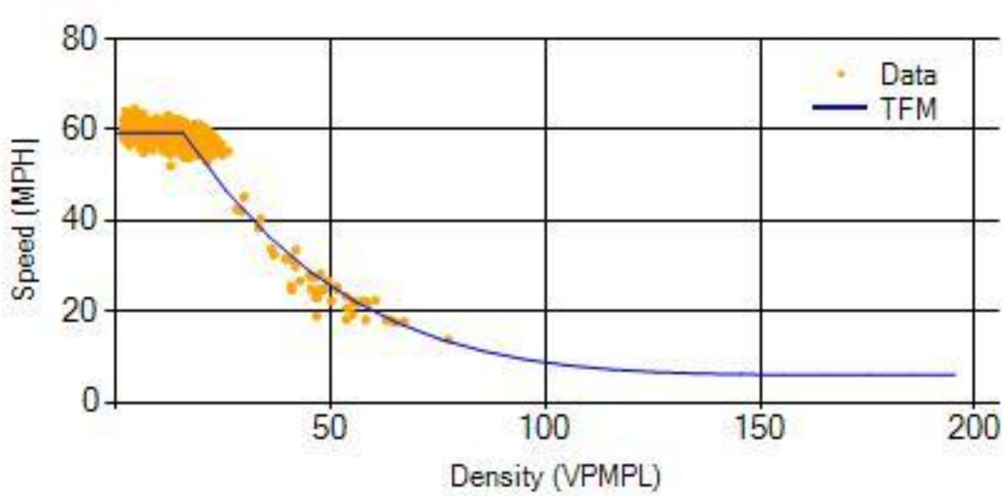


Figure 5-7 Traffic Flow Model Fitting Results

Table 5-1 Dual-Regime Modified Greenshields Model Fitting Results in Case Study 5

Parameters	Values
Break-Point Density	15.5 vpmpl
Free-Flow Speed	59 mph
Speed-Intercept	85 mph
α	4.86
Minimum Speed	6 mph
Jam Density	200 vpmpl
R-Square for Regime 2	0.915

5.4. References

Chiu, Y.C., E. Nava, J. Shelton, K. Milster, T. Smith, M. Munden, B. Bustillos, A. Sherrick, A. Paz, and C. Kou. DynusT Online User's Manual, Version 3.0.1., http://dynust.net/wikibin/doku.php?id=start, Accessed May 20, 2012.

Mahmassani, H.S., J. Dong, J. Kim, R.B. Chen, and B. Park. Incorporating Weather Impacts in Traffic Estimation and Prediction Systems. Final Report, FHWA-JPO-09-065, EDL# 14497, September 2009.

The MathWorks Inc. *MATLAB version 2010a*. Natick, Massachusetts, 2010.

Transportation Research Board, *Highway Capacity Manual*, Washington, D.C., 2010.

6. Map Visualization

Even though the estimated performance measurements and traffic flow parameters can be visualized through table or chart format in other modules of ITSDCAP, it is useful to visualize these spatial-related results on an ArcGIS map. Therefore, a Map Visualization module is provided in ITSDCAP and is described in this chapter.

6.1. Map Visualization Module Description

The Map Visualization module in ITSDCAP includes two tabs, as shown in Figure 6-1. The first tab is for visualization settings and the second tab is used to display the map and visualize animations. Depending on the performance measures or traffic parameters of interest, the shape file for map visualization can be either a point-based or link-based file. For example, the speed and volume reported by detectors are point measurements while the travel time and speed from INRIX Data are roadway segment-based measurements. The GIS shape file and the corresponding data file based on previous module results are input through the visualization settings tab, providing the user the flexibility to visualize any spatial-dependent performance measurements. For maximum flexibility, the column numbers for time variables, spatial locations and traffic parameters to be displayed on the map based on the input data file are also specified by users. Note that the input of "column number for label within the shape file" is used to display the labels on maps, helping users to have a better understanding of the locations.

The map visualization shows the parameter variations at all shape locations. However, sometimes the user may be interested to see one or more parameters at specific locations. For example, a user may want to visualize the changes in speed on a map while visualizing the corresponding volume or density changes. Considering this, the ITSDCAP tool allows the users to select shapes on maps and simultaneously display the time variation of one or more parameters in the charts, as illustrated in Figure 6-4. The data file for the chart and the corresponding column numbers for plotting the charts are inputted through the Visualization – Settings tab.

As shown in Figure 6-2, there are three major components in the display windows: GIS map, map toolbar, and animation toolbar. The GIS map area is used to display the results. The map toolbar provides basic map functions, including selection, zoom in, zoom out, and zoom to extent, pan, adding or removing labels and legends, and configuring display colors. Figure 6-3 presents an example of color settings. As shown in this figure, the user can specify a threshold for each category and the associated display color. If the parameter to be visualized on the map is a time-dependent variable, the animation toolbar will be enabled, allowing the users to play the animation and also control the animation speed. As mentioned above, an additional chart for simultaneously visualizing the parameters at certain locations can appear in this window, depending on the user input.

6.2. Examples of Map Visualization

In this section, three case studies are presented to illustrate the map visualization in ITSDCAP. Case Study 1 includes visualizing the average speed and volume estimated from point-detector data for I-95 NB in Miami-Dade County and is depicted in Figures 6-1, 6-2, 6-3 and 6-4. Case Study 2, shown in Figure 6-5 and 6-6, includes visualizing lane incident frequency for I-95 SB in Miami-Dade County. Different from Case Study 1, the incident frequency in this case is a time-independent variable. Case Study 3, depicted in Figure 6-7 and 6-8, involves visualizing the time-dependent travel time and speed from INRIX Data for I-95 NB in Miami-Dade County. In this case study, the travel time and speed are link-based parameters.

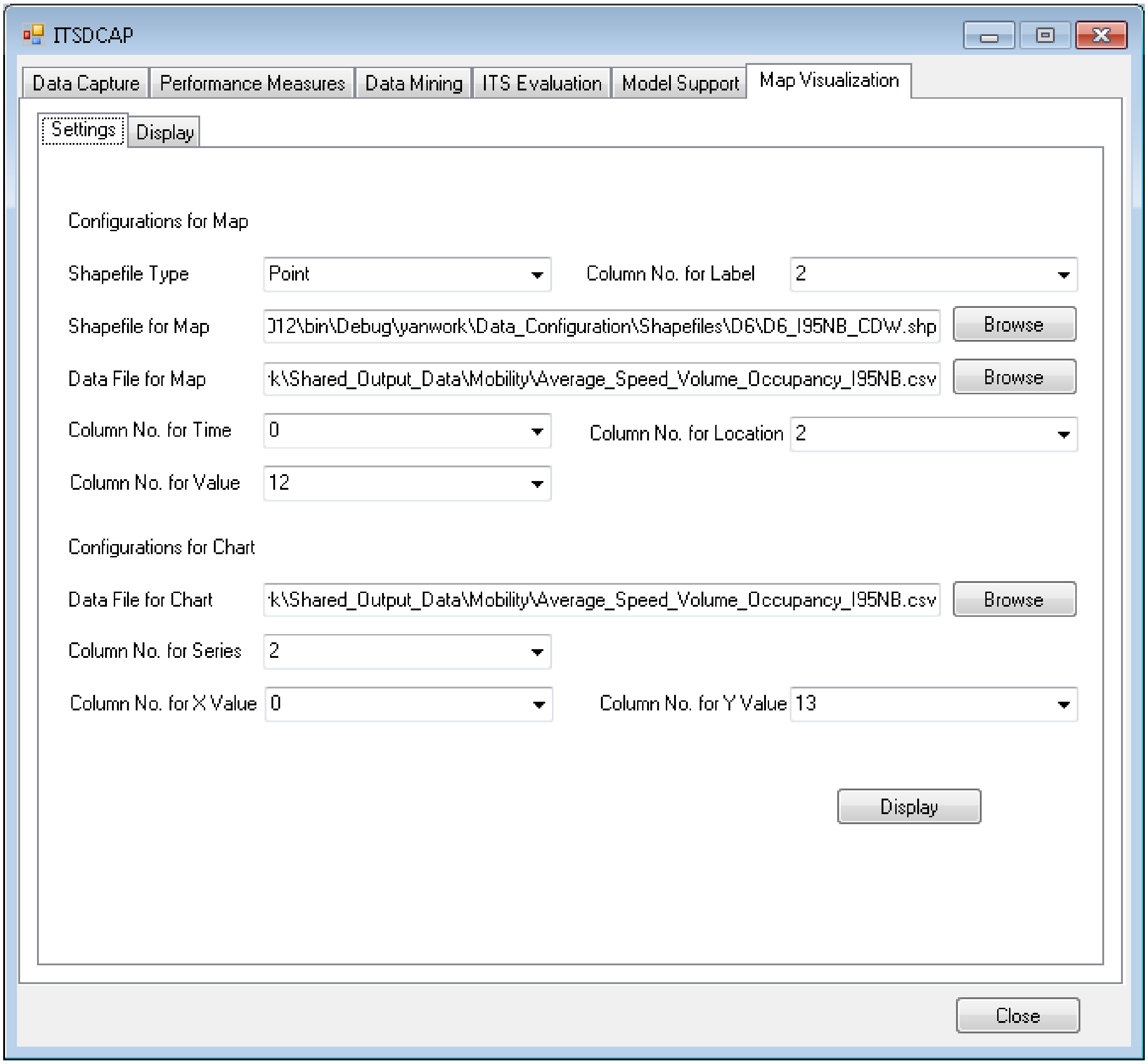


Figure 6-1 Input Interface for Map Visualization Case Study 1

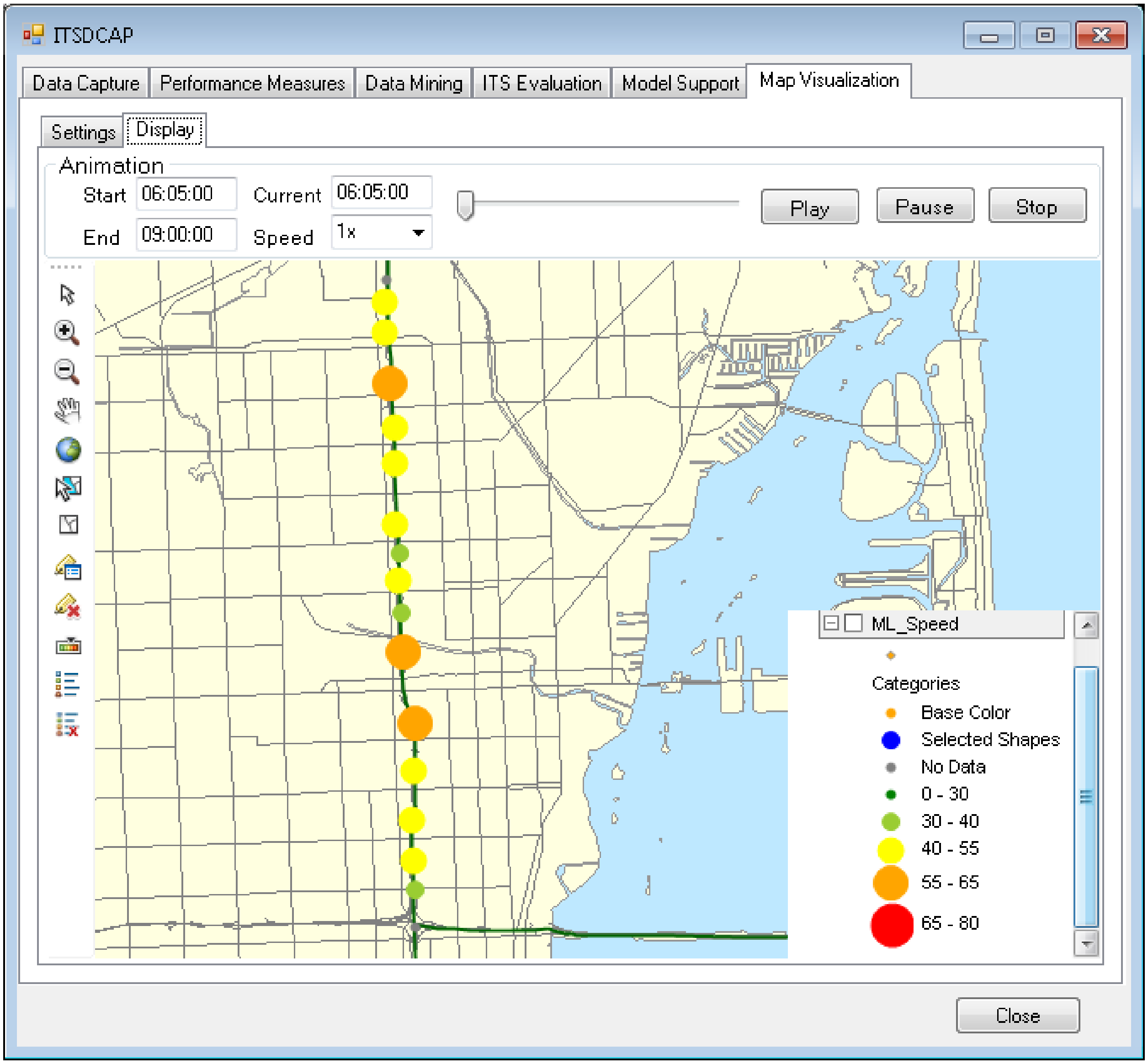


Figure 6-2 Display Interface for Map Visualization Case Study 1

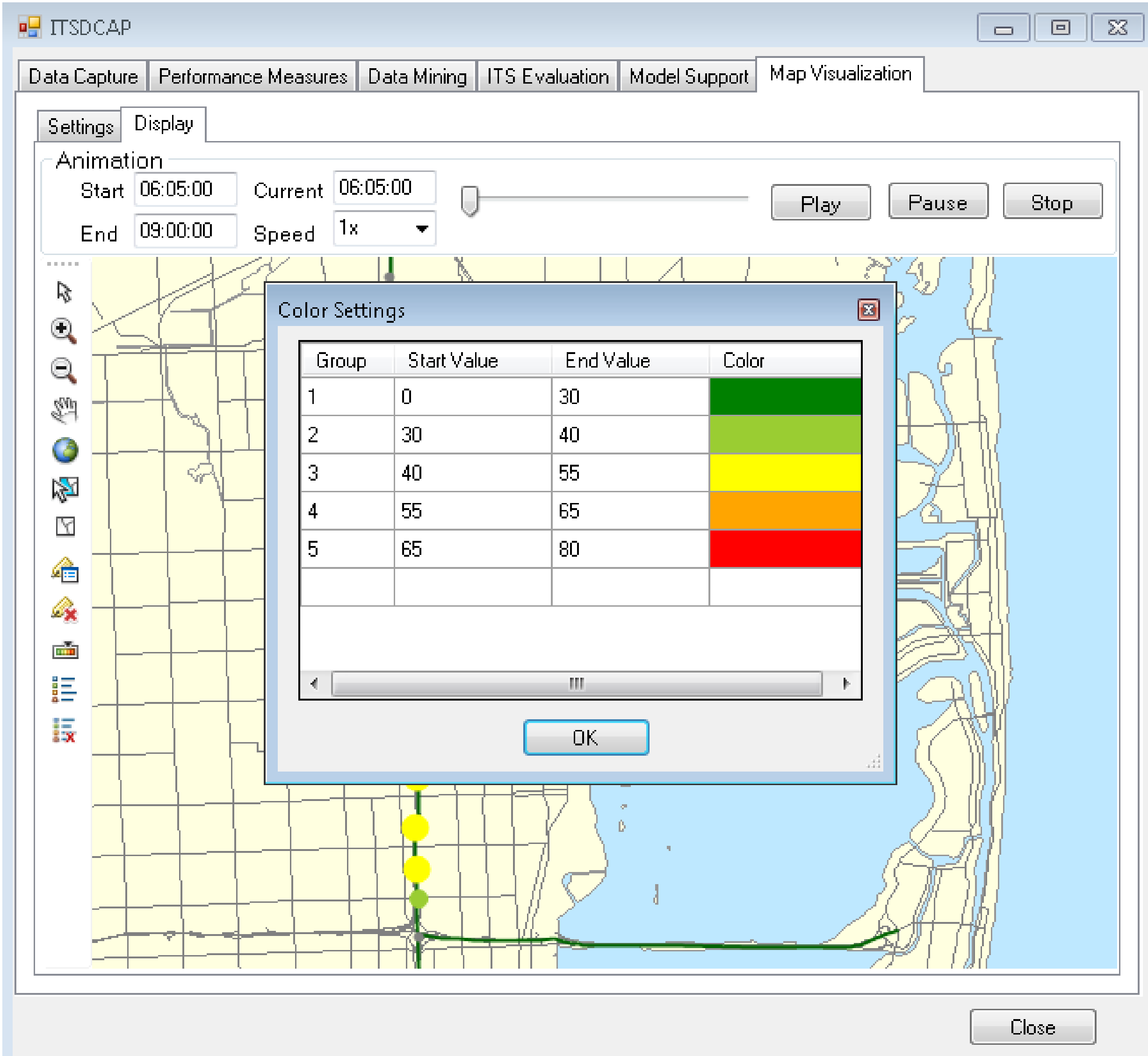


Figure 6-3 Color Settings for Map Visualization Case Study 1

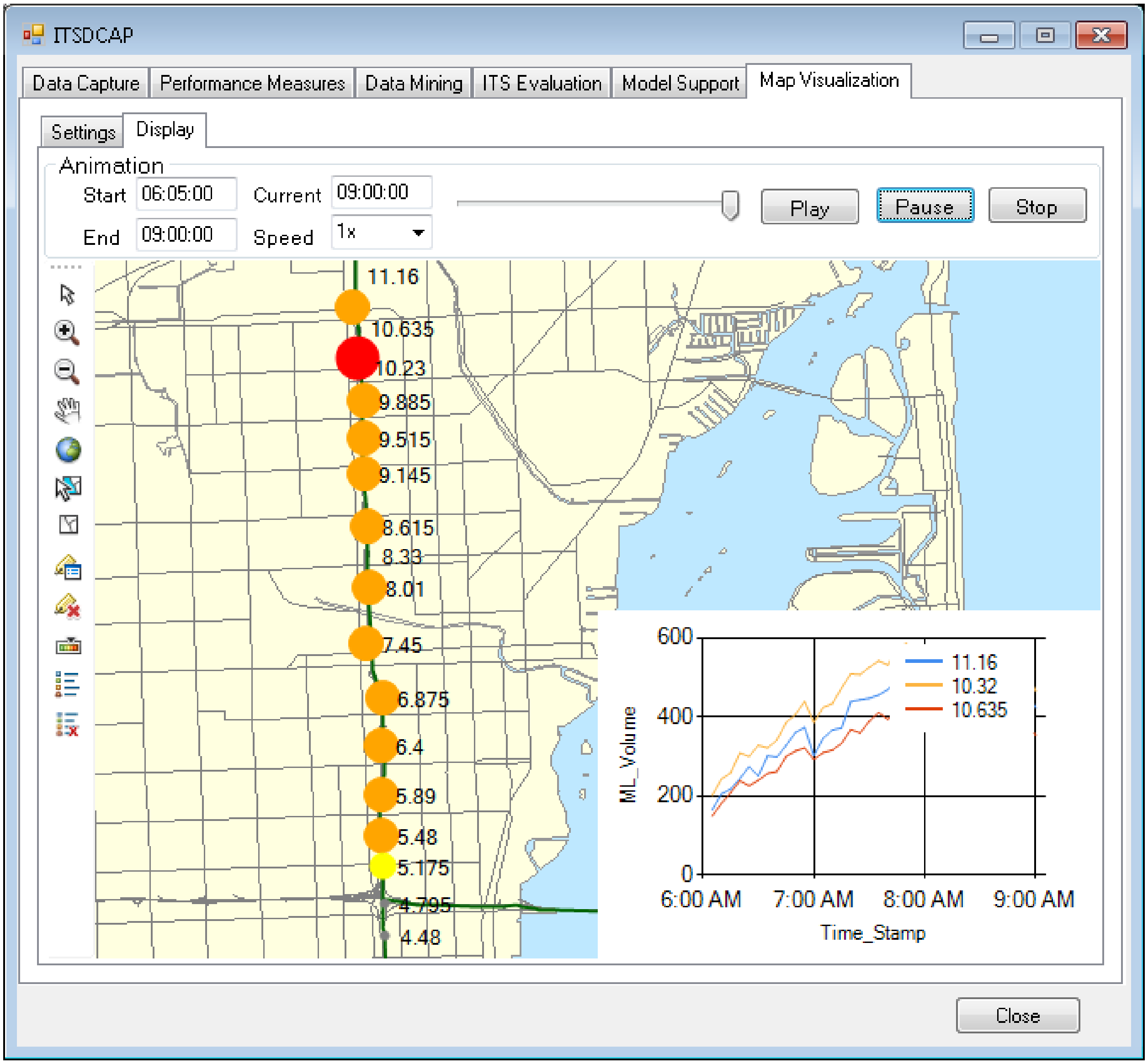


Figure 6-4 Chart Display for Map Visualization Case Study 1

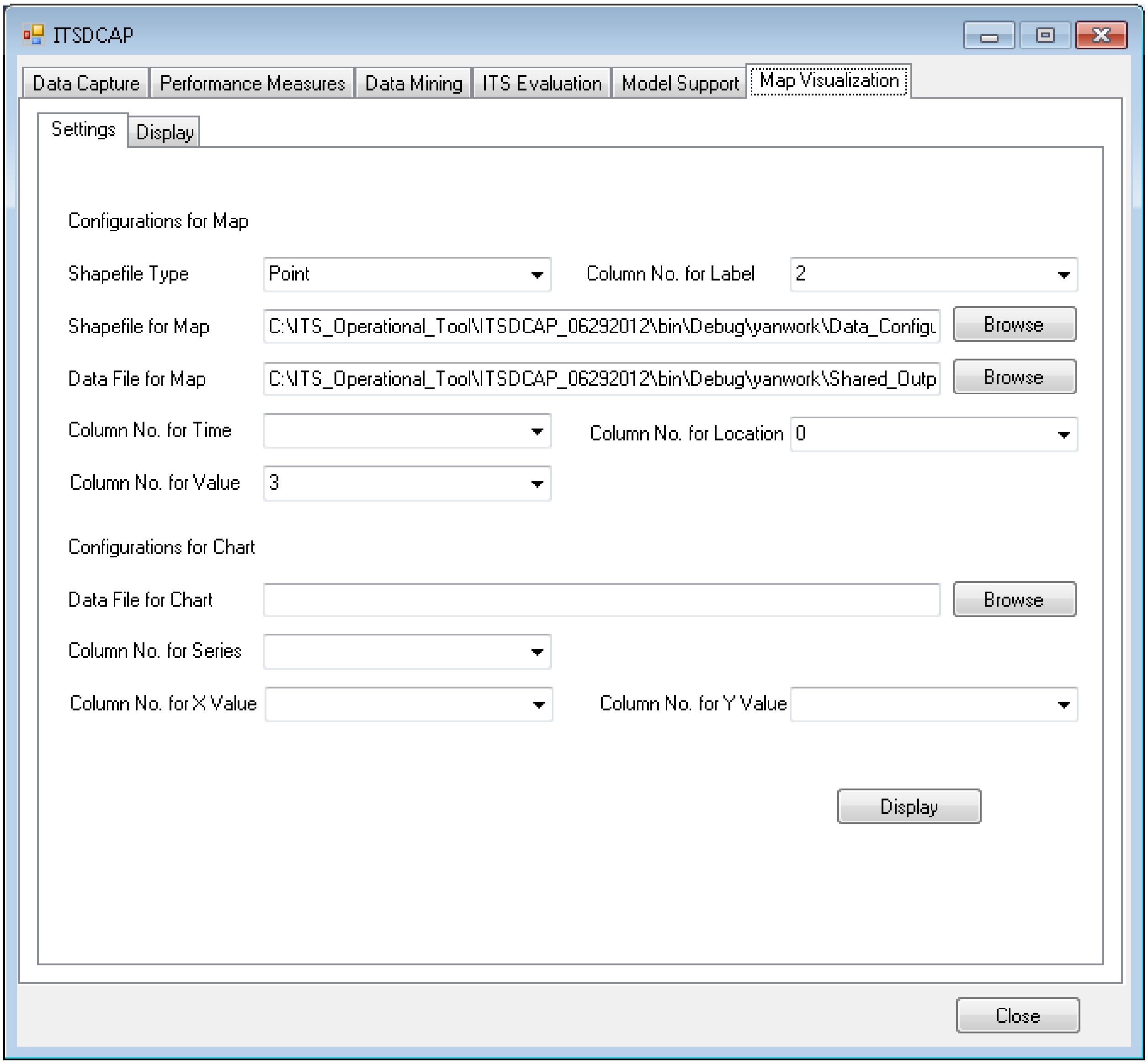


Figure 6-5 Input Interface for Map Visualization Case Study 2

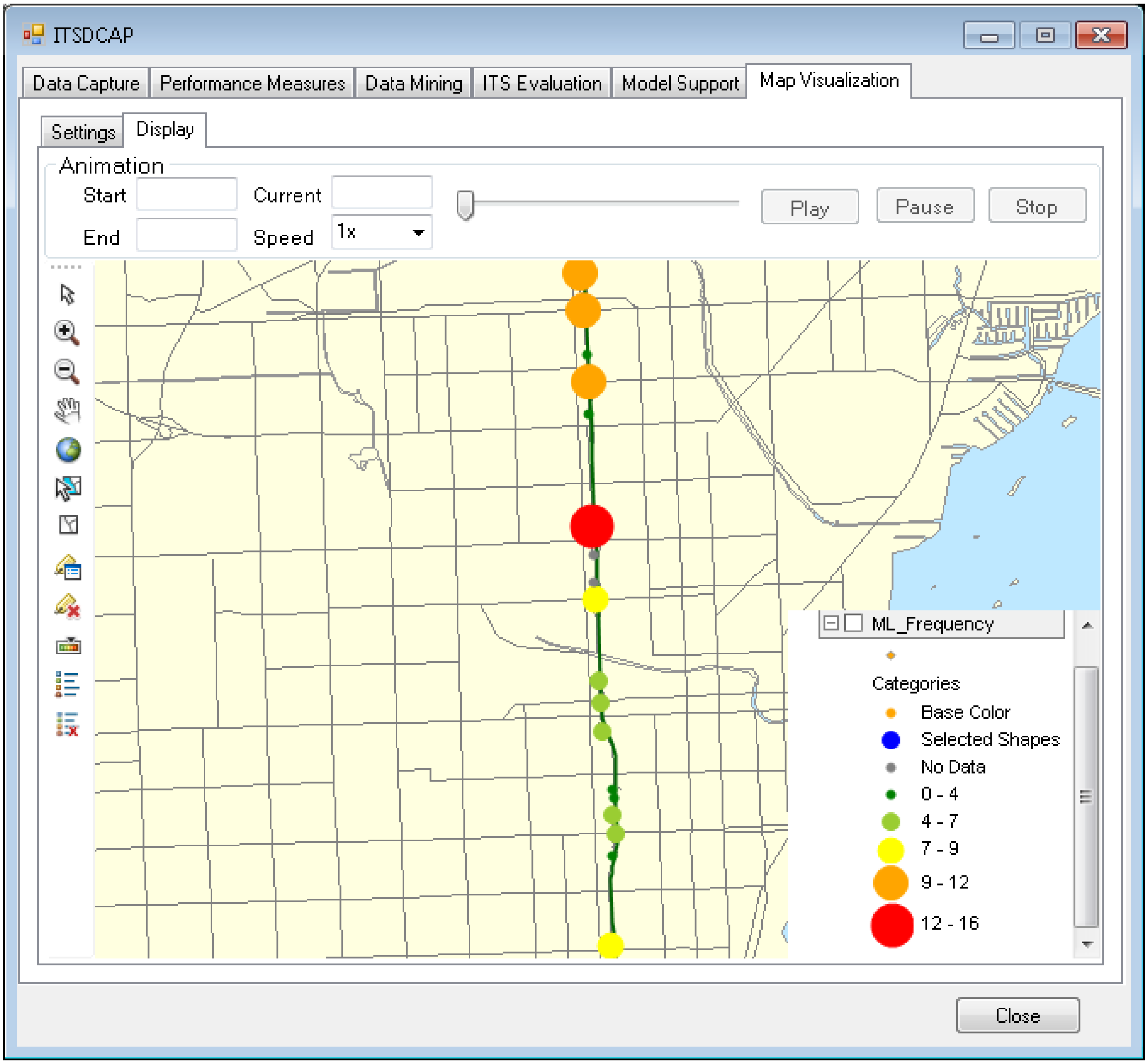


Figure 6-6 Display Interface for Map Visualization Case Study 2

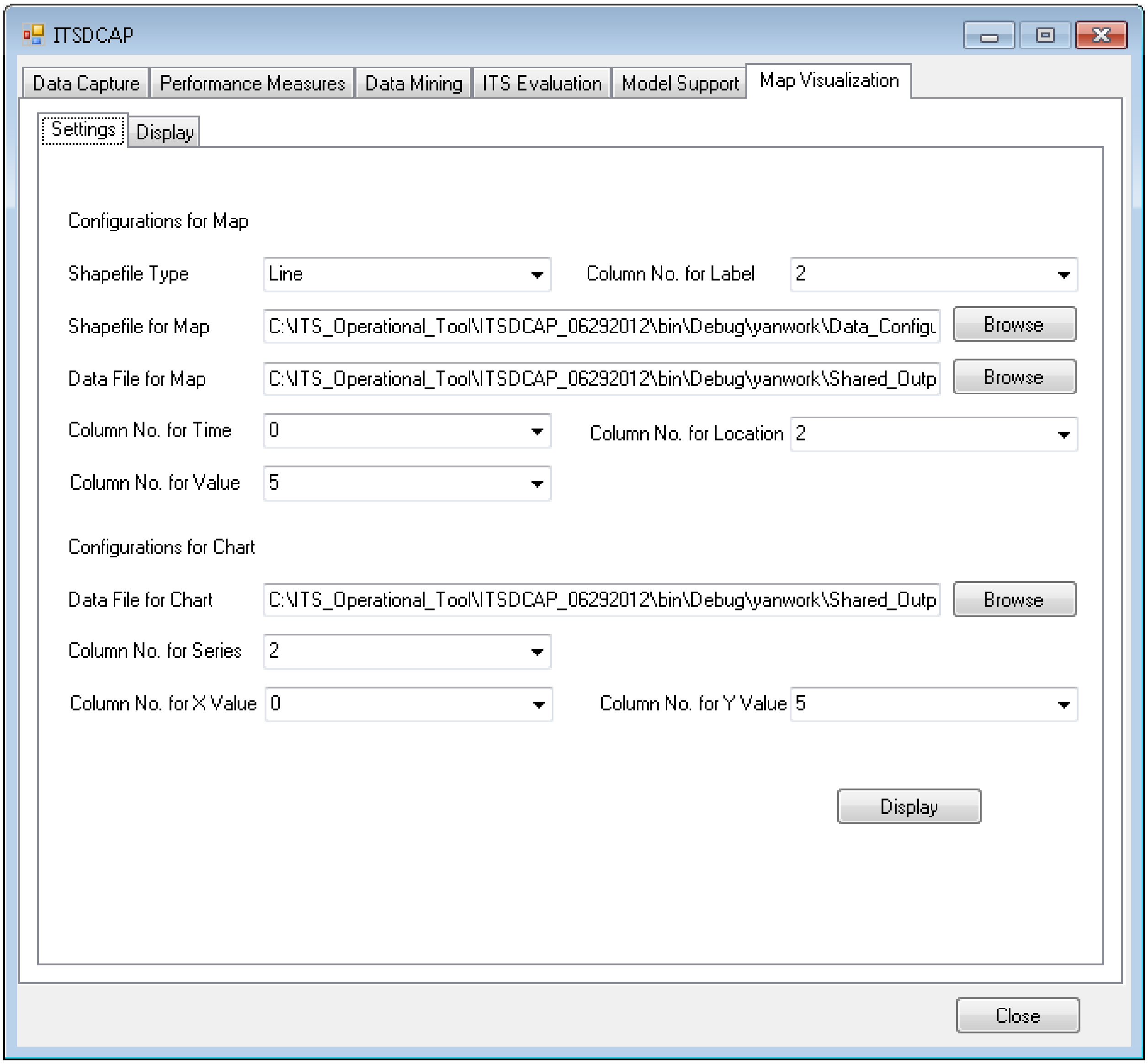


Figure 6-7 Input Interface for Map Visualization Case Study 3

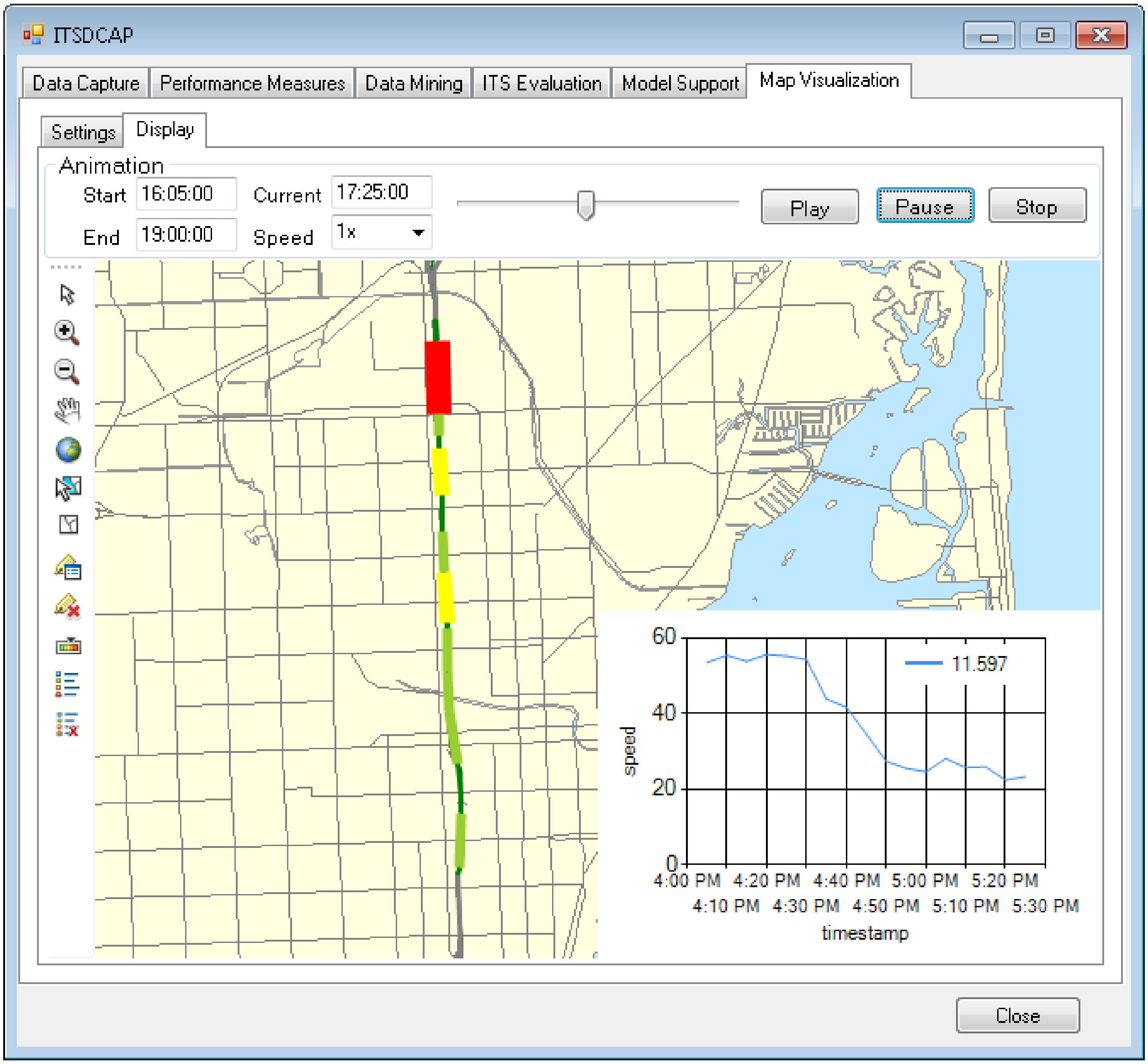


Figure 6-8 Display Interface for Map Visualization Case Study 3

7. ITS Evaluation

Evaluating the benefits of ITS implementation is necessary for both planning and operational purposes. As mentioned in the introduction, a Florida ITS evaluation tool, FITSEVAL, has been developed for use at the planning level. With the availability of rich ITS data and wide implementations of ITS, it becomes feasible to evaluate the impacts of ITS based on real-world data. Furthermore, impact factors based on ITS data can also be derived and applied to enhance the FITSEVAL evaluation accuracy. To achieve this, an ITS Evaluation module was developed and incorporated in the ITSDCAP tool. Procedures are developed in this project to evaluate five types of freeway-related ITS implementations based on ITS data, including incident management, ramp metering, managed lanes, smart work zone, and road weather information system.

7.1. Incident Management

Incident Management (IM) programs aim at reducing incident durations and minimize the negative impacts of incidents by facilitating efficient incident detection, response, and clearance. Figure 7-1 presents the ITSDCAP interface for incident management evaluation. As shown in this figure, the users can choose to evaluate incident management under the option "ITS Type." Other related inputs include the selection of the data source, study corridor, study time period, and types of impacts to be considered for before and after study. The impacts that can be evaluated using ITSDCAP, as shown in Figure 7-1, include the following:

- Incident statistics in terms of incident duration and frequency
- Incident rate
- Demand, queue length, and secondary incident probability for individual incidents
- Incident delay, safety, fuel consumption and emissions for benefit/cost analysis.

The evaluation of these impacts is described below in detail.

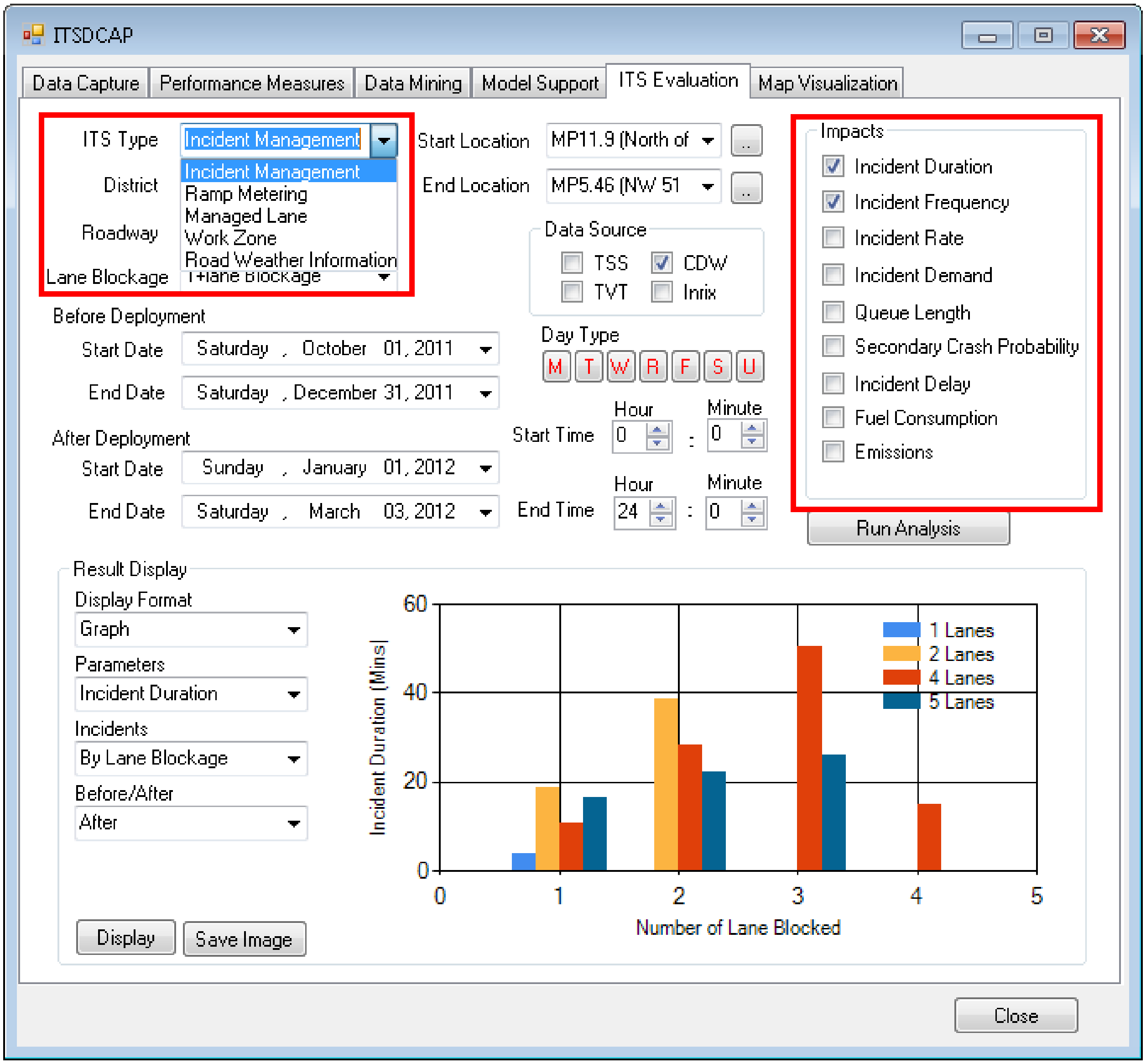


Figure 7-1 Incident Management Evaluation Interface

7.1.1. Incident Statistics

Incident duration and incident frequency is essential input for planning-level ITS evaluation such as FITSEVAL. Such information is also useful for TMC operations to adjust their operations according to these statistics. In ITSDCAP, the incident frequency and average incident duration are summarized by time, location, and the number of blocked lanes, giving the user a picture of the temporal and spatial distribution of incidents.

7.1.2. Incident Rate

Incident rate is defined as the number of incidents per million vehicle-mile traveled (MVMT) by the lane blockage type. This is also an input that is required for the FITSEVAL tool that includes default values calculated based on FDOT District 4 data. In ITSDCAP, in order to calculate the MVMT, the total vehicle-mile traveled during the study period for the selected corridor is calculated based on the normal day traffic volumes for each period of the analysis, estimated using the procedures described in previous sections. The incident rate estimates by the type of lane blockages are outputted by ITSDCAP tool for each period of the analysis.

7.1.3. Demand, Mobility Impacts, and Secondary Incident Probability for Individual Incidents

Demands during incidents, queue lengths, and associated secondary incident probabilities are important factors that need to be evaluated for incident management assessment. Since the volume counts upstream of the incidents are not actual demand due to capacity constraints of incidents, in this project, the historical normal day volume count at the incident location is used to estimate the demands during the incidents. The procedures for estimating historical demand are the same as those utilized for other modules of ITSDCAP, as described earlier.

As described in Mobility Measure Estimation Module, travel time and queue length can be estimated based on the detector data using three different methods; they are, speed threshold-based method, occupancy threshold-based method, and clustering analysis-based method. In this ITS Evaluation module, the users can apply any one of these methods to estimate the queue length during the incident depending on their preference. The maximum queue length associated with the incident is reported in the output. INRIX data can also be used to estimate travel time. Under certain conditions, it may be desired to calculate the delays based on queuing analysis based on demands, capacity drop, and incident duration rather than measuring the impacts on travel time based on detector data. For example, this option may produce better results if the detectors are not able to capture the full queues due to incidents. Thus, the option of calculating delays based on queuing analysis is given to the user.

An enhanced logistic regression model, developed in a previous effort by the research team (Zhan et al., 2009), was applied in this project to assess the potential for secondary crashes. This model was developed based on the FDOT District 4 incident database, and relates the probability of secondary incidents to factors that were found to have statistically significant influence on secondary incident occurrence; including, time of day, incident location, incident type, lane blockage duration, and queue length. Equation 7-1 shows the derived expression of the logistic regression model for secondary crash likelihood.

$$\begin{aligned}\mathrm{Prob}(\mathit{SecondaryCrash}) = \exp(&-6.100 + 0.462 \times \ln(\mathrm{LaneBlockage}) + 0.170 \times \mathrm{QueueLength} \\ &+ 0.236 \times \mathrm{I95NB} + 0.702 \times \mathrm{PM} + 0.959 \times \mathrm{Midday} \\ &+ 1.397 \times \mathrm{AM} + 0.451 \times \mathrm{Accident})\end{aligned} \tag{7-1}$$

Where, LaneBlockage represents the total length of lane blockage in minutes and QueueLength denotes the maximum queue length in miles caused by the incident. All the other variables in Equation 7-1 are binary variables with a value of 0 or 1. The variable of I95NB indicates whether the incident occurred on I-95 northbound. The variables of AM, Midday, and PM have values of 1 if the incident occurred during the weekday AM peak period, midday period, or PM peak period, respectively. If the incident type is crash, the variable of Accident has a value of 1.

7.1.4. Benefit/Cost Analysis

When conducting the benefit/cost analysis for incident management, four types of performance measures are considered: incident delay, fuel consumption, safety, and emissions. Instead of using queuing analysis as in planning-level ITS evaluation tool (FITSEVAL); in ITSDCAP, the incident delay is calculated based on the incident day's vehicle-hour traveled compared to the normal day's vehicle-hour traveled for those timestamps with incident conditions, including the recovery time period. Note that the delays for those demands that cannot pass the incident location due to the reduced capacity are captured by considering the VHT changes during the incident recovery time period.

For some incidents, the associated records in the incident database only report the timestamp at which the lane blockage ends and there is no information about the timestamp when the traffic returns to normal conditions. In this case, a procedure is developed in this project to identify the incident recovery ending time based on detector data. In this algorithm, starting from the timestamp that the lane blockage ends, the speeds of neighboring detectors around the incident are compared to the normal day values. When the difference of these speeds from their normal day values are consistently less than certain thresholds, the timestamp is considered as the ending time of the incident recovery. Below is the mathematical expression for this criterion of comparison,

$$\begin{cases} s_{i,t} - s_{i,n,t} \geq 0 \\ s_{i,t} - s_{i,n,t} < 0 \quad and \quad \left| s_{i,t} - s_{i,n,t} \right| < \varepsilon \end{cases} \tag{7-2}$$

where s symbolizes the speed. The subscript j represents the detector station, t indicates the time interval, n refers to the normal day value. Empirically, the speeds at the three upstream detectors and one downstream detector of the incident location are examined in this project. The ε is the speed threshold, whose default value is selected as 5 mph. To avoid the fluctuations in detector data, this algorithm requires those detector speeds to satisfy the above threshold for at a specified number of time intervals. The default is two intervals considering the detector data aggregation level of 5 minutes. However, the users can modify these default values based on their needs. The benefits of incident management between any two given periods are calculated by summing the delays caused by all incidents in each period and calculating the difference between the before and after period. Another option is given to the user to calculate the delays based on queuing analysis that is desirable when traffic detector locations do not allow capturing the full lengths of queues due to incidents, as described in the previous section.

In addition to the incident delay, fuel consumption and emission impacts of incidents are calculated based on the method used by Skabardonis and Mauch (2005) and Lin et al. (2012). The equation for fuel consumption and pollutant emission calculation is as follows:

$$F_i = D \times e_{si} \tag{7-3}$$

where F_i represents either the fuel consumption or CO, HC, NOx emissions. D is the incident-induced delays and e_{si} is the fuel consumption rate or emission rate at speed s. Compared to the method described in the performance measure calculation section, this method can better capture the fuel consumption and emissions under the stop-and-go conditions caused by incidents.

The above mentioned performance measures are converted to dollars by considering the value of time, safety, fuel costs and emission costs. The resulting benefits are then compared to the costs of implementing incident management to produce the benefit/cost ratio.

7.2. Ramp Metering

Ramp metering uses traffic signals to control the number of vehicles entering the freeways. It aims at reducing the disturbance generated by the on-ramp vehicles at the merge area and ensuring smoother traffic flow on the freeway. Based on the literature review conducted as part of the FITSEVAL project (Hadi et al., 2008), the major impacts of ramp metering include the improvement in freeway throughput and reduction in crashes. Figure 7-2 illustrates the user interface for ramp metering evaluation. As shown in this interface, the users can start the evaluation of ramp metering by selecting the corresponding option under the "ITS Type." Accordingly, the impacts of ramp metering will be listed in the interface. In the ITSDCAP tool, the following impacts of ramp metering can be calculated:

- Average speed, travel time and maximum throughput
- Vehicle-mile traveled and vehicle-hour traveled
- Travel time reliability
- Safety
- Fuel consumption and emissions.

Based on the selected impacts, benefits of ramp metering in dollar values will be estimated and used to determine the benefit/cost ratio.

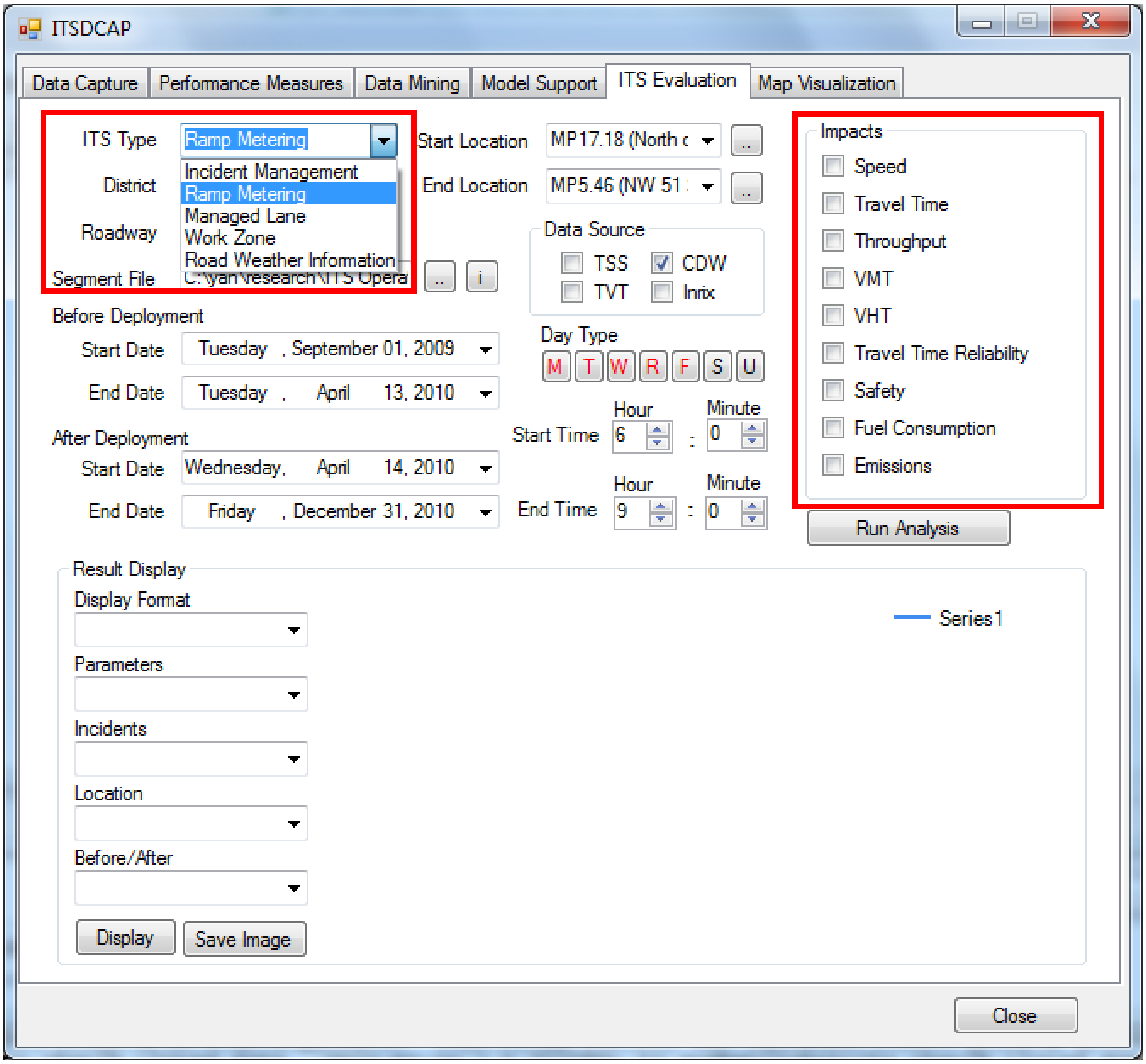


Figure 7-2 Ramp Metering Evaluation Interface

7.2.1. Average Speed, Travel Time, and Maximum Throughput

As mentioned above, the implementation of ramp metering can impact freeway speed and throughput. In ITSDCAP, users have the option of using either raw TSS data, STEWARD data, or Inrix Data to assess the average speed and travel time for before and after the implementation. The travel time can also be estimated based on TVT data depending on user selection. Since the Inrix data does not have volume counts, the maximum throughput has to be calculated based on the detector data within the specified study time period. Note that the changes in travel time on the metered ramps are not measured in the analysis, due to the lack of data. However, queuing analysis can be performed to estimate the delays at these ramps--if the demands on the ramps are available.

7.2.2. Vehicle-Mile Traveled (VMT) and Vehicle-Hour Traveled (VHT)

Based on user specified time periods for ramp metering before and after study; the total VMT and VHT for the two periods are calculated by multiplying the volume counts by the corresponding distance and travel time. The difference in VHT is the time savings due to the implementation of ramp metering, which can be further converted to dollar values by multiplying by the value of time.

7.2.3. Travel Time Reliability

According to the analysis presented Section 4.2, the 85th percentile travel time index is a better metric for quantifying travel time reliability. Therefore, it is used in this module to assess the impacts of ramp metering on travel time reliability. The 85th percentile travel time is then multiplied by the normal day volumes and value of reliability (VOR) to quantify the benefits of travel time reliability in dollar values. At this stage, the VOR is not clear. A recent advanced traffic and demand management analysis effort being conducted for the Federal Highway Administration assumed VOR values close to those of value of time (VOT). Thus, VOR is assumed to equal VOT in this study.

7.2.4. Safety

To evaluate the safety benefits of ramp metering, the number of fatalities, injuries, and property-damage-only (PDO) for affected freeways and metered ramps are obtained from the crash database for the before and after period. These numbers can be used externally at this stage in combination with procedures of the Highway Safety Manual (HSM) such as the Empirical Bayes before and after safety effectiveness evaluation method (AASHTO, 2010) to estimate the difference in crashes between the before and after conditions. The differences in crashes are multiplied by the corresponding crash costs to obtain the monetary benefits.

7.2.5. Fuel Consumption and Emissions

In addition to the above-mentioned impacts, the impacts of ramp metering on fuel consumption and emissions are also considered in this study. The fuel consumption and emissions for before and after time periods are calculated based on fuel consumption rate/emission rate and corresponding vehicle-mile traveled. These performance measures are transformed to monetary benefits by multiplying the fuel costs and pollutant costs.

7.3. Managed Lanes

Managed lanes aim at maximizing freeway capacity by implementing proactive operational strategies and managing them in response to changing traffic conditions (FHWA, 2012). Managed lanes may be a high-occupancy vehicle (HOV) lane, or an exclusive-use lane for use by bus or truck lanes, or a high-occupancy toll (HOT) lane allowing high-occuancy vehicle use for free while other vehicles access this facility by paying tolls (Kuhn et al., 2005). In advanced applications, tolls for HOT lanes are adjusted based on dynamic traffic conditions.

Various models have been reported in literature to assess managed lane operations, including nested logit models (McDonald and Noland, 2001; DeCorla-Souza, 2003), traffic simulation models, dynamic assignment models (He et al., 2000), optimization models (Li and Govind, 2003), equilibrium assignment-based travel demand forecasting models (Brunk and Middleton, 1999; Rodier and Johnston, 2002), or combination of these models (Murray et al., 2001). In ITSDCAP, however, the evaluation of managed lanes is directly based on real-world data. As shown in Figure 7-3, the impacts of managed lanes considered in this tool include:

- Average speed, travel time and maximum throughput
- Vehicle-mile traveled and vehicle-hour traveled
- Travel time reliability
- Safety
- Fuel consumption and emissions
- Toll Revenue.

For most of these impacts, the evaluation procedures are very similar to those implemented for ramp metering, and therefore, the description of these procedures is omitted here for brevity. The only impact that is different from ramp metering is the toll revenue. In ITSDCAP, such impact is calculated by multiplying the number of users by the toll rate. Again, the benefits converted to dollar values are divided by the total costs for managed lanes to produce the benefit/cost ratio.

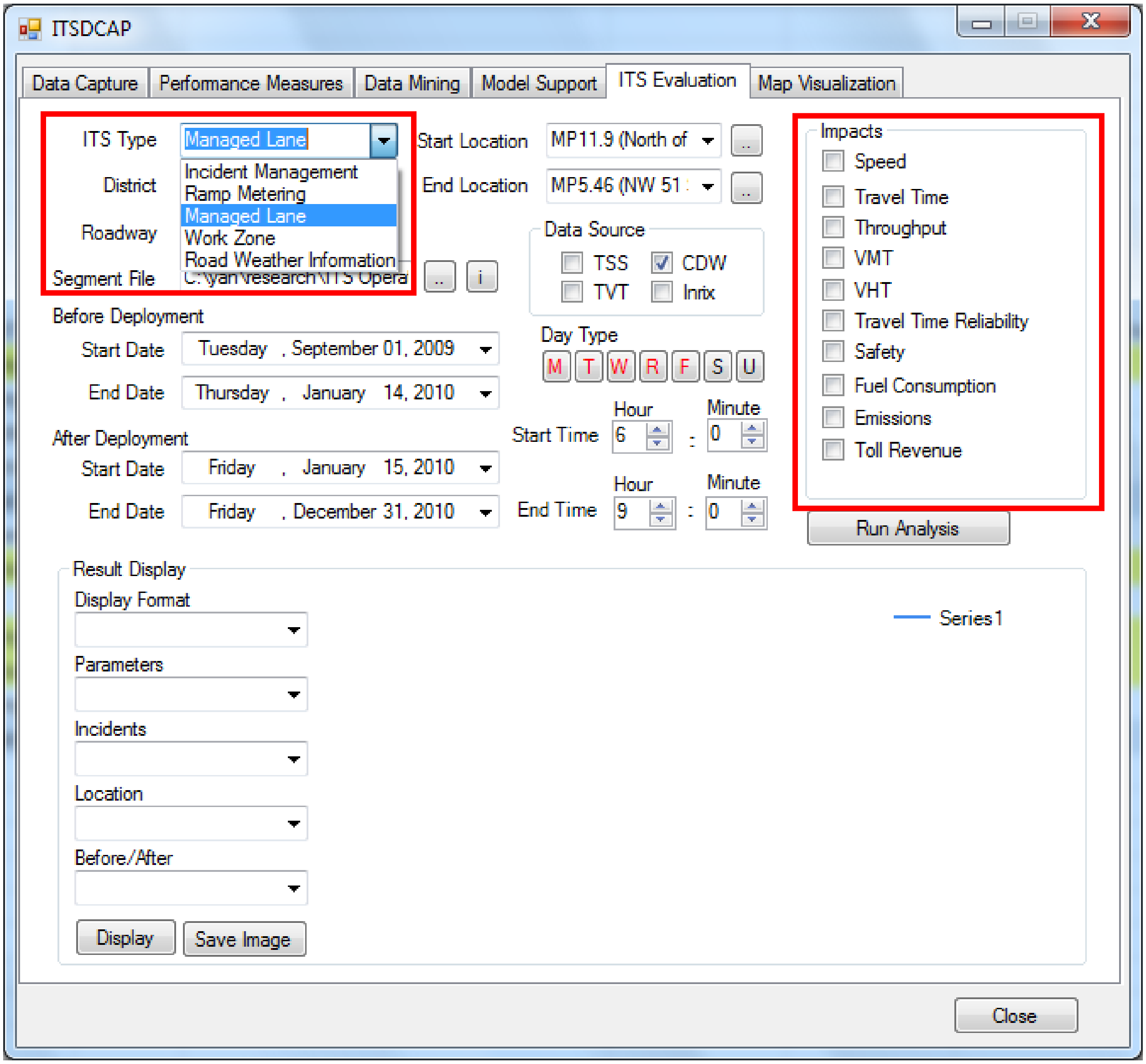


Figure 7-3 Managed Lanes Evaluation Interface

7.4. Smart Work Zone

Smart work zone is an automated system that aims at improving safety and mobility for motorists by providing them with real-time information on work zone traffic conditions. Figure 7-4 illustrates the interface for smart work zone evaluations. As shown in this figure, the evaluation of smart work zone mainly focuses on the safety impacts in this version of ITSDCAP. This will be extended to include mobility benefits in future versions. A list of work zone information has to be provided by users, including the information on the highway segment, direction, area type for determining crash costs, roadway segment number, starting milepost, ending milepost, and construction starting and ending date and time. Figure 7-5 provides an example of the input. Once the user specifies the roadway and study period through the work zone evaluation interface, the program automatically finds the work zone information within this time period based on the user-input list of work zone information under the input folder for work zone evaluation. One of two methods can be used in ITSDCAP to assess the benefits:

- The associate numbers of fatalities, injuries, and PDO crashes as well as the crash rates for both the before and after time periods are calculated in ITSDCAP. Then, the changes in crash number for these two time periods are further converted to dollar values and used to determine the benefit/cost ratio.
- When it is not possible to use before and after analysis, a methodology similar to that implemented in FITSEVAL (Hadi et al., 2008) is used. In this case, crash reduction factors are used to produce the number of crashes with smart work zone implementations, allowing the calculation of the benefits.

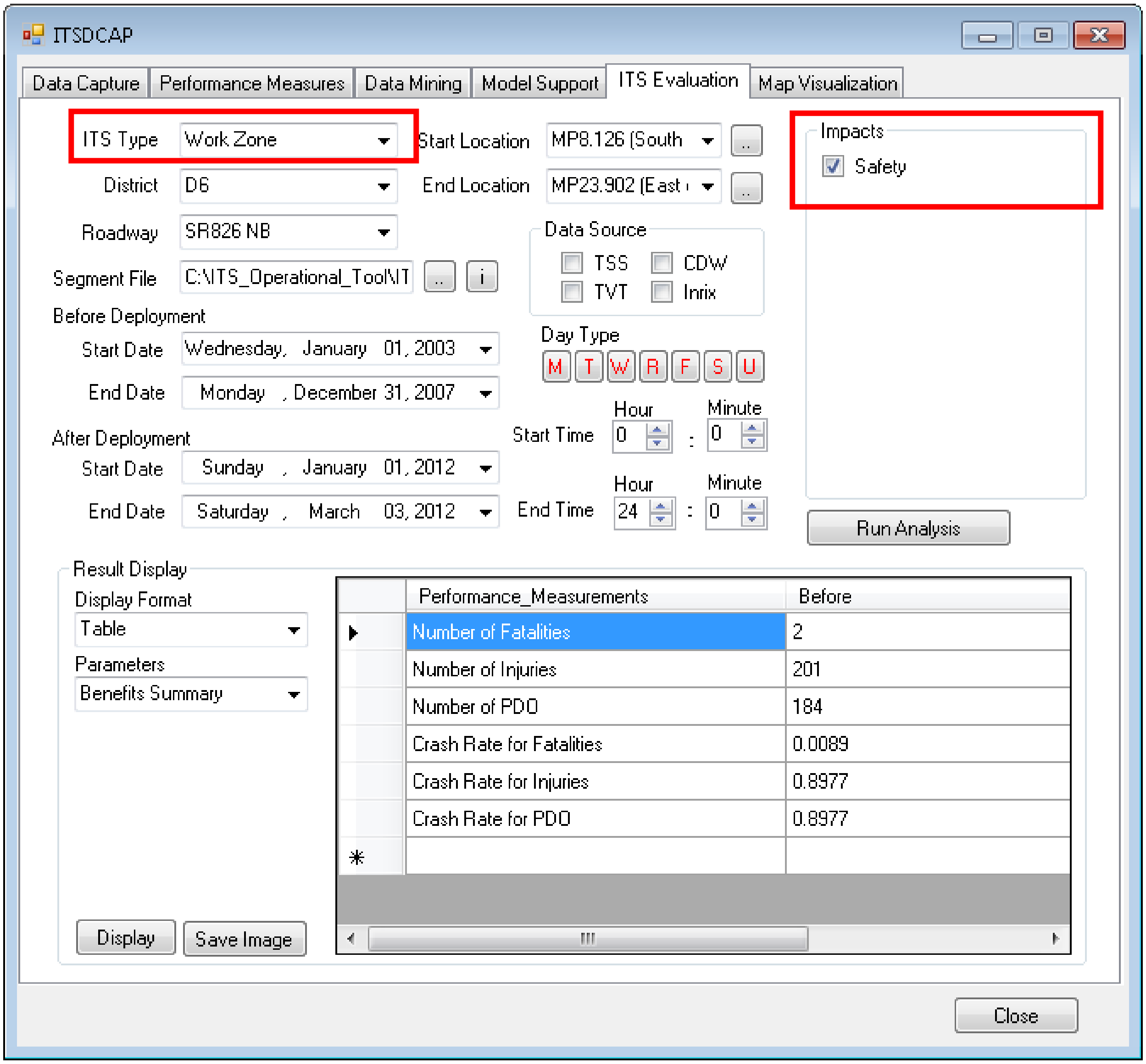


Figure 7-4 Smart Work Zone Evaluation Interface

	A	B	C	D	E	F	G	H	I	J	K
1	Index	Roadway_Name	Directon	Area_Type	Roadway_SegmentNo	Starting_Milepost	Ending_Milepost	Starting_Date	Ending_Date	Starting_Time	Ending_Time
2	1	SR826	NB	Urban Interstate	87260000	11.468	12.998	3/24/2003	7/11/2007	0:00:00	24:00:00
3											

Figure 7-5 Example of Work Zone Input Information

7.5. Road Weather Information System

Road weather information system (RWIS) monitors the pavement and weather conditions. RWIS then processes and disseminates the weather information to the travelers, road operators and maintenance staffs, aiding them in making informed decisions. ITSDCAP focuses on the safety impacts of road weather formation system, as shown in Figure 7-6. The crash numbers and crash rates by severity types (that is, fatality, injury, and PDO) during the rainy conditions within the study periods of time are estimated based on the crash database. One of two methods can be used in ITSDCAP to assess the benefits:

- The associated numbers of fatalities, injuries, and PDO crashes as well as the crash rates for both the before and after time periods are calculated in ITSDCAP. Then, the changes in crash number for these two time periods are further converted to dollar values and used to determine the benefit/cost ratio.
- When it is not possible to use before and after analysis, a methodology similar to that implemented in FITSEVAL (Hadi et al. 2008) is used. In this case, crash reduction factors are used to produce the number of crashes with RWIS implementations, allowing the calculation of the benefits.

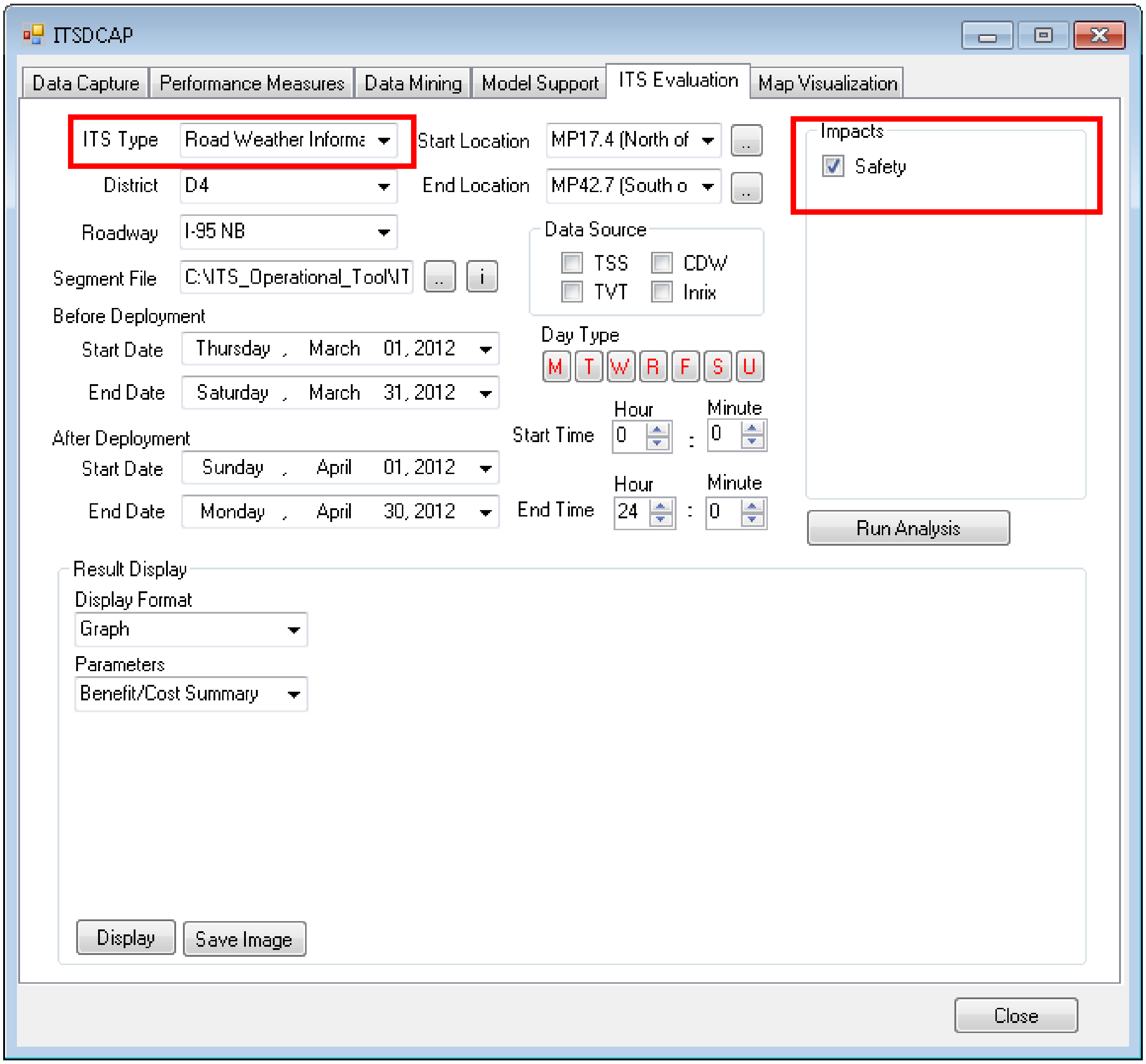


Figure 7-6 Example of Work Zone Input Information

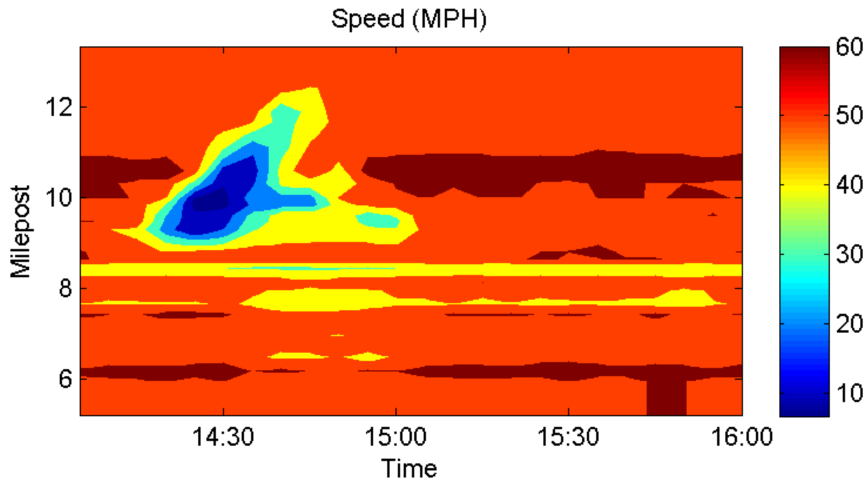
7.6. Case Studies

As examples, this section presents several case studies to illustrate the functions of ITS Evaluation module in ITSDCAP.

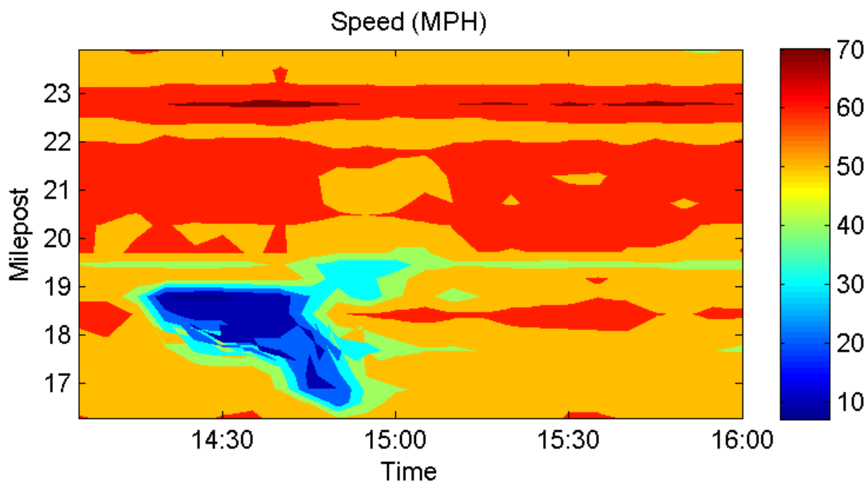
7.6.1. Determine the Incident Recovery Ending Time

Two case studies are presented in this report to illustrate the function of determining incident recovery ending times. In case study 1, a one-lane out of four-lane blockage incident is considered. This incident occurs at NW 103rd St. along I- 95 SB Corridor at the timestamp 2:18 PM on September 6, 2011. In the incident database, there is no record that indicates when the traffic returned to normal conditions. Based on the developed procedure, the estimated ending time of incident recovery is 3:05 PM, which can be further confirmed through the speed contour plot presented in Figure 7-7(a).

Case study 2 involves a two out of three-lane blockage incident that occurred at NW 57th Ave. along SR-826 eastbound. The incident was detected on January 24, 2012, at 2:13 PM. The estimated timestamp that traffic recovered to normal conditions in this study is 3:20 PM. Again, this can also be verified based on speed contour in Figure 7-7(b).



(a)



(b)

Figure 7-7 Example of Incident Recovery Ending Time

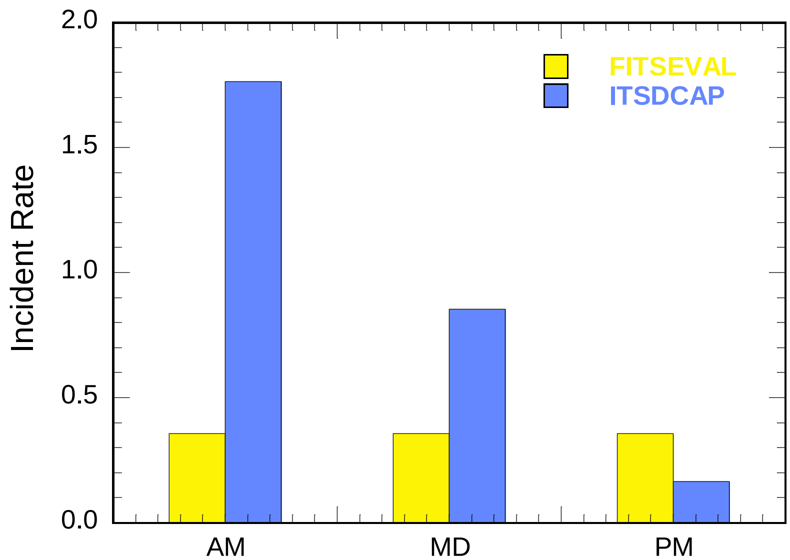
7.6.2. Comparison between FITSEVAL and ITSDCAP for Incident Management

This case study aims at comparing the evaluation of incident management (IM) using the planning-level evaluation tool, FITSEVAL, versus the data-based evaluation tool, ITSDCAP. Several differences in the evaluation methodologies of the two tools need to be highlighted, as listed below:

- In ITSDCAP, the demand level is based on historical normal day traffic count rather than utilizing demand forecasting models as in FITSEVAL.
- Incident rates and durations are obtained from the local incident database in ITSDCAP while FITSEVAL utilizes default values for these parameters based on specific corridors managed by FDOT District 4 in Broward County but allows the users to change the defaults based on local conditions.
- While queuing analysis is the only method used to assess the mobility impacts in FITSEVAL, incident mobility benefits are calculated in ITSDCAP based on the VHT difference between incident days and normal days, although the utilization of queuing analysis is also possible in ITSDCAP.
- Emission and fuel consumption are calculated using the same methods in ITSDCAP and FITSEVAL. However, the speed within the queue is calculated based on detector data in ITSDCAP but using a traffic model in FITSEVAL.

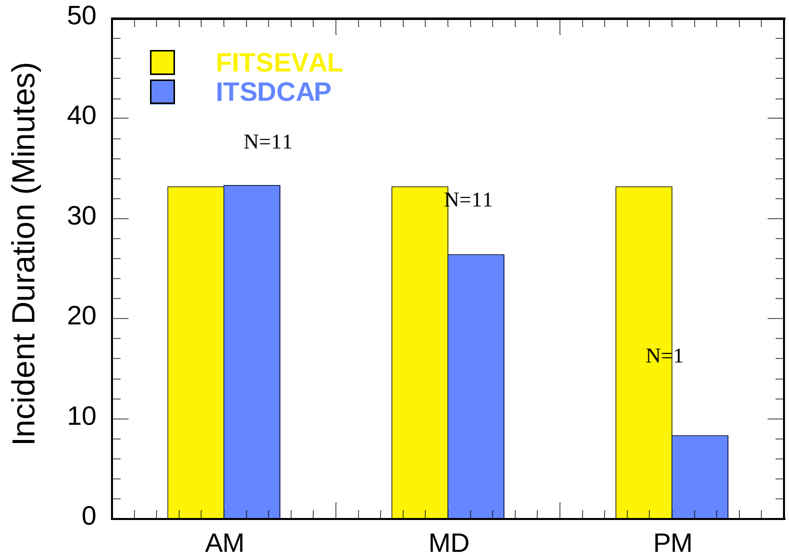
In this case study, SR-826 eastbound in Miami-Dade County is selected as the study corridor to determine the impacts of one-lane incidents during the time period from Jan. 1, 2012 to Mar. 31, 2012. Figure 7-8 illustrates the location of the study corridor and Figure 7-9 presents the estimated parameters from the ITSDCAP compared to the default values used in the FITSEVAL. As shown in Figure 7-9, the actual parameters for the study period as estimated using ITSDCAP is quite different from the default values in FITSEVAL. Note that the number of one-lane blockage incidents is very limited in the study period for the PM peak, which may result in unreliable results for this period.

Figure 7-8 Study Corridor Utilized in Incident Mobility Impacts Assessments

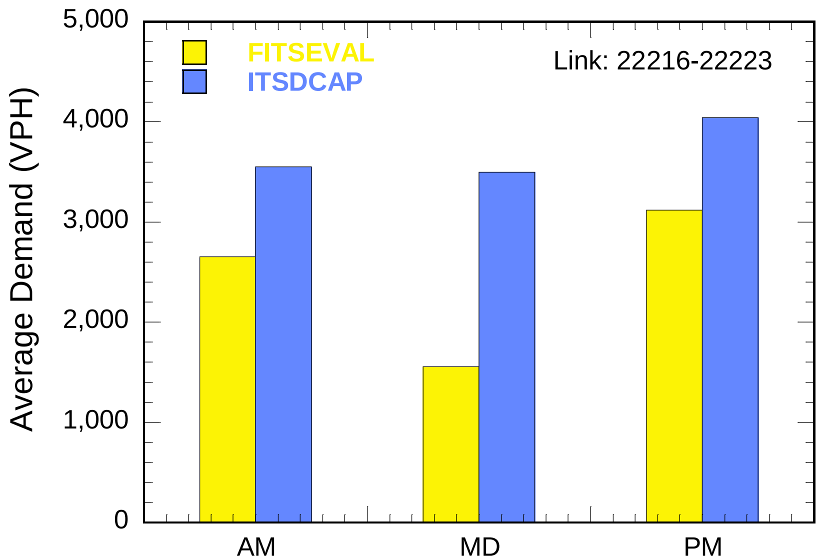


(a)

Figure 7-9 Estimated Parameters from ITSDCAP (Continued on Next Page)



(b) N: Sample Size



(c)

Figure 7-9 Estimated Parameters from ITSDCAP

Four case studies were included in the comparison, which are:

- FITSEVAL-Default: using the default values in FITSEVAL
- FITSEVAL-Modified 1: FITSEVAL evaluation using incident duration and incident rate estimated using ITSDCAP but utilizing the demands estimated by FITSEVAL

- FITSEVAL-Modified 2: FITSEVAL evaluation using incident duration, incident rate and demands estimated using ITSDCAP (this method should produce the same results as using queuing analysis in ITSDCAP)
- ITSDCAP: using the real-world data. With the incident mobility impacts are calculations in ITSDCAP based on the VHT

Table 7-1 summarizes the comparison results. As shown in this table, when using the default parameters in FITSEVAL, the performance measures are quite different from those based on real-world data. As the incident duration and incident rates are updated to reflect real-world measurements, the delay, fuel consumption, and emissions estimates based on queuing analysis in FITSEVAL become closer to those based on ITSDCAP measurements, but there are still significant differences between the estimates. When the demand is further updated, the incident management evaluation results become close to the real-world values. Figure 7-10 presents the estimates of the delay monetary values. Again, this plot also indicates that the update of assessment parameters based on real-world conditions in FITSEVAL can bring the evaluation results close to real-world measurements of mobility impacts.

Table 7-1 Comparison of IM Evaluation Results between FITSEVAL and ITSDCAP

		FITSEVAL-Default	**FITSEVAL-Modified 1**	**FITSEVAL-Modified 2**	**ITSDCAP**
Incident Delay (VHT)	AM	146	728	3249	3086
	MD	0	0	936	915
	PM	231	7	15	5
	Total	377	735	4200	4006
Gas Consumption (Gallons)	AM	211	1053	4447	4596
	MD	0	0	1289	1736
	PM	329	10	21	10
	Total	540	1063	5757	6342
CO Emission (Tons)	AM	0.04	0.21	0.62	1.14
	MD	0	0	0.17	0.40
	PM	0.06	0.001	0.003	0.002
	Total	0.10	0.211	0.793	1.542

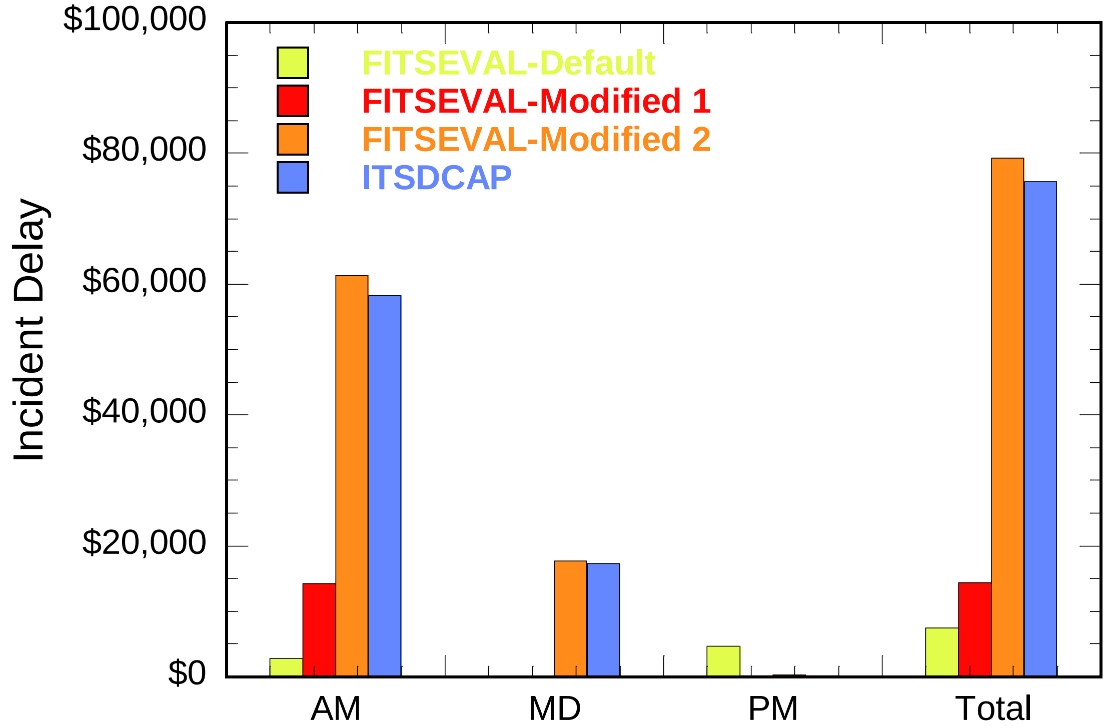


Figure 7-10 Comparison of Incident Delay in Dollar Values

7.6.3. Ramp Metering and Express Lane Evaluation

In this case study, the impacts of ramp metering and express lane implementation along I-95 SB Corridor in Miami-Dade County, FL, will be examined. The activation date for I-95 SB Express Lane was January 15, 2010. The ramp metering activation date was Apr. 14, 2010, and ramp signals are only activated during the morning peak period between 6:00 AM and 9:00 AM. Figure 7-11 shows the map of study corridor. The impacts of these two ITS implementations were evaluated in terms of throughput, safety, and travel time reliability.

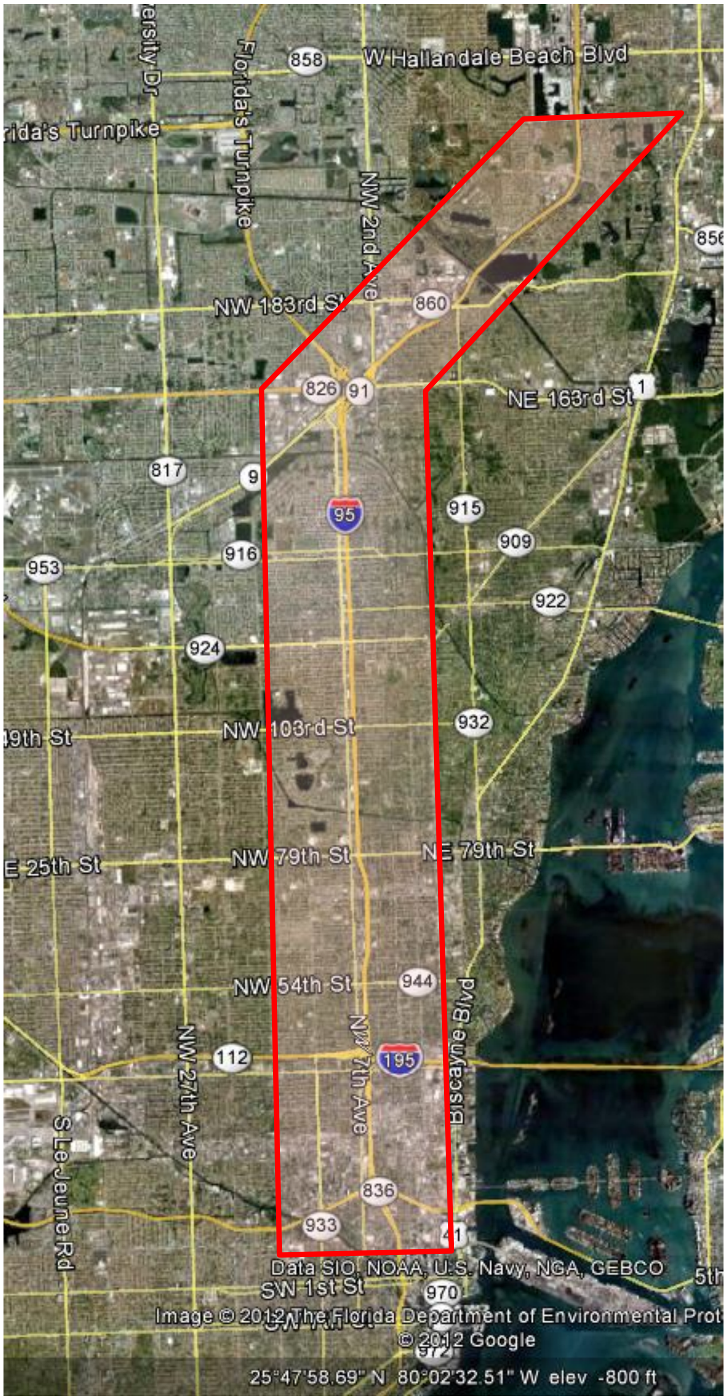


Figure 7-11 Map of Study Corridor for RM and HOT Evaluation

Figure 7-12 presents the results of maximum throughput during the AM peak period for roadway section between Golden Glades Interchange and SR 836. Note that this section has both express lane and ramp metering during the morning peak period. There are three curves within this figure: one corresponds to the results based on the data on September 2009, which is before the activation of express lane; the second one shows the throughput results according to the detector data in March 2010, after the activation of express lane; and the third curve represents the results for time period of May 2010, with the activation of both ITS components. This figure indicates that the implementation of ITS improves the roadway throughput.

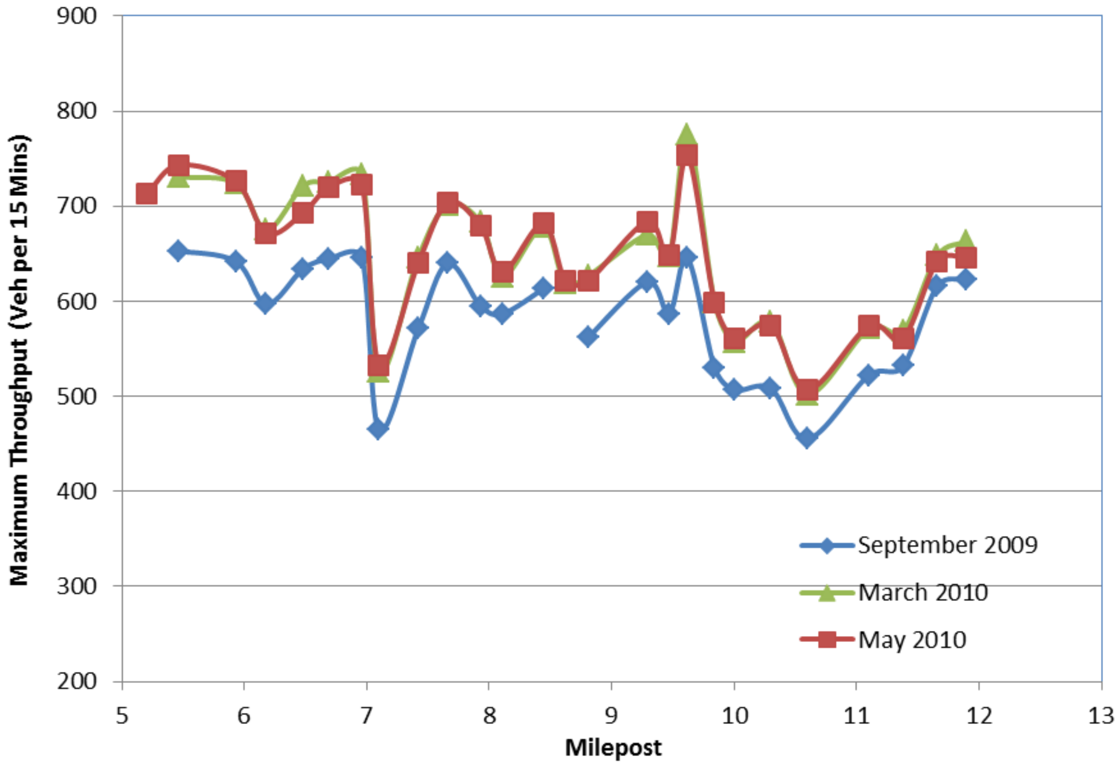


Figure 7-12 Maximum Throughput Results for I95SB between Golden Glades Interchange and SR 836

7.7. References

American Association of State Highway and Transportation Officials (AASHTO). Highway Safety Manual, Washington, D.C., 2010.

Brunk, J.C. and M.D. Middleton, Incorporating Toll-Lane and High-Occupancy Toll-Lane Effects into the Dallas System Planning Methodology. In *Transportation Research Record: Journal of the transportation Research Board*, No.1659, TRB, National Research Council, Washington, D.C., 1999, pp. 105-110.

DeCorla-Souza, P. Evaluation of Toll Options Using Quick-Response Analysis Tools: A Case Study of the Capital Beltway. CD-ROM, Transportation Research Board, National Research Council, Washington, D.C., 2003.

Federal Highway Administration Freeway Management Program (FHWA). *Managed Lanes*. http://ops.fhwa.dot.gov/freewaymgmt/managed_lanes.htm. Accessed June 6, 2012.

Hadi, M., Y. Xiao, H. Ozen, and P. Alvarez. *Evaluation Tools to Support ITS Planning Process: Development of a Sketch Planning Tool in FSUTMS/Cube Environment*. Final Report, Lehman Center for Transportation Research, Florida International University, October 2008.

He, R.R., B. Ran and K. Choi. Decision Support System for Value Pricing Strategies on I-394 Corridor Network. CD-ROM, Transportation Research Board, National Research Council, Washington, D.C., 2000.

ITS Deployment Analysis System (IDAS) User's Manual, Prepared for the Federal Highway Administration by Cambridge Systematics, Oakland, CA, November 2001.

Kuhn, B., G. Goodin, A. Ballard, M. Brewer, R. Brydia, J. Carson, S. Chrysler, T. Collier, K. Fitzpatrick, D. Jasek, C. Toycen, and G. Ullman. *Managed Lanes Handbook*. FHWA/TX-06/0-4160-24. Texas Transportation Institute, October 2005.

Li, J. and S. Govind. An Optimization Model For Assessing Pricing Strategies of Managed Lanes. CD-ROM, Transportation Research Board, National Research Council, Washington, D.C. 2003.

Lin, P., A. Fabregas, H. Chen, H. Zhou, and Q. Wang. *Review and Update of Road Ranger Cost Benefit Analysis*. Final Report, Center for Urban Transportation Research, University of South Florida, January 2012.

McDonald, N.C., and R.B. Noland. Simulated Travel Impacts of High Occupancy Vehicle Lane Conversion Alternatives. In *Transportation Research Record: Journal of the Transportation Research Board,* No.1765, TRB, National Research Council, Washington D.C. 2001, pp.1-7.

Murray, P.M., H.S. Mahmassani and K.F. Abdelghany. Methodology for Assessing High-Occupancy Toll-Lane Usage and Network Performance. In *Transportation Research Record: Journal of the transportation Research Board*, No.1765, TRB, National Research Council, Washington, D.C., 2001, pp. 8-15.

Rodier. C.J. and R.A. Johnston. A comparison of High Occupancy Vehicle, High Occupancy Toll and Truck Only Lanes in the Sacramento Region. Presented at 80th Annual Meeting of the Transportation Research Board, Washington, D.C., 2002.

Skabardonis A. and M. Mauch. *FSP* [Freeway Service Patrol] *Beat Evaluation and Predictor Models: Users Manual,* Research Report, No. UCB-ITS-RR-2005-XX, Institute of Transportation Studies, University of California-Berkeley, California, 2005.

Zhan, C., A. Gan, and M. Hadi. Identify Secondary Crashes and Their Contributing Factors. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 2102, 2009, pp. 68-75.

8. Data Mining

The data mining module in ITSDCAP explores the relationship between different attributes of the data from one or more data sources. A simple built-in data mining module is included in the current version of ITSDCAP. Additional data mining functions will be added in the future. However, more advanced data mining applications can be conducted utilizing data outputted by ITSDCAP. This chapter discusses first the built-in Data Mining module in ITSDCAP. Then, this chapter presents more advanced data mining methods conducted in this project outside ITSDCAP, based on the data provided by the tool, to illustrate the benefits of combining data from multiple sources in the analysis.

8.1. Data Mining Procedures in ITSDCAP

The current version of the built-in ITSDCAP data mining module allows the association of data from the following sources: crash data, incident data, DMS data, and 511 data. However, as will be discussed later in this chapter, fused data from other sources can be output by ITSDCAP allowing powerful data mining analysis outside ITSDCAP.

In the built-in data mining module, once the user selects each of the four databases mentioned above, the attributes available in the database are listed in the “Database Attributes” list box, allowing the users to further select their desired attributes, as shown in Figure 8-1. The user can select the type of the output (e.g., graph, table, etc.) and run the analysis to produce the desired outputs.

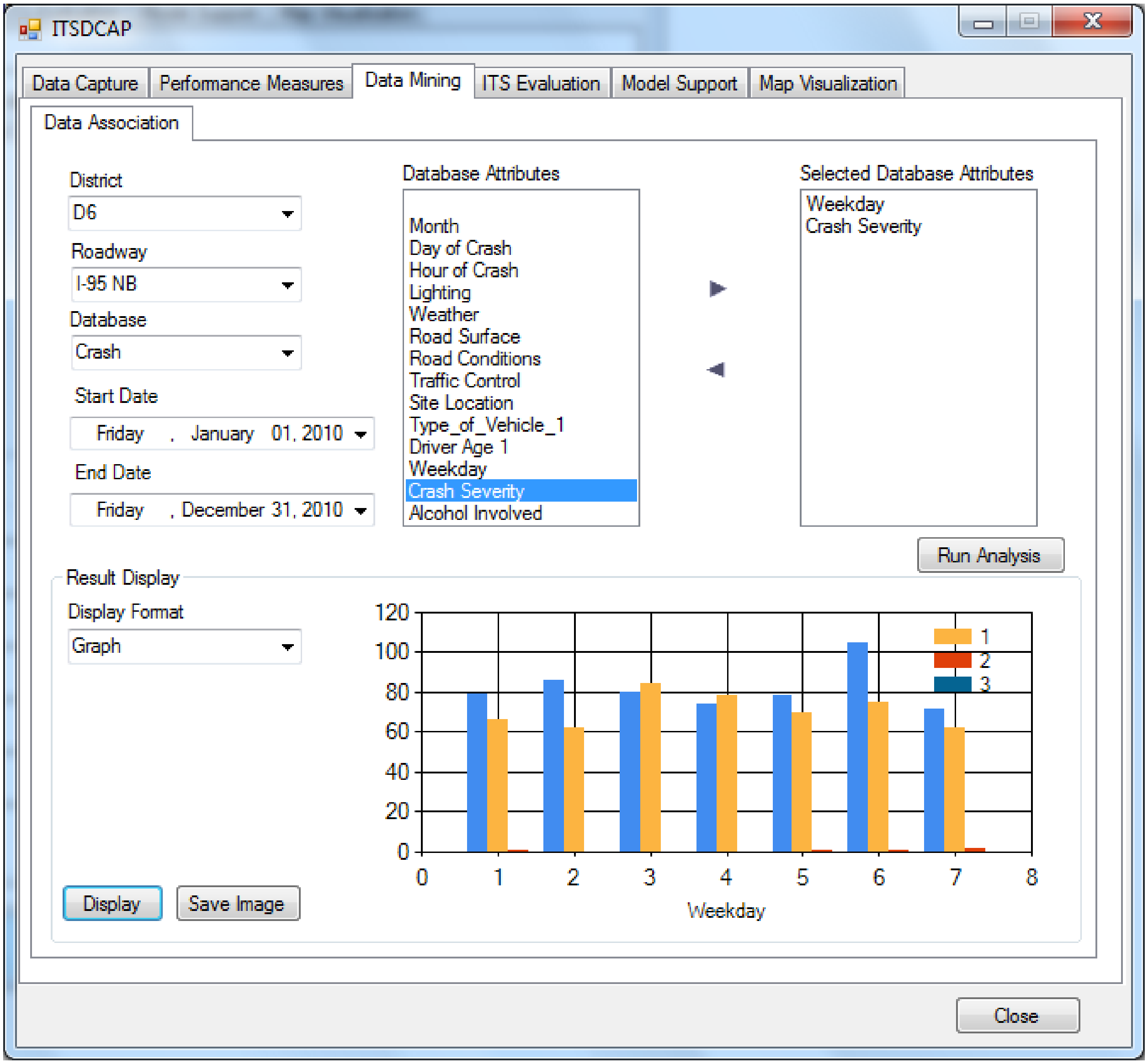


Figure 8-1 Data Mining Interface in ITSDCAP

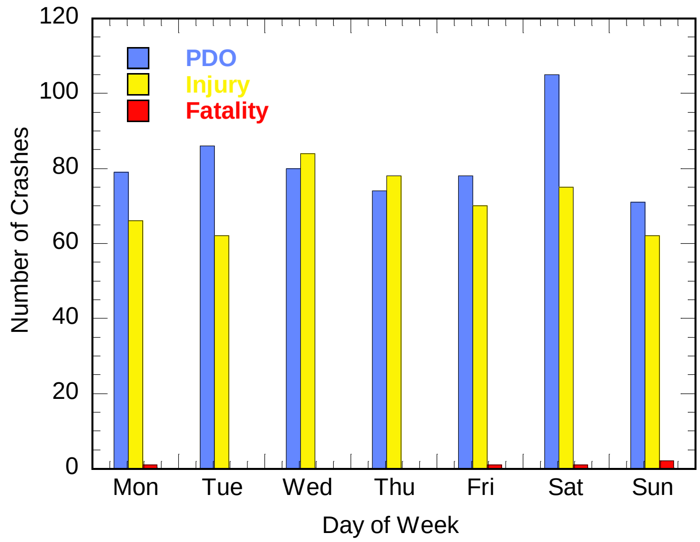
Table 8-1 summarizes the data source and associated attributes considered in the built-in Data Mining Module in ITSDCAP. As shown in this table, attributes from the four databases can be associated with each other or with attributes from other databases. For example, the average number of DMS activations during incidents and the average durations of the activations can be related to various incident attributes. Another example is to associate the number of 511 calls with incident attributes.

Table 8-1 Data Source and Associated Attributes Considered in the Built-In Data Mining Module

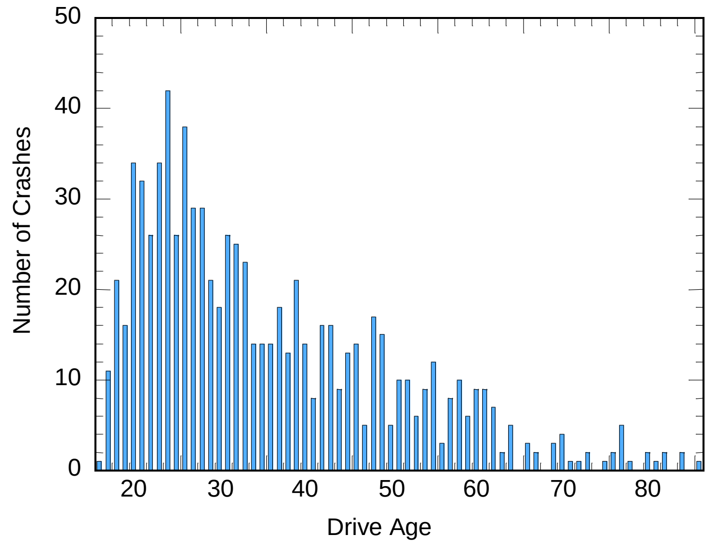
Data Source	**Crash Database**	**Incident Database**	**DMS Data**	**511 Data**
Attributes	• Month • Day of Month • Day of Week • Hour of Crash • Lighting • Weather • Road Surface • Traffic Control • Site Location • Type of Involved Vehicle 1 • Age of Driver 1 • Crash Severity • Alcohol/Drug Involvement	• Roadway Number • Direction • Number of Travel Lane Blocked • Number of Entry Lane Blocked • Number of Exit Lane Blocked • Number of Shoulder Blocked • Shoulder Lane Blocked Only • Last Event Response Plan ID • Is Severity Other • Severity ID • Dispatch Assisted • Dispatch Vehicle Count • User Error Count • Activity Count • Road Ranger Assisted • Road Ranger Number • Road Ranger User Error Count • Road Ranger Activity Count • Severe Incident Response Vehicle (SIRV) Assisted • SIRV Number • SIRV User Error Count • SIRV Activity Count • Year • Month • Day of Month • Day of Week • Hour • Surface • Weather • Lighting • Event Type	• Activated DMS Number • DMS Activation Duration	• Date • Day of Week • Hour • Number of 511 Calls

8.2. Use Cases of the Built-In Data Mining Module

This section presents results from using the basic built-in data mining module to associate data attributes with each other. Figure 8-2 shows the crash frequency as a function of a number of attributes for I-95 northbound in Miami-Dade County. The time period for the analysis is between Jan. 1, 2010, and Dec. 31, 2010. As can be seen in Figure 8-2, by associating different attributes from the crash database, important conclusions can be made, such as time of the day/day of the week with high crash frequency, age of drivers involved in crashes, alcohol-involved crashes, etc.



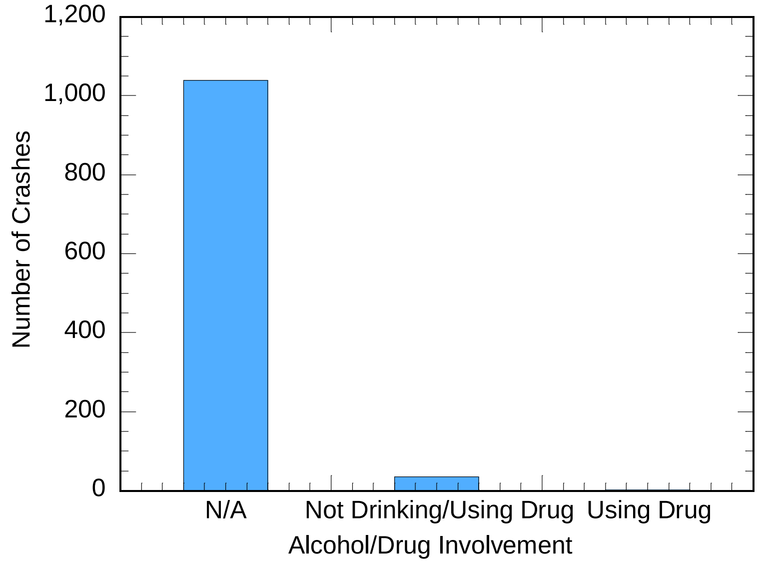
(a) Relationship between Day of Week, Crash Severity and Number of Crashes



(b) Relationship between Driver Age and Number of Crashes

Figure 8-2 Data Mining Case Study Results for Crash Data

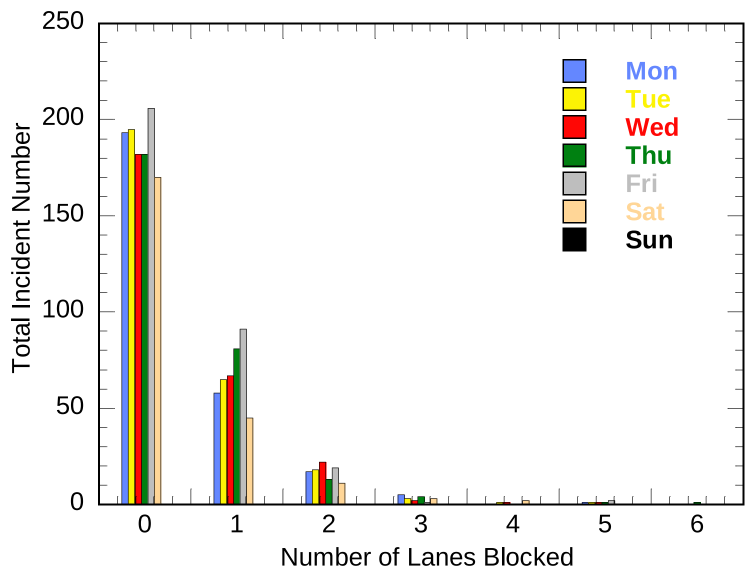
(Continued on Next Page)



(c) Relationship between Alcohol/Drug Involvement and Number of Crashes

Figure 8-2 Data Mining Case Study Results for Crash Data

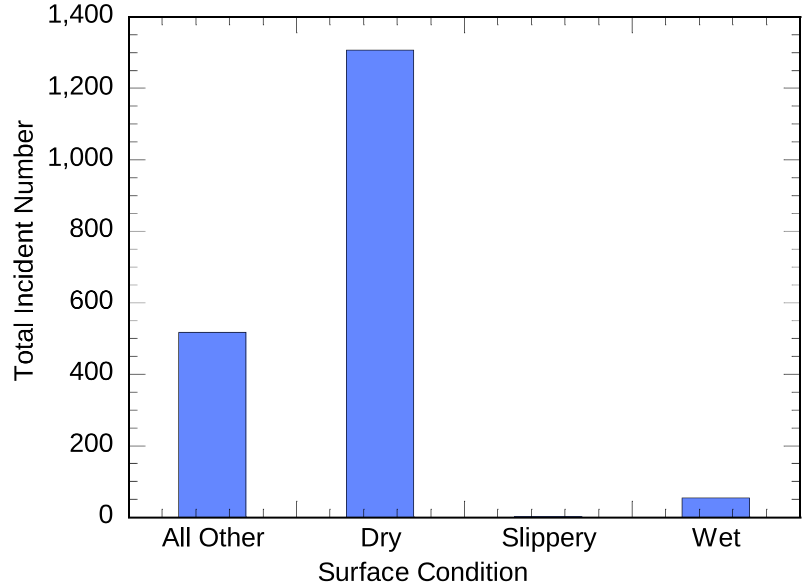
Figure 8-3 shows the relationship between incident attributes and the number of incidents and incident duration for I-95 NB in Miami-Dade County. The time period is between Jan. 1, 2012 and Mar. 31, 2012. Such association allows analysts to determine the attributes of incidents and lane blockages by time of day/day of week, weather, and pavement conditions.



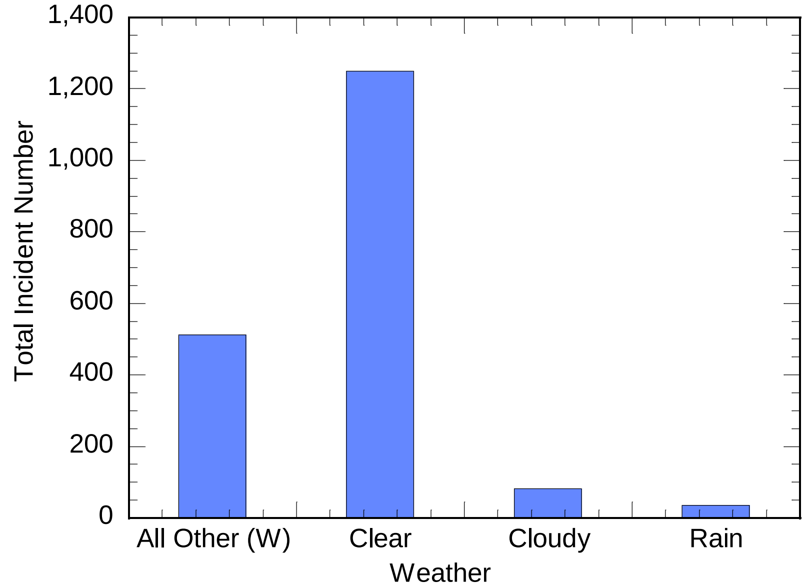
(a) Relationship between Number of Travel Lanes Blocked, Day of Week and Incident Number

Figure 8-3 Data Mining Case Study Results for Incident Data

(Continued on Next Page)



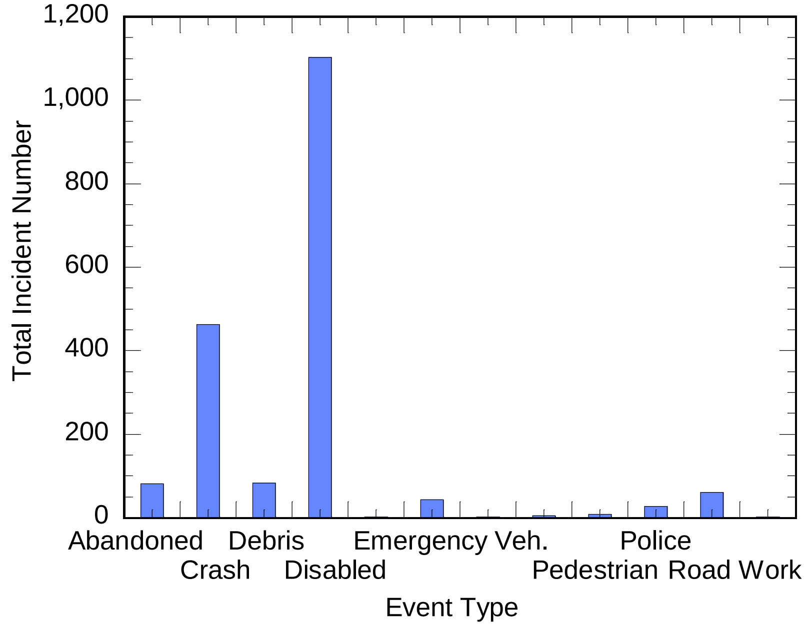
(b) Relationship between Road Surface Conditions and Incident Number



(c) Relationship between Weather Conditions and Incident Number

Figure 8-3 Data Mining Case Study Results for Incident Data

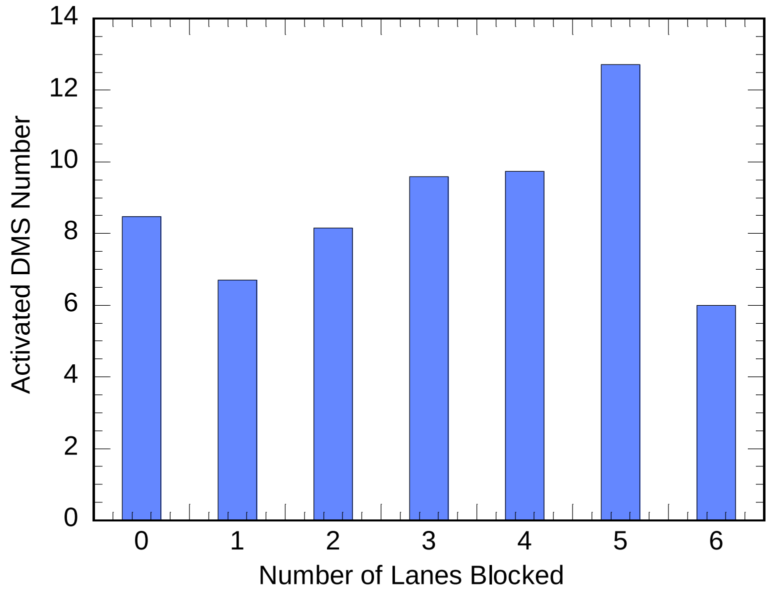
(Continued on Next Page)



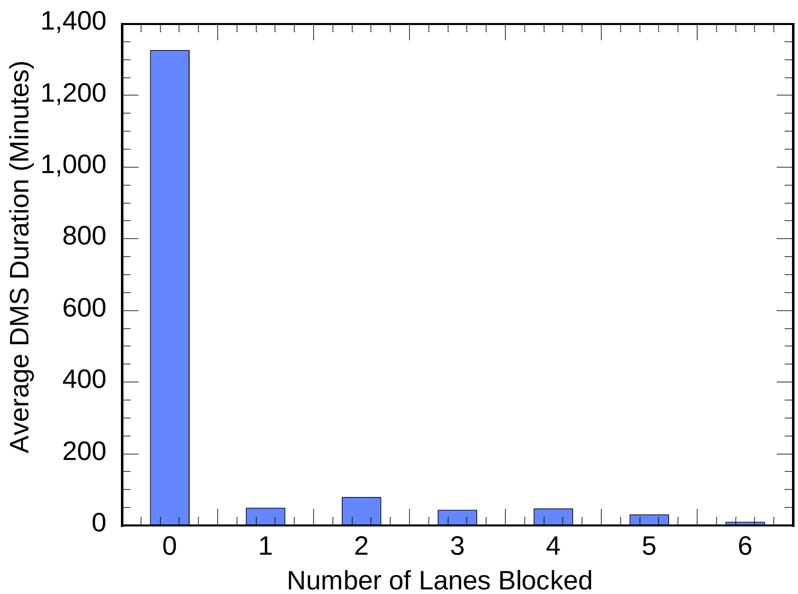
(d) Relationship between Event Type and Incident Number

Figure 8-3 Data Mining Case Study Results for Incident Data

Figure 8-4 shows the relationship between incident attributes and the number and durations of DMS activations for I-95 NB in Miami-Dade County. The time period is between Jan. 1, 2012 and Mar. 31, 2012. Figure 8-5 shows the relationship between incident data and 511 data for the I-95 Corridor in Miami-Dade County. The time period is between Feb. 1, 2011 and Apr. 30, 2011.



(a) Relationship between Activated Number of DMS and Number of Travel Lanes Blocked



(b) Relationship between Average DMS Activation Duration and Number of Travel Lanes Blocked

Figure 8-4 Data Mining Case Study Results for Incident and DMS Data

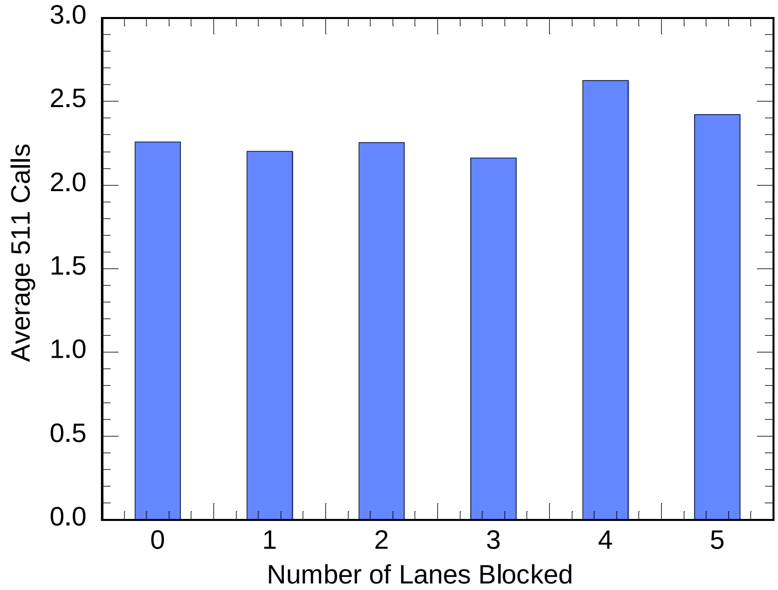


Figure 8-5 Relationship between Average 511 Calls and Number of Travel Lanes Blocked

8.3. Comparison of Crash Data in the CAR and Incident Database

As stated earlier, in addition to the built-in data mining module in ITSDCAP, advanced data mining analyses can be performed based on the extracted data. This section and the remaining sections of this chapter present few examples.

In this section, an initial exploration was made for potentially using the data from the incident management databases, possibly in a combination with the CAR system database in crash analysis. Although the incident databases collected by the TMCs may not have the details of crash attributes available in the CAR system database, it may be useful to complement the analysis under certain conditions. For example, the CAR database does not include minor PDO crashes with estimated damage values less than $1,000. The incident management database does not have this threshold. In addition, the CAR database for a given year does not become available until sometime after the end of the calendar year. Thus, utilizing the incident database may be useful as an interim measure.

An initial exploration was done of the ability of matching crashes that are contained in the FDOT CAR system to crashes that are contained in the SunGuide incident management database. It is important to note that while the CAR database contains crashes exclusively, the SunGuide database contains numerous types of incidents that may affect traffic (e.g., crashes, disabled vehicles, debris, etc.). Because this effort aims to match crashes in the CAR database to the corresponding crashes in the SunGuide system, only the incidents of type crashes (EVENT_TYPE = “1”) in the incident database are included in the analysis. In Florida, all crashes that occur on State roads; which result in a fatality, injury, or property-damage-only (PDO) higher than $1,000, are included in the CAR System. The data in the system is updated yearly. There are thirty-eight data elements in the FDOT CAR database. The incident data in the SunGuide Incident Management Database are mainly collected through FDOT TMC and Road Ranger operations.

In this study, initial attempts were made to match crashes from the CAR system to crash incidents from the SunGuide system using common or similar data elements including time, location and the travel direction of the involved vehicles. The idea was that once the crashes in the CAR database are matched to the corresponding incidents in the SunGuide database, the data elements from the two databases could be merged and used in the analysis. Crash and incident data for the last three months of 2008 for I-95 in Miami-Dade County were examined.

Table 8-2 shows the number of crashes contained in the CAR and the SunGuide systems for each month in our study period, as well as for the total period. In all three months, the incident management database has higher numbers of crashes than the CAR system, probably due to the PDO reporting threshold in the CAR system. Overall, there are 33% more crashes in the incident management system.

Table 8-2 Total Possible Matches

	SunGuide	**CAR**	**Maximum Possible Matches**
October, 2008	212	175	175
November, 2008	160	115	115
December, 2008	185	129	129
Total	557	419	419

A spreadsheet was developed to identify possible matches based on the occurrence times, the travel directions of the vehicles involved, and the locations of the crashes. An issue to consider in the matching was the difference in the location coding between the two databases. Before the locations of crashes could be compared to those of incidents, the mileposts had to be converted to longitude and latitude values, which is how the locations are coded in incident database. However, it should be mentioned that incident longitudes and latitudes as recorded in the SunGuide system are for the closest reference point rather than that of the incident itself. This became more apparent as the incident locations were plotted on a map in Google Earth with many incidents occurring at the exact same location. By definition, a reference point on a highway could be an exit, a mile marker, or a toll plaza. When plotting the reference points for the SunGuide incidents on I-95 in District 6 in Google Earth, it was established that all reference points in our study area represented exits and/or entrances to the highway. Using a 0.50 mile maximum matching distance between the CAR and incident databases, it was possible to match about 30% of the crashes in the two databases. Using a 1 mile matching distance increased the percentage of matching to 41.5 %.

8.4. User Choice of I-95 Express

The I-95 Express HOT lanes in Miami have been in operation for about two years. One of the important aspects of HOT lane operations is the user response to pricing strategies in terms of using the priced HOT lanes. Traditionally, user behavior related to road pricing has been estimated using data collected by means of stated preferences surveys (SP), revealed preferences surveys (RP), simulation, or by means of a mix of the approaches. However, the survey approach is expensive, particularly for large systems. In addition, the data collected using such methods are normally representative of one or a few days and may not capture fine temporal variations in user preferences and other features like user decisions under non-recurrent congestion conditions (incidents, weather, etc.). With respect to the alternative survey methods, SP surveys tend to have serious limitations due to the fact that their answers to the surveys may not reflect the individual’s actual behavior but just an intention. In case of using simulation techniques, the main criticism is the fact that neither the simulated transportation system is a real system, nor are the users actually commuting.

ITS data archives can provide information for estimating actual responses to road pricing based on measured managed lane utilization. Using ITS data for this purpose can provide a significantly lower cost and a cost-efficient data collection method compared to traditional methods. The additional details provided by the ITS data, both in time and space resolutions, allow better representations of real-world environments in transportation models. For example, the use of archived ITS data allows the consideration of drivers' behavior in a number of traffic conditions; including, special events, accidents, work zones, weather events, and incident management strategies.

In addition to travel time, there is a growing body of empirical evidence that travelers value reliability as an important factor in their trip making decisions. Travel time reliability becomes especially important in dynamic pricing applications, where tolls are adjusted based on traffic congestion level in order to maintain a specified level of service. In the case of the HOT lanes, the users may experience only a small reduction in their average travel time over non-toll lanes, but enjoy a substantial reduction in their travel time day-to-day variability. This increased reliability can be critical for travelers with rigid schedule requirements.

The goal of this use case is to examine the use of ITS data to determine user sensitivities to travel time and travel time reliability in a system managed using road pricing strategies. In particular, this study examines the correlation between the observed proportion of users of the HOT lanes and system performance measures considering different treatments of the performance measures. The hypothesis behind this is that different treatments of the system performance measures are expected to correlate differently with the observed proportion of HOT lane users.

8.4.1. Historical Information Considered by Users

The first question to be answered is to determine the period of time prior to the date of travel from which information is significantly correlated with user decisions to select the HOT lanes.

Because of the variations from day to day, it is expected that the users utilize average values of past experiences to decide whether they are going to use the managed lane at a particular time of the day. An investigation was done of the period of time that provides the best correlation between the proportion of HOT users and the differences in travel time between the HOT and the GP lanes. The difference in HOT and GP travel time estimates were computed considering different periods of time prior to the day of travel; including 1 day, 5, 10, 15, 30, 60, and 90 days. In order to assess the degree of correlation between the proportion of the HOT users and the difference in travel time, the coefficient of correlation (R^2) was estimated for the dataset. The results are shown in Figure 8-6 for the AM and PM peak periods. In general, it is observed that the larger the period of time considered, the larger is the coefficient of correlation. However no further increments are observed beyond 60 days where the coefficient of correlation remains constant. The latter suggest 60 days as the number of days from which information is averaged in a traveler's mind when selecting HOT lanes.

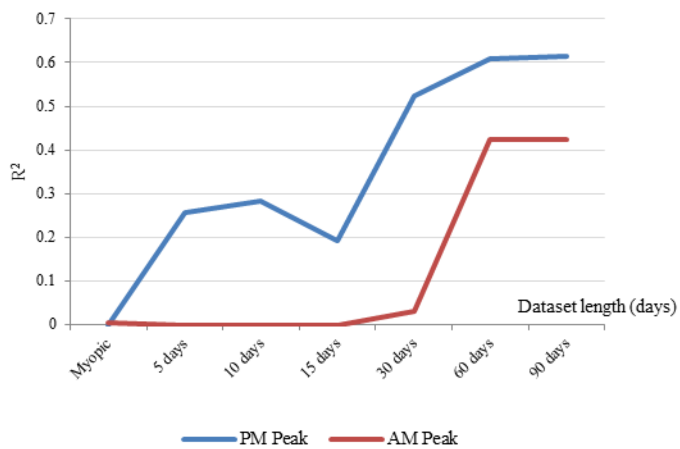


Figure 8-6 R^2 Considering Different Dataset Lengths in AM and PM Periods

Another question to be answered was if there is evidence from the data that the travelers put more emphasis on their latest travel experience (more recent days of travel) compared to relatively older experiences when making their choices. This was examined by determining if utilizing a weighted average travel time (with more weights on the latest experiences) provides better correlations with the proportion of HOT users, compared to average travel times. For this purpose, different exponentially decreasing weights were applied to the historical differences in travel time between HOT and GP lanes. The parameter, which affects how the exponential weight function changes, is referred to as b in this study. A b value of 1.0 represents the case where all the experiences from the past 60 days receive the same weights thereby resulting in a utilized travel time difference that is the simple arithmetic mean of the differences of the past 60 days. The higher the b value the higher the weights put on the latest experiences in calculating the weighted mean. It was determined in this study that a simple average travel time correlates better with the proportion of HOT lane users, as indicated by the results in Figure 8-7. Thus, there is no evidence that the users put more emphasis on their latest experience when making their decisions.

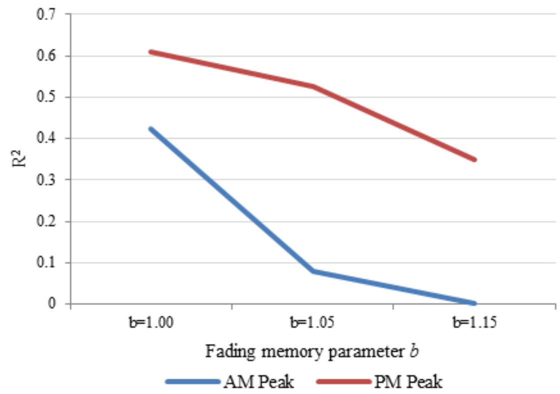


Figure 8-7 *R^2* Considering Different Fading Memory Rules

Finally, the impact of current day conditions on the utilization of the lane was explored. It should be emphasized that this was done for a "normal" day, not for an incident day. Further investigation is needed to reexamine the results for incident days. In this study, several hypothetical cases were explored ranging from *only real-time information is used* (a=0.0) to *only historic information is used* (a=1.0). In all the cases explored, the historical information was based on the mean travel time computed over 60 days of data. The analysis of the results show that the current day difference in travel time between GP and HOT lanes does not correlate well with the proportion of HOT lane users. In some cases, in the AM peak, the results showed that drivers were paying for the use of the HOT lanes, while actually the GP lanes had lower travel times. Figure 8-8 depicts the R^2 values between the observed proportion of HOT users and the travel time difference between GP and HOT lanes. It can be observed that there is a relationship between the coefficient of correlation and the weight considered for the current day information. In particular, a better correlation between the quantities of interest is obtained when current day information is minimal or not considered at all, particularly for AM peak. It should be mentioned again that this conclusion is limited to "normal day conditions" and also for current information dissemination setup, in which the differences in travel time between the HOT and GP lanes are not communicated to the users in real-time.

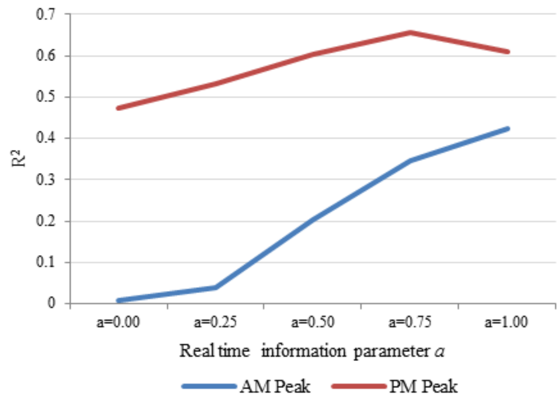


Figure 8-8 R^2 Considering Different Real-Time Information Parameter *a*

8.4.2. Exploration of Performance Measures that Impact HOT Lane Choices

Next, an initial investigation was made to determine if it is possible to utilize ITS data to identify the impacts of the differences between various performance measures of GP and HOT lanes on the proportions of HOT lane users by time of day. Measures that may impact the choice and can be measured based on the collected data, can include travel time, travel time reliability, occupancy, and toll.

The probability of selecting the HOT lanes is modeled using logistic regression. Logistic regression is appropriate when the dependent variable is the population proportion or probability, as is the case in this study. Linear regression analysis (which assumes that the independent variable is distributed following a continuous normal distribution) is inappropriate in this case. The general form of HOT lane utilization probability in a logistic model is as follows:

$$P(y_i = 1 \mid x_i) = p_i = \frac{e^{\alpha+\beta x_i}}{1+e^{\alpha+\beta x_i}} \tag{8-1}$$

The odds of an event occurring (odds ratio) is defined as follows:

$$\frac{p_i}{1-p_i} = e^{\alpha+\beta x_i} \tag{8-2}$$

where p_i is the probability that an instance i will occur, α is the constant, β is the vector of coefficients for independent variables, and x_i is the vector of independent variables.

Estimation techniques like regression are seriously affected by collinearity in the data. Collinearity arises as the explanatory variables used in the model move further and further away from orthogonality, which causes the estimates to become less reliable (high variance). The presence of the collinearity can be ascertained by computing the correlation matrix of the regressors. It has been recommended in the literature that correlation coefficients greater than 0.5, indicate the presence of data collinearity. Based on data visualization, it was observed that travel time, toll, PTI 95th, PTI 80th, and occupancy are all strongly correlated during the AM and PM periods. This is an undesirable characteristic of the data used in this study for parameter estimation because it prevents the estimation of joint models with more than one explanatory variable at a time. Table 8-3 shows the typical correlation factors between candidate explanatory variables based on the collected data.

Table 8-3 Typical Correlation Factors

Base Variable	Average R^2 Values Observed During AM Period			Average R^2 Values Observed During PM Period		
	Fare	PTI 95th	PTI 80th	Fare	PTI 95th	PTI 80th
Travel time	0.81	0.81	0.82	0.82	0.87	0.83

Table 8-4 shows the results from the logistic regression analysis.

Table 8-4 Results of Logistic Regression

Peak Period	Independent Variable	Bo	Bo Significance	B1	B1 Significance	VOT or VOR ($)
PM	Travel Time	1.14	0.007	-0.24	0.16	14.36
	PTI 80th	2.73	<0.001	1.45	0.02	1.4