Final Report

to the

CENTER FOR MULTIMODAL SOLUTIONS FOR CONGESTION MITIGATION (CMS)

Title: Statewide Transportation Engineering Warehouse for Archived Regional Data (STEWARD)

Phase III Final Report

Prepared by the University of Florida
Transportation Research Center

Submitted to

The Florida Department of Transportation
FDOT Project BD545-93
UF Project 72734
and
The University of Florida
Center for Multimodal Solutions for Congestion Mitigation (CMS)
Project 72869

December 15, 2009

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

APPROXIMATE CONVERSIONS FROM SI / METRIC UNITS TO STANDARD / US CUSTOMARY 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
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				
^oF	Fahrenheit	(F-32) x 5 / 9 or (F-32) / 1.8	Celsius	^oC
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

APPROXIMATE CONVERSIONS FROM SI / METRIC UNITS TO STANDARD / US CUSTOMARY UNITS (Continued)				
SYMBOL	**WHEN YOU KNOW**	**MULTIPLY BY**	**TO FIND**	**SYMBOL**
LENGTH				
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
AREA				
mm^2	millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
VOLUME				
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
MASS				
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
TEMPERATURE				
^oC	Celsius	1.8C + 32	Fahrenheit	^oF
ILLUMINATION				
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
FORCE and PRESSURE or STRESS				
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle Development of a Central Data Warehouse for Statewide ITS and Transportation Data in Florida Phase III: Final Report	5. Report Date December 15, 2009 6. Performing Organization Code
7. Author(s) Kenneth G Courage and Seokjoo Lee	8. Performing Organization Report No. FR-72734
9. Performing Organization Name and Address University of Florida Transportation Research Center PO Box 116580 Gainesville, FL 32611-6580	10. Work Unit No. (TRAIS) 11. Contract or Grant No. BD545, RPWO # 93
12. Sponsoring Agency Name and Address Florida Department of Transportation,605 Suwannee Street, MS 30 Tallahassee, FL 32399 University of Florida Center for Multimodal Solutions for Congestion Mitigation, 512 Weil Hall, Gainesville, FL 32611	13. Type of Report and Period Covered Final Report 2008-2009 14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

This report documents Phase III of the development and operation of a prototype for the Statewide Transportation Engineering Warehouse for Archived Regional Data (STEWARD). It reflects the progress on the development and operation of STEWARD since the completion of Phase II in April 2008. The previous effort focused on FDOT District 2 to demonstrate that data from a traffic management center can be centrally archived in a practical manner and that a variety of useful reports and other products can be produced. The current effort included the addition of data from four more FDOT districts (4, 5, 6 and 7) to the STEWARD database. A fully functional Web site was implemented to support users in retrieving data and creating reports. Support was provided to a variety of users for research and operational studies.

Analyses were conducted using the data to demonstrate the extraction of traffic counts from detectors, evaluation of managed lanes, travel time reliability reporting, evaluation of the effect of an incident on freeway performance and a comparison of the detector data characteristics with the principles of traffic flow theory. As a result of this project it was recommended that the STEWARD operation be continued and expanded.

The work was performed under two parallel projects, one of which was supported by the FDOT and the other by the University of Florida. While the objectives of these two projects were stated separately, some of the activities overlapped the project boundaries in a manner that was mutually beneficial to both projects. Therefore, the projects have been combined for reporting purposes to facilitate review and assimilation of the “big picture” by stakeholders. This document serves as the final report for both projects.

17. Key Word ITS data archive, Central data warehouse	18. Distribution Statement No Restrictions

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 138	22. Price

Acknowledgements

The effort described in this report was carried out in connection with two parallel projects. One project was funded by the Florida Department of Transportation Research Office and the other was funded by the University of Florida (UF) Center for Multimodal Solutions for Congestion Mitigation, a federally funded University Transportation Center.

The project mangers were Trey Tillander for the FDOT project and Ines Aviles-Spadoni for the UF project. Technical support was provided by David Chang and Clay Packard from the FDOT Traffic Engineering Research Laboratory (TERL) and by Richard Reel and Steve Bentz from the FDOT Statistics Office. Cooperation and support from the district TMCs in the provision of archived data was excellent, and was a strong factor in the success of the project.

The research team appreciates the support and cooperation of the funding agencies, Central Office staff, district SunGuide managers and STEWARD users in the development and operation of this system.

Preface

STEWARD was designed and implemented in 2008 under a proof of concept project sponsored by the Florida Department of Transportation (FDOT). Its primary function is to provide a repository for data generated by SunGuide traffic management centers (TMCs) within Florida. Its operation was continued and the system was expanded under two parallel projects supported by the FDOT and by the University of Florida's Center for Multimodal Solutions for Congestion Mitigation (CMS), a federally funded University Transportation Center.

The objectives of the FDOT sponsored project were as follows:

1. Integrate the CDW functions with other FDOT data management programs.
2. Transfer the CDW operations to the TERL center in Tallahassee.
3. Automate and fine tune the transfer of daily archive data from district TMCs and the statewide monthly, quarterly, and annual reporting processes.
4. Expand the CDW to include other data sources.
5. Incorporate the CDW functionality into requirements for future SunGuide versions.

The objectives of the CMS project were as follows:

1. Present a series of workshops for potential providers and users of archived data.
2. Expand the scope of the CDW database to include additional SunGuide detectors.
3. Operate the system to provide data and reports to agencies and researchers.
4. Analyze the data to explore congestion modeling relationships.

While the objectives of these two projects were stated separately, some of the activities overlapped the project boundaries in a manner that was mutually beneficial to both projects. Therefore, the projects have been combined for reporting purposes to facilitate review and assimilation of the "big picture" by stakeholders. This document serves as the final report for both projects.

EXECUTIVE SUMMARY

Problem Statement

SunGuide, a system of hardware and software that was developed specifically for the Florida Department of Transportation (FDOT), operates most of the traffic management centers (TMCs) in Florida. SunGuide includes a rudimentary archive element that creates a daily text file containing the basic data produced by each of its sensors during each reporting interval (usually 20 seconds). While the data are numerically accurate, the information is not useful until it is organized geographically within the system, stored in a database that can be interrogated and presented in the form of useful reports.

The Statewide Transportation Engineering Warehouse for Archived Regional Data (STEWARD) was designed and implemented in 2008 year under a proof of concept project sponsored by the FDOT. Its primary function is to provide a repository for data generated by SunGuide TMCs within Florida. As such, it serves as a central data warehouse (CDW) for SunGuide data. The proof of concept project demonstrated data from regional TMCs around the state can be centrally archived in a practical manner and that a variety of useful reports and other products can be produced to meet the requirements of a wide range of users and to provide researchers with a rich supply of data for various purposes. The benefits will not, however, be realized unless a more permanent system can be put in place. With this in mind, the STEWARD operation was continued and the system was expanded under two parallel projects supported by the FDOT and by the University of Florida's Center for Multimodal Solutions for Congestion Mitigation (CMS), a federally funded University Transportation Center.

Project Objectives

The objectives of the FDOT sponsored project include:

1. Integrate the STEWARD functions with other FDOT data management programs.
2. Transfer the STEWARD operations to the traffic Engineering Research Laboratory (TERL) center in Tallahassee.
3. Automate and fine-tune the transfer of daily archive data from district TMCs and the statewide reporting processes.
4. Expand STEWARD to include other data sources.

The objectives of the CMS project include:

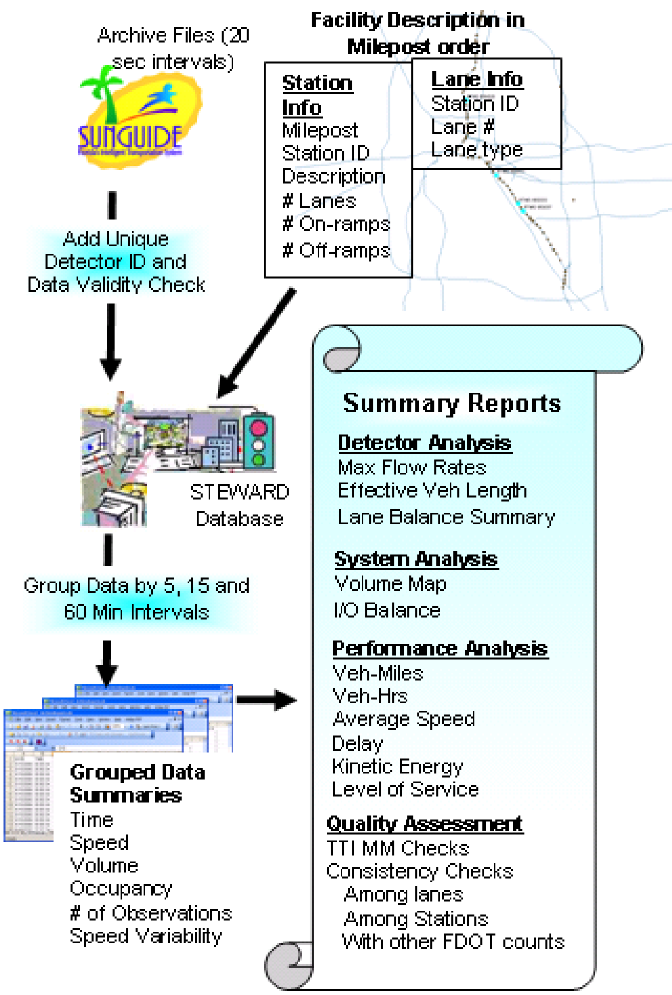
1. Present a series of workshops for potential providers and users of archived data.
2. Expand the scope of the STEWARD database to include additional SunGuide detectors.
3. Operate the system to provide data and reports to agencies and researchers.
4. Analyze the data to explore congestion modeling relationships.

While the objectives of these two projects were stated separately, some of the activities overlapped the project boundaries in a manner that was mutually beneficial to both projects. Therefore, the projects have been combined for reporting purposes to facilitate assimilation of the "big picture" by stakeholders.

Background

There are six separate processes involved in the operation of STEWARD:

1. Configuration of the detector systems in each of the SunGuide centers that provide data. Each sensor in the system is assigned a unique identification, location (milepost, coordinates, etc.), direction, lane number and several operating parameters.
2. Daily transmission and assimilation of the archived data. An automated process has been established whereby each of the SunGuide centers transmits daily data to STEWARD. The configuration data are combined with the archived data to transform the information into the required format.
3. Generation of diagnostic reports for each day to assist the SunGuide system operators in identifying detectors that require maintenance attention.
4. Application of a quality assurance (QA) procedure to identify invalid data: Data that do not pass the QA tests are rejected.
5. Posting the valid data on the project website for general access.
6. Downloading and use of the data from the website by a variety of stakeholders.



STEWARD Overview

The figure above illustrates how these processes fit together.

The final report for the project describes each of these processes in terms of their requirements, development and implementation. The following topics are discussed:

1. *Literature Review:* The literature was reviewed with respect to data archiving activities in other states, the characteristics of the Florida SunGuide traffic management system and quality control of traffic data.
2. *Analysis of the Traffic Archive Data:* A number of analysis tools have been developed to verify the STEWARD data characteristics. Traffic flow principles have also been incorporated into the diverse research applications.
3. *Data Management:* A data management system was designed, established and verified with two years of traffic data warehouse operations. The information is available to general users through the STEWARD Web site.

4. *Statewide Deployment:* STEWARD was designed to accommodate statewide traffic data. At this point, data from Districts 2, 4, 5, 6 and 7 have been incorporated into STEWARD.
5. *Quality Assurance:* A systematic approach has been developed to improve the quality of traffic data.
6. *Workshop Presentations:* As a part of the technology transfer component of this project, a series of workshops was developed and presented at selected locations throughout Florida. Each workshop covered a nominal half-day period.

Products and Deliverables

One of the main products of this research is the system itself. The current status of the system is presented at the right in terms of the number of stations that are covered in each district

District	Facility	Stations
2	I-95, I-295	192
4	I-95, I-595, SR869	334
5	I-4, I-95	452
6	I-75, I-95, I-195, SR-826, US-1	233
7	I-4, I-275	150

The following products were produced and delivered as a part of this project:

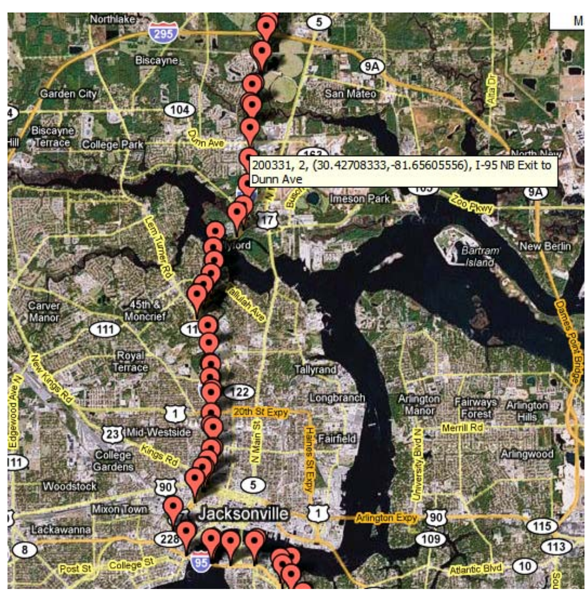
1. The software, including the database and data management components
2. The Web site manager, which provides database query, mapping and report generation components
3. Several desktop utility programs that process data from and create special analysis reports
4. The final report, which summarizes the project activities and presents examples of the analysis capabilities

The STEWARD Web site provides the interface between the archived data and the users of those data. The following items are available on the Web site for each of the SunGuide TMCs:

1. *Station Level Reports*: A single detector station can be selected from a list of available stations. Report data (e.g., traffic counts) may be downloaded for the selected station for further processing.
2. *Section Level Reports:* A contiguous group of stations may be selected by choosing the stations that define the section boundaries. Reports that summarize performance measures for the entire section may be downloaded.
3. *Facility level reports:* All of the data for an entire facility may be downloaded and processed.

Several options exist for reports at all levels. The date and time range may be specified as well as the aggregation level (5, 15 or 60 minutes).

Each facility may be viewed graphically as a satellite photo map, an example of which is shown at the right. All of the stations are identified by clicking their icon on the map.



The Web site also includes a System Resources Page that provides access to project reports, example results, training materials and software utilities that are used for processing data files that have been downloaded.

A summary of Web site utilization statistics for the period from June to August of 2009 is presented in the table at the right.

	Jun. 2009	Jul. 2009	Aug. 2009
Total Visitors	2,176	2,572	1,452
Average Visitors per Day	70	80	64
Total Unique IPs	243	280	293
Total Page Views	10,784	12,566	11,525
Average Page Views per Visitor	4.96	4.89	5.57
Total Hits	35,570	33,337	32,651
Average Hits per Day	1,147	1,041	1,020
Average Hits per Visitor	16.35	12.96	15.77

Results

STEWARD provides TMC managers, district ITS program managers, transportation engineers and management with several useful functions:

- Identification of detector malfunctions
- Calibration guidance for detectors
- Quality assessment and data reliability tests
- Daily performance measures
- Support for periodic reporting requirements
- Extraction of traffic counts as an input to the central and district office traffic counting programs
- Support for the analysis of traffic volume trends

It also provides data for research and special studies. This report provides four examples involving analysis of data from the STEWARD Web site:

1. Analysis of speed, flow and density relationships in the detector data.
2. Analysis of the effect of a selected incident.
3. Analysis of a managed lane.
4. Travel time reliability analysis.

There are several projects and activities that have already benefited from the available data. For example, University of Florida and Florida International University researchers have already made good use of the data. The Web site has shown a continued high level of activity. It is anticipated that activity levels will increase as more data become available and awareness of the STEWARD capabilities increases.

As a result of this project it is recommended that the STEWARD operation be continued and expanded. More specific recommendations are included in the report.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AADT	Annual Average Daily Traffic
ADMS	Archived Data Management Subsystem
ASP	Active Server Pages
ATMS	Advanced Traffic Management System
ATRI	American Transportation Research Institute
AVL	Automatic Vehicle Location
Caltrans	California Department of Transportation
CARS	Crash Analysis Reporting System
CCTV	Closed Circuit Television
CDW	Central Data Warehouse
CHART	Coordinated Highways Action Response Team
CHP	California Highway Patrol
CMS	Center for Multimodal Solutions for Congestion Mitigation
CSS	Cascading Style Sheets
CSV	Comma Separated Values
CV	Coefficient of variation
DB	Database
DMS	Dynamic Message Signs
ETL	Extraction Transformation and Loading
EVL	Effective Vehicle Length
FDOT	Florida Department of Transportation
FFS	Free Flow Speed
FHWA	Federal Highway Administration
FTP	File Transfer Protocol
HAR	Highway Advisory Radios
HICOMP	Highway Congestion Monitoring Program
HCM	Highway Capacity Manual
HOT	High-Occupancy Toll
HOTTER	High-Occupancy Toll and high-occupancy vehicle Traffic Evaluation Report
HOV	High-Occupancy Vehicle
HTTP	Hypertext Transfer Protocol
ID	Identification
IIS	Internet Information Services
ITS	Intelligent Transportation System
LOS	Level of Service
LVBR	Lane Volume Balance Ratio
MM	Mobility Monitoring
MPH	Miles per Hour
NCHRP	National Cooperative Highway Research Program
OWB	Oracle Warehouse Builder
PORTAL	Portland Oregon Regional Transportation Archive Listing
PVM	Price per Vehicle Mile
PeMS	Freeway Performance Measurement System
PATH	California Partners for Advanced Transit and Highways

QA	Quality Assurance
RCI	Roadway Characteristics Inventory
RITIS	Regional Integrated Transportation Information System
RTMC	Regional Traffic Management Center
RTMS	Road Traffic Microwave Sensor
RWIS	Roadway Weather Information Systems
SIS	Strategic Intermodal System
SITSA	Florida Statewide Intelligent Transportation System Architecture
SQL	Structured Query Language
SR	State Road
STEWARD	Statewide Transportation Engineering Warehouse for Archived Regional Data
SWRI	Southwest Research Institute
TASAS	Traffic Accident Surveillance and Analysis System
TCP	Transmission Control Protocol
TERL	Traffic Engineering Research Laboratory
TMC	Traffic Management Center
TriMet	Tri-County Metropolitan Transportation District of Oregon
TSS	Traffic Sensor Subsystem
TTI	Texas Transportation Institute
TVT	Travel Time Subsystem
XML	Extensible Markup Language
UF	University of Florida
VHTT	Vehicle Hours of Travel Time
VMS	Variable Message Signs
VMT	Vehicle Miles of Travel

1 INTRODUCTION AND SUMMARY

The potential benefits of maintaining an archive of data produced by transportation management centers (TMCs) are well recognized [1]. With that in mind, the University of Florida has developed a prototype central data warehouse (CDW) to demonstrate that data from TMCs around the state can be centrally archived in a practical manner and that a variety of useful reports and other products can be produced [2]. This report focuses on the challenges involved in the development of a CDW and on the use of the archived data for various operational and research purposes. The product of this research project is known as the Statewide Transportation Engineering Warehouse for Archived Regional Data (STEWARD).

1.1 Statement of the Problem

SunGuide, a system of hardware and software that was developed specifically for the Florida Department of Transportation (FDOT), operates most of the TMCs in Florida. SunGuide includes a rudimentary archive element that creates a daily text file containing the basic data produced by each of its sensors during each reporting interval (usually 20 seconds). While the data values are numerically accurate, the information is not useful until it is organized geographically within the system, stored in a database that can be interrogated and presented in the form of reports. The problem addressed by this project is the design, implementation and operation of a storage and retrieval system that uses the basic archive files from SunGuide to generate reports that meet the requirements of a wide range of users and to provide researchers with a rich supply of data for various purposes.

1.2 Project Objectives

The following specific objectives have been formulated to address the problem as described above:

- Review the literature as it pertains to traffic management and data archiving.
- Establish the basis for analyzing archived data in terms of freeway traffic flow principles.
- Design and implement a data management scheme to accommodate the archived data.
- Collect and archive data from participating TMCs throughout Florida.
- Develop a quality assurance methodology that makes maximum use of the system aspects of the archived data.
- Identify and explore potential operational and research applications for the archived data.

1.3 Summary of Project Tasks

The following tasks were carried out in support of the stated objectives:

1.3.1 Literature Review

The project team reviewed the literature with respect to data archiving activities in other states, the characteristics of the Florida SunGuide traffic management system, and quality control of traffic data.

1.3.2 Analysis of the Traffic Archive Data

A number of analysis tools have been developed to verify the archived data characteristics. Traffic flow principles have also been incorporated into the diverse research applications.

1.3.3 Data Management System Development

A data management system was designed, established and verified with two years of traffic data warehouse operation.

1.3.4 Statewide Deployment

STEWARD was designed to accommodate statewide traffic data. At this point, data from Districts 2, 4 5, 6 and 7 have been incorporated into STEWARD.

1.3.5 Quality Assurance

Current quality assurance methods focus on individual lane detectors. A systematic approach has been developed to improve the quality of traffic data by adding additional data quality tests based on relationships among individual lane data at a detector station and consistency of data between adjacent stations.

1.3.6 Workshop Presentations

As a part of the technology transfer component of this project, a series of workshops was developed and presented at selected locations throughout Florida. Each workshop covered a nominal half-day period and was followed by discussions with selected personnel to resolve local issues. The program consisted of six sessions supported by PowerPoint presentations and other handout material. The presentation material is included in a separate volume of this report to facilitate future presentations.

The project team presented six workshops at various locations throughout Florida. The presentation schedule was as follows:

- Central Office: TERL/TRANSTAT**,** Tallahassee, Nov. 5, 2009
- District 2: 513 Weil Hall (UF Campus), Gainesville, Dec. 7, 2009
- District 4: SunGuide Center, Ft Lauderdale, Dec. 2, 2009
- District 5: District Office, Deland Dec. 8, 2009
- District 6 SunGuide Conference Room 2, Miami, Dec. 9, 2009
- District 7: Tampa Bay SunGuide Center, Tampa, Dec. 1, 2009

1.4 Organization of the Report

The report and its supplemental material are organized as follows:

1.4.1 Content of the Report Body

The sections that follow this introduction are summarized as follows:

- Section 2 presents the literature review on traffic data warehouses in other states, the SunGuide system in Florida and traffic data quality.
- Section 3 describes the traffic flow theory and performance measures that are applied in the development of summary reports.

- Section 4 describes the CDW requirements, including functional requirements, data transfer requirements and reporting requirements.
- Section 5 summarizes the STEWARD design, operation, web interface, extraction, transformation and loading (ETL) process and operational status. Several items are described in more detail in technical appendices that are presented as separate documents.
- Section 6 describes the development and application of quality assurance procedures, including additional data quality tests proposed to supplement those that are found in the literature.
- Section 7 describes the operational features and reports available from STEWARD. It also summarizes the current and potential operational applications for the archived data.
- Section 8 describes current and future research applications that are supportable by STEWARD data.
- Section 9 presents some interesting examples that demonstrate the use of the data for traffic count extraction, travel time reliability reporting, managed lane analysis and incident analysis.
- Section 10 presents the conclusions and recommendations generated by the project.

1.4.2 Supplemental Material

Supplemental material bound and delivered separately includes a technical appendix volume and the presentation material for a workshop for prospective users of STEWARD. The following technical appendices are included:

- *Appendix 1, STEWARD System Description:* STEWARD consists of three main elements including ETL process, the database and the web-user interface. This appendix describes the details of the system architecture and implementation with respect to those elements.

- *Appendix 2, STEWARD Operation:* A description of the steps required to install the STEWARD software and databases is presented in this appendix. This material has been developed to assist FDOT and ITS contract personnel in setting up STEWARD in their own systems. The topics include the Oracle database program installation, STEWARD deployment and the STEWARD Web site installation.

- *Appendix 3, STEWARD Web Interface:* While some use of STEWARD will be made within FDOT by accessing the databases directly, most users in the future will gain access to the archived data via the Internet. This appendix describes the Internet-based features of STEWARD from the perspective of a user who seeks to query the database and produce reports.

- *Appendix 4, STEWARD Software Utility Documentation:* This appendix describes several utility programs developed for use with STEWARD. Because each of the utilities will be of interest to a different group of users, their documentation has been presented in stand-alone format as a separate section in this appendix. The following documents are included:
    - Appendix 4a: ETL Utility
    - Appendix 4b: ITSCounts
    - Appendix 4c: SunVol Analysis Utility
    - Appendix 4d: HOV/HOT Lane Analysis Utility

1.5 Key Definitions for Geographical Elements

To promote a better understanding of STEWARD, it is necessary to define a few terms related to the geographical structure of the data before any discussion takes place. The following definitions apply to the grouping of geographical elements represented in the STEWARD databases:

- *Lane:* The Wikipedia definition of a lane is “a portion of a paved road [that] is intended for a single line of vehicles and is marked by white or yellow lines.” For purposes of this project, all lanes are assumed to have detection devices capable of monitoring traffic in a manner that is independent of other lanes. The attributes of a lane include:
    - Direction of travel
    - Function (freeway, ramp, auxiliary, HOV, etc)
    - Position on the roadway, indicated by a number representing its order on the roadway from left to right
    - Type of detection device

- *Station:* A station includes all of the lanes that carry traffic in the same direction at the same point on the roadway. For some purposes, ramp lanes at the same location may be considered as a part of the station. The attributes of a station include:
    - Location on the road, indicated by milepost
    - Geographical coordinates for mapping purposes
    - Number of lanes of each type
    - Freeway speed limit
    - FDOT Count station number, if the station used for traffic counting purposes

- *Segment:* A segment refers to the portion of the road that is bounded by two adjacent stations. Segments are unidirectional. Since all detection takes place at the segment boundaries, the station data is generally used to represent the traffic conditions on the segment. For purposes of analysis, the traffic conditions must be assumed to be uniform within the segment.

- *Facility:* A facility includes all of the segments identified with a specific road. Each district may choose to define its own facility structure. Since they are made up of unidirectional segments, facilities are unidirectional as well. The facility structure normally conforms to the highway numbering system. An example of a facility would be “I-95 northbound” within a given district. Facilities do not overlap district boundaries. Within the STEWARD database structure a district may designate up to 10 facilities.

- *Section:* A section is a subset of a facility represented by two or more contiguous segments within a facility that are grouped for reporting purposes. Like segments, sections are unidirectional. Performance measures are accumulated over a section by aggregating the individual performance measures for each segment. An example of a section would be “I-95 northbound between Baymeadows Road and I-10.” Sections do not cross facility boundaries.

- The presentation material for a workshop for prospective users of STEWARD is provided in its own volume to facilitate reproduction and distribution of the workshop handout notes. The program consists of six sessions supported by PowerPoint presentations and other handout material. The workshop content is summarized as follows:
  - Session 1: Executive Summary
  - Session 2: Facility Data Configuration and Maintenance
  - Session 3: Internet Access Features
  - Session 4: Available Reports
  - Session 5: Traffic Count Support
  - Session 6: Other Data Applications

2 LITERATURE REVIEW

To meet the challenges of developing a CDW and demonstrating its capability to perform useful functions, it is first necessary to review the body of past research. The main areas that must be covered include similar traffic data archive systems implemented in other states, the SunGuide traffic management system that will provide the raw data and the quality control concepts that are typically applied in existing systems. This section will cover the state of the practice in each of those areas. It is noted that each traffic data archive system has its own architecture to satisfy its diverse requirements and interface. The details of each system will be identified and described.

There is a potential ambiguity in the terminology found in the literature as it relates to the definition of occupancy. This term is used to describe two entirely different characteristics. The first definition refers to the proportion of a given time interval that a detector senses the presence of a vehicle. It is generally expressed as a percentage value. This characteristic has been shown to be an indicator of the density of traffic on a roadway and therefore a measure of traffic congestion. The second definition relates to the number of persons per vehicle (PPV) within the traffic stream. Vehicles with a specified minimum number of occupants are referred to as high-occupancy vehicles (HOVs). The distinction between these definitions must be inferred from the context in which they are used.

2.1 Traffic Data Warehouses in Other States

Hranac presented the progress of archived data user services (ADUS) [3] as the state of the practice of traffic data warehouses. He defined the following five stages:

Data → Reports (for decision support system) → Application (web 2.0) → Prediction → Control Automation

The reported stages represent the traditional outputs derived from established data archive systems. The application stage is the principal focus of applications that take advantage of traffic data warehouse technology.

A complete and comprehensive review of all data archiving activities is beyond the scope of this project. Instead, three traffic data warehouses were selected to provide an overview of the state of the practice in different locations within the USA. These systems include the Freeway Performance Measurement System (PeMS) system in California, the Portland Transportation Archive Listing (PORTAL) System in Oregon and the Coordinated Highways Action Response Team CHART system in Maryland. These systems are fully deployed and are active in operation and research. The following reviews indicate the current status and future direction of each system.

2.1.1 California – PeMS System

PeMS uses traffic data collection, processing and analysis software developed by the California Department of Transportation (Caltrans), the University of California, Berkeley, and the California Partners for Advanced Transit and Highways (PATH) [4].

System Design

PeMS collects data from 8,100 detector stations in nine districts in California. District TMCs send loop-detector data into PeMS. The raw freeway detector data are sent to PeMS from each Caltrans district over the Caltrans wide-area network in real-time. These data are processed, archived and available from PeMS system immediately. All of the processed traffic information is available to the web users. It also archives the incident data from the California Highway Patrol (CHP) and Traffic Accident Surveillance and Analysis System (TASAS). It also archives the lane closure data from Caltrans in real-time.

Data Quality Control

PeMS developed its own error detection algorithms, using a time series method for the detection of errors. When this algorithm was developed, PeMS collected data from 16,000 loop detectors at 30 seconds resolution. The validity checks were performed on the daily data from each loop detector. If it fails the validity check, the detector is identified as invalid for that day. PeMS marks a daily error flag for each detector as good or bad. To exclude the low-traffic volume conditions at night, the test is performed only on the data collected from 5 a.m to 10 p.m (2,041 samples per day). This serves to prevent the misinterpretation of very low volumes as a case of detector malfunction.

PeMS checks four types of detector errors: stuck-off, two types of hanging-on (non-zero occupancy and zero flow case, or very high occupancy case), and stuck on/off. A summary of the error detection algorithms is presented in Table 1.

- Stuck-off: Detector data are considered bad if 1,200 or more observations have zero occupancy per day.
- Hanging-on: (case 1) If the non-zero occupancy and zero flow case happens more than 50 times per day.
- Hanging-on: (case 2) If the occupancy values are more than 35 percent for more than 200 times.
- Stuck-on/off: If the entropy of the occupancy samples is less than 4.

The entropy of the occupancy is defined as

$$E(x) = -\sum_{x,p(x)>0} (p(x) \times \log(p(x)))$$

where p(x) is the probability that a variable will have the value x.

Data are considered to be invalid if their entropy is less than 4. A low-entropy value indicates that data values are not changing much over time. Originally implemented in PeMS, the entropy criterion has since been replaced with a "consecutive identical values" criterion for easy understanding. If the test results of the detector data are not valid, PeMS discards the entire daily sample and imputes the contents using the data neighborhood.

Table 1. Evaluation Criteria for the PeMS Loop Detector Data

Rules	Description	Parameter
1	Zero occupancy observations > P_1	P_1 = 1200
2	Zero volume and non-zero occupancy observations > P_2	P_2 = 50
3	Occupancy > 0.35 observations > P_3	P_3 = 200
4	Entropy of occupancy < P_4	P_4 = 4

As a single day example, the statistics for 6/4/2009 show that detector errors accounted for 24.3% of the total errors that occurred on that day. Controller errors and communications malfunctions accounted for the remainder. Table 2 summarizes the data quality statistics.

Table 2. PeMS Traffic Data Quality Statistics

# of Detectors	% Good	% Bad	Suspected error								
			Line Down	Controller Down	No Data	Insufficient Data	Card Off	High Value	Inter-mittent	Constant	Feed unstable
26,865	75.7	24.3	2.0	6.8	5.0	1.4	6.3	2.1	0.6	0.0	0.0

(Based on PeMS traffic data on 6/4/2009)

Performance Measures

PeMS has developed several performance reports and congestion analysis tools [5]. This section will summarize the main performance measures provided by PeMS.

- Vehicle-miles traveled and vehicle-hours accrued: Vehicle-miles (VMT) and vehicle-hours (VHTT) are provided for every roadway and facility. These are fundamental measures used to evaluate the movement of goods and people in a transportation system.

- The Highway Congestion Monitoring Program (HICOMP) report is provided for each year to measure freeway system performance. This report measures delay by county and district and identifies the most congested locations. The delay used in this report is defined by Caltrans as the difference between actual travel time and the travel time of the same trip at a constant speed of 35 mph. Negative values are set to zero.

- The Annual Average Daily Traffic (AADT) report is provided each year to evaluate freeway system performance. It included the average daily volumes at 751 locations along the freeway in 2007.

2.1.2 Oregon – PORTAL System

The PORTAL system is a traffic data archive system in Portland, OR. It was developed by Portland State University with the support of Oregon DOT, Metro, the City of Portland, and the TriMet transit agency [6].

System Design

The PORTAL system archives data from approximately 600 loop detectors in Portland, Ore., and Vancouver, Wa.. It receives traffic detector data every 20 seconds from the Oregon DOT and then processes and archives them into the PORTAL database server [6]. Besides the traffic detector data, PORTAL archives incident data, bus data, weather data, dynamic message sign (DMS) data and truck weigh-in-motion records. All of these data are available via the PORTAL web site and some of the data, including traffic detector data, incident and weather data, are available in real-time.

Data Quality Control

The PORTAL system applies two types of detector data tests, including a detector configuration test and communication failure test [7]. The detector configuration test was adapted from the test sets used in an urban freeway-monitoring project. This test verifies that the detector data are within an acceptable operational range. Six conditions are used for the PORTAL system. These rules were developed for 20 sec values from loop detectors. Table 3 summarizes the evaluation criteria for the PORTAL detector data.

Table 3. Evaluation Criteria for the PORTAL Detector Data

Rules	Description
1	Count > 17
2	Occupancy > 95%
3	Speed > 100 mph
4	Speed < 5 mph
5	Low (maximum occupancy per day)
6	Low (average occupancy in peak period per day)

The communication failure test checks the percentage of communication failures or zero traffic volumes during the peak period.

These two test results are used to create the detector status report and provide maintenance requests for suspicious detectors. The PORTAL system policy is to flag and filter out the erroneous data but not to impute the data. PORTAL provides the processed data to various users.

Performance measures

Basic performance measures, such as VMT, VHT, travel time, and delay are provided in the PORTAL Web site. These measures are aggregated over time (5, 15 and 60 minutes) and over lanes (station, corridor and system) [7]. Other performance measures, such as the green measures (emissions and energy consumption), delay cost, and person mobility are under development [8].

2.1.3 Maryland - CHART System

The CHART system is a traffic management system in Maryland. It was developed by The Maryland State Highway Administration, Maryland Transportation Authority, Maryland State Police, Federal Highway Administration, and University of Maryland [9].

System Design

The CHART system supports 155 miles of roadway traffic speed sensors, 70 DMS units, 30 highway advisory radios (HARs), 220 closed-circuit television cameras (CCTV) and 55 roadway weather information systems (RWIS) [10]. The CHART Web site provides most of its traffic reports in real-time. Incident reports, route restrictions/lane closures, live traffic cameras, local weather station data, speed sensor data, and highway message signs are available from CHART [10].

CHART Applications

After the successful deployment of the CHART system in Maryland, two applications were developed from this traffic data management system, including the Regional Integrated Transportation Information System (RITIS) and an incident data visualization tool (Fervor).

RITIS is a system that integrates the existing transportation management data in Virginia, Maryland and Washington, D.C. It receives regional data from multiple agencies. It then fuses, translates and standardizes the data to achieve integrated results. In this project, participating agencies are able to view the entire regional traffic information and use it to improve their operations and emergency preparedness. The traveler information system uses RITIS to provide regional standardized data for traveler information, including web sites, paging systems, and 511 traveler information [11].

Incident data visualization by Fervor is another example of an application developed in the Maryland traffic data management system. The existing incident data analysis tools are defined in the traffic data management system and generate incident reports from the pre-developed reports. The user may also generate an offline data file, download it, and then perform graphing and statistical processes independent of the website application [12]. Fervor also provides web-based, visual analytics applications, with an interactive user interface, geo-spatial analysis, statistical ranking functions, and multi-dimensional data exploration capabilities. A screen capture of the Internet interface is shown in Figure 1.

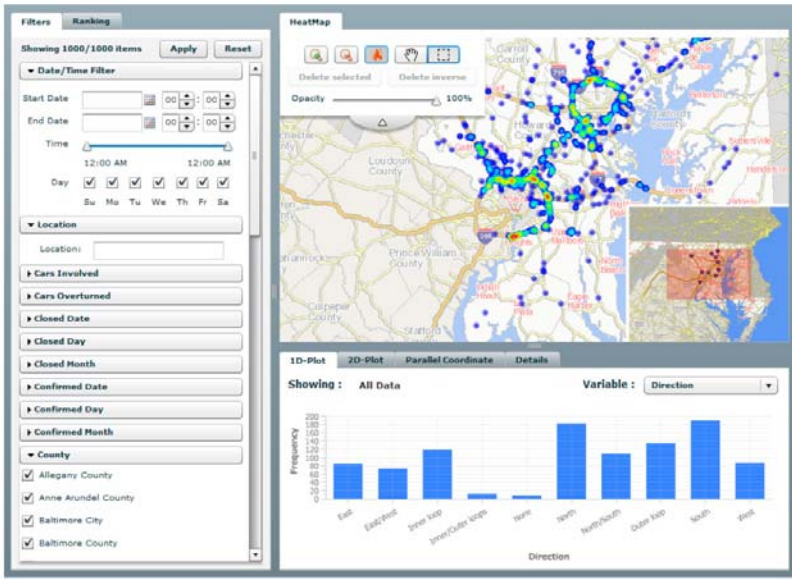


Figure 1: Demonstration of the Fervor application

2.1.4 Summary of Traffic Data Archive Systems

Table 4 presents a summary of the functionality of the three traffic data warehouses discussed previously.

2.2 The SunGuide Traffic Management System in Florida

The SunGuide system is a traffic-management system developed by Southwest Research Institute (SwRI) for the FDOT. Its design goals are to provide the most technically comprehensive advanced traffic management system (ATMS) software available and to establish a standard traffic-management center for use throughout the State of Florida [12].

The SunGuide software is comprised of various subsystems that interact with each other in a cooperative environment. Each subsystem allows the control of roadway devices as well as information exchange across a variety of transportation agencies. This software provides a common base that utilizes a communication interface based on the Extensible Markup Language (XML). Data are stored in an underlying Oracle database [13].

2.2.1 Operational Features

Figure 2 shows the overall architecture of the SunGuide system. A data bus, which provides the communication interface for the entire system, is located in the middle of the figure. Below this data bus, there are 11 subsystems that provide the interface between the external equipment and the SunGuide system. Each traffic equipment component has its own subsystem. For example, the traffic detection subsystem takes care of all the roadway traffic detectors connected to SunGuide. The external subsystems include DMS, CCTV control, video switching, video wall, traffic detection, highway advisory radio, ramp meters, safety barriers, RWIS, AVL/Road Ranger and incident detection.

Table 4: Comparison of Traffic Data Warehouses in other States

	PeMS	*PORTAL*	*CHART*
Coverage	9 out of 12 districts in California	Portland, OR and Vancouver, WA	Northern parts of Washington D.C. and Baltimore, MD
Data sources	- 8,100 detectors - Incident data - Lane closure data	- 600 loop detectors - Incident data - Bus data - Weather data - VMS data - Truck weight-in-motion records	- Detectors at 155 mi of freeways - Dynamic message signs - Highway advisory radios - Closed-circuit television - Roadway weather information systems
Main reports	- Vehicle-miles traveled - Vehicle-hours traveled - Highway congestion monitoring program report - Annual average daily traffic report	- Vehicle-miles traveled - Vehicle-hours traveled - Travel time - Delay	- Incident reports - Route Restrictions/lane closures - Live traffic cameras - Local weather - Station images - Local weather station data - Speed sensor data - Highway message signs
System software	Oracle, PHP, and Google map	Linux, PostgreSQL, Apache, Adobe flash and Google map	Oracle, CORBA, Apache, Javascript
Applications	PeMS 10.1	PORTAL 2.0	- Chart Release 3 - Regional Integrated Transportation Information System (RITIS) - Incident data visualization tool (Fervor).

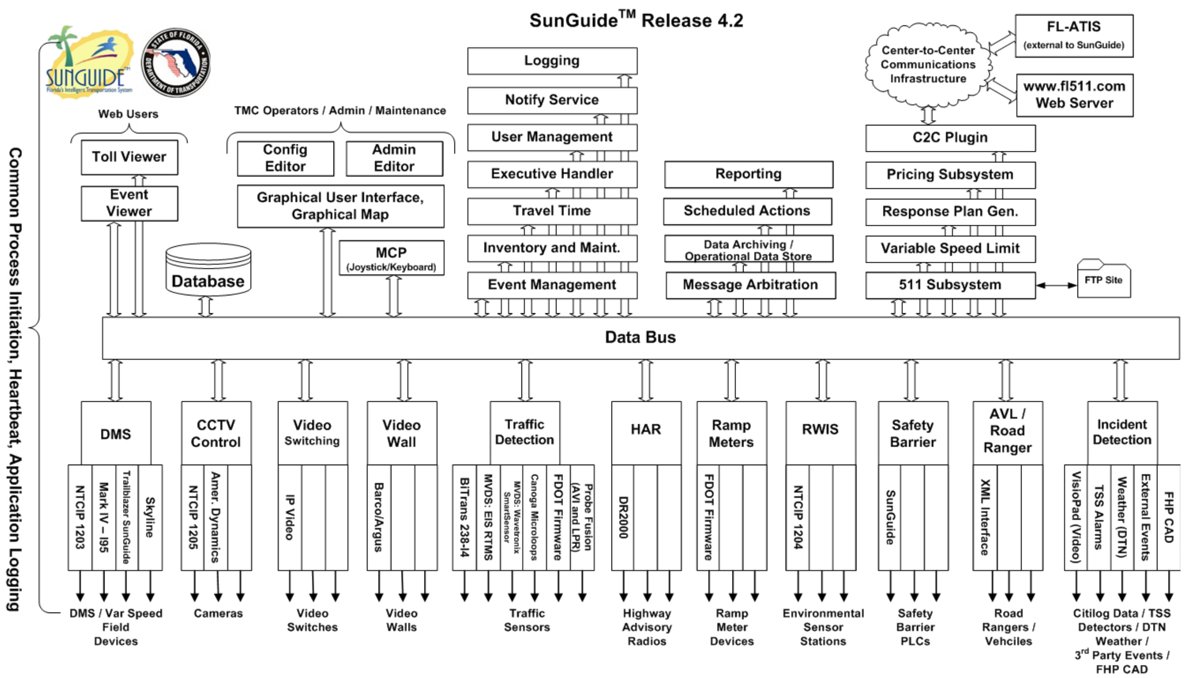


Figure 2. SunGuide system architecture

The SunGuide database archives all of the data from these 11 subsystems as a repository of real-time, configuration and historical data for the system [14]. Other SunGuide subsystems include administrative, operative, and informative subsystems. In Figure 2, these subsystems are located above the data bus. As one of the informative subsystems, the data archive subsystem allows the administrator to query the database and retrieve traffic data. This subsystem is the data source for STEWARD.

2.2.2 Archive Functions

Florida TMCs in Districts 2, 4, 5, 6 and 7 have deployed the SunGuide system as their main operating system for traffic management and control. The traffic sensor subsystem (TSS) from SunGuide is the main data source for STEWARD. TSS data are delivered from the external equipment into the SunGuide traffic detection subsystem every 20 or 30 seconds. Once a day, they are saved as an archive data file by the data archive subsystem. This data file is transferred into STEWARD every day.

The travel time (TVT) data are calculated from TSS data and archived in SunGuide. In the early stages of the STEWARD development, the TVT data were also loaded into the STEWARD database. This practice was discontinued for three reasons:

1. It was difficult to maintain stable travel time link configuration files because of frequent changes while the districts were setting up their TVT links.
2. All of the necessary travel time reporting information can be obtained by analysis of the TSS data.
3. The TSS-based travel time reliability report is more flexible because it is not constrained to an established travel time link system. Any contiguous set of segments may be selected for this purpose.

2.3 Quality Control of Traffic Data

The Quality Control (QC) methods presented in a Federal Highway Administration (FHWA) report on urban freeway monitoring [15] are used as the basis of the STEWARD data quality assessment. The proposed QC methods were developed by the Texas Transportation Institute (TTI) and have been widely applied to traffic data in 30 cities with about 3,000 miles of freeway. One of the main reasons to apply this method to STEWARD is to make the performance measures comply with those of FHWA's urban mobility program.

The suggested QC criteria are divided into three categories:

- *Completeness testing*: The data completeness (availability) measures the number of available data values to the number of total possible values that one could expect.
- *Basic Rules:* Table 5 shows the basic rules. These data quality checks can be characterized as basic validity checks and should detect major problems with the data.

Table 5: Basic Rules of the QC Criteria

Quality Control Rules	Sample Code with Threshold Values
Controller error codes	If VOLUME={code} or OCC={code} or SPEED={code} where {code} typically equals "-1" or "255"
No vehicles present	If SPEED=0 and VOLUME=0 (and OCC=0)
Consistency of elapsed time between records	Elapsed time between consecutive records exceeds a predefined limit or is not consistent
Duplicate records	Detector and date/time stamp combination are identical.

- *Quality Control Criteria:* The quality control criteria expand on the basic rules by establishing quantitative thresholds against which the data values may be checked. Table 6 shows the details of the criteria. These data quality checks are designed as quality control criteria to detect more subtle erroneous or suspect data that could potentially go undetected with these basic rules. The thresholds presented in this table were obtained from Reference [15].

Table 6: Quality Control Criteria

Quality Control Test	Sample Code with Threshold Values
QC1-QC3: Logical consistency tests	If DATE={valid date value} (QC1) If TIME={valid time value} (QC2) If DET_ID={valid detector location value} (QC3)
QC4: Maximum volume	If VOLUME > 17 (20 sec.) If VOLUME > 25 (30 sec.) If VOLUME > 250 (5 min.) If VPHPL > 3,000 (any time period length)
QC5: Maximum occupancy	If OCC > 95% (20 to 30 sec.) If OCC > 80% (1 to 5 min.)
QC6: Minimum speed	If SPEED < 5 mph
QC7: Maximum speed	If SPEED > 100 mph (20 to 30 sec.) If SPEED > 80 mph (1 to 5 min.)
QC8: Multi-variate consistency	If SPEED = 0 and VOLUME > 0 (and OCC > 0)
QC9: Multi-variate consistency	If VOLUME = 0 and SPEED > 0
QC10: Multi-variate consistency	If SPEED = 0 and VOLUME = 0 and OCC > 0
QC11: Truncated occupancy values of zero	If OCC = 0 and VOLUME > MAXVOL where MAXVOL=(2.932*ELAPTIME*SPEED)/600
QC12: Maximum estimated density	IF ((VOLUME*(3600/NOM_POLL))/SPEED) > 220 where NOM_POLL is the nominal polling cycle length in seconds.
QC13: Consecutive identical volume-occupancy speed values	No more than eight consecutive identical volume-occupancy-speed values. That is, the volume AND occupancy AND speed values have more than eight consecutive identical values, respectively. Zero (“0”) values are included in this check.

3 APPLICABLE FREEWAY TRAFFIC-FLOW PRINCIPLES

The daily archive data from SunGuide for each polling interval (typically 20 seconds) includes only the volume (number of vehicles), average vehicle speed and the proportion of time that each detector was occupied by a vehicle (denoted as occupancy). The raw data must be manipulated and interpreted in terms of standard traffic flow principles and measures to provide information that is of value to the end user.

This section will review the quantitative measures that are commonly used to describe the flow of traffic. It will identify the relationships that have been developed in the past between the various measures. It will also identify the ways in which the descriptive measures may be incorporated into measures that evaluate the performance of a freeway facility. Later sections will describe the application of the principles put forth in this chapter for the following purposes:

- Development of requirements for performance measures from the CDW
- Development of the computational methodology by which those requirements can be met
- Application of traffic flow principles to the evaluation of the quality of the archived data
- Investigation of the archived data to determine how well their internal relationships conform to established principles

3.1 Speed Flow Rate and Density Relationships

The macroscopic descriptors of traffic flow are flow rate, speed and density. Mathematically, these descriptors are related by a simple equation:

$$Q = K*U$$

Where (using the common symbols from the literature and commonly applied dimensions)

- Q = Flow rate (Vehicles per hour)
- K = Density (Vehicles per mile)
- U = Speed (Miles per hour)

Thus, any of the three parameters may be computed deterministically given the other two. The nature of traffic flow creates certain internal dependencies between the parameters based on the widely observed phenomenon that speed drops as density increases. These internal relationships have been incorporated into several empirical models that make it possible to compute the value of any two parameters given the third. The speed-flow density relationships from the archive data will be investigated in a later section of this report. Note that some of the research projects identified later used the STEWARD data to investigate speed-flow relationships.

3.1.1 The Fundamental Diagram

Greenshields developed the original model of these relationships in 1935 [16]. Greenshields proposed a linear relationship between speed and density, thereby creating parabolic speed-flow and flow-density relationships. Known at the time as the “fundamental diagram,” the Greenshields relationships endured for many years. The fundamental diagram is illustrated in Figure 3.

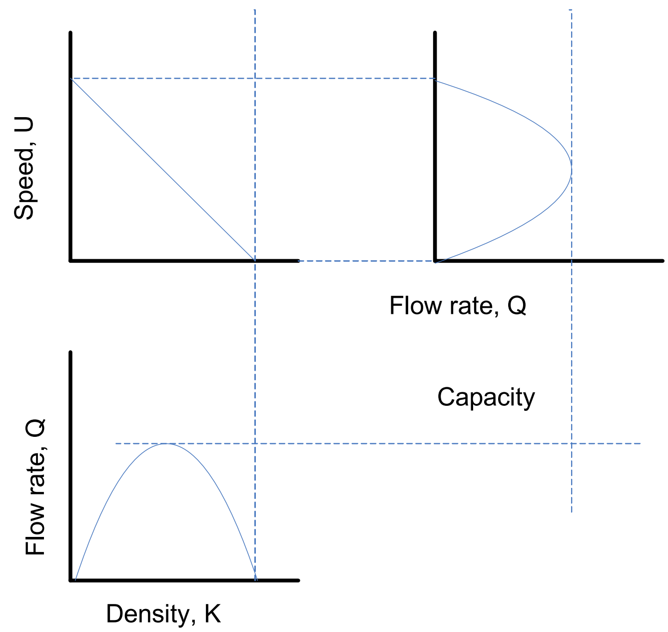


Figure 3. Fundamental diagram relating speed, flow rate and density

Some other important parameters can be derived from the individual relationships in the fundamental diagram shown in Figure 4.

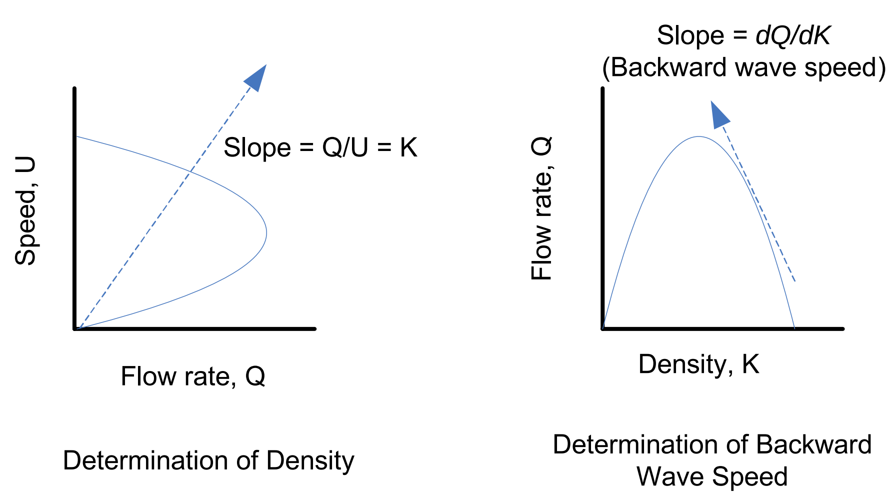


Figure 4. Determination of parameters from the fundamental diagram

The density at any point on the speed-flow curve may be determined as the slope of the radius vector from the origin to that point. The speed of a backward wave during a shift in the operating point of the flow-density curve may be obtained as the rate of change of flow with respect to density or dQ / dK. Backward wave speed speeds are generally computed numerically from a shift in operation from Point 1 to Point 2 points as $(Q_2\text{-}Q_1)/(K_2\text{-}K_1)$.

3.1.2 Other Speed-Flow-Density Models

Several other models have been proposed since the Greenshields classic paper was published. Greenberg's model is one of the nonlinear relationships for speed and density [16].

$$u = u_0 \times \ln(\frac{K_j}{K})$$

where K_j = Jam Density (veh/mi)
u_0 = Optimum Speed (mi/hr)

Optimum speed is defined as the speed at which the traffic flow is at capacity level and the jam density is defined as the density at which vehicles are approaching bumper-to-bumper spacing and stopped.

Current thinking, based on empirical observation, is that it is necessary to divide the model space into two regimes representing oversaturated and undersaturated operation, respectively. It is also accepted that there is a region of unstable flow that separates these regimes and that this region is difficult to quantify mathematically. The current concept of the speed-flow density relationships is illustrated in Figure 5.

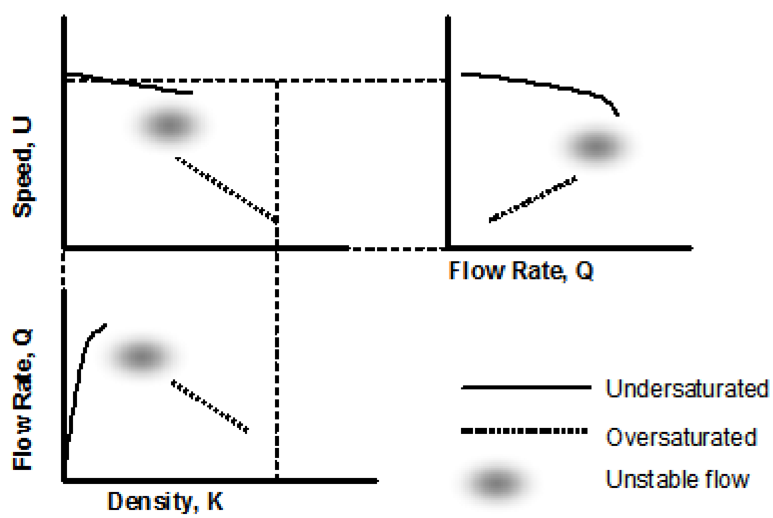


Figure 5. Current concept of the speed-flow-density relationship

The speed-flow-density relationships from the archive data will be investigated in a later task. Note that several of the research projects identified later used the STEWARD data to investigate speed-flow relationships.

3.1.3 Highway Capacity Manual Treatment of Speed, Flow and Density

The Highway Capacity Manual (HCM) presents an example of a typical empirical speed-flow relationship is illustrated in Figure 6, which shows a number of individual observations taken over short intervals of time. The source: of this figure is Exhibit 11-1 of the 2010 HCM [17]. It is important to note that the observations in the bottom (oversaturated) part of Figure 6 are associated with backup from a downstream bottleneck. The conditions reflected in these observations will not occur in a basic freeway segment with no downstream bottleneck.

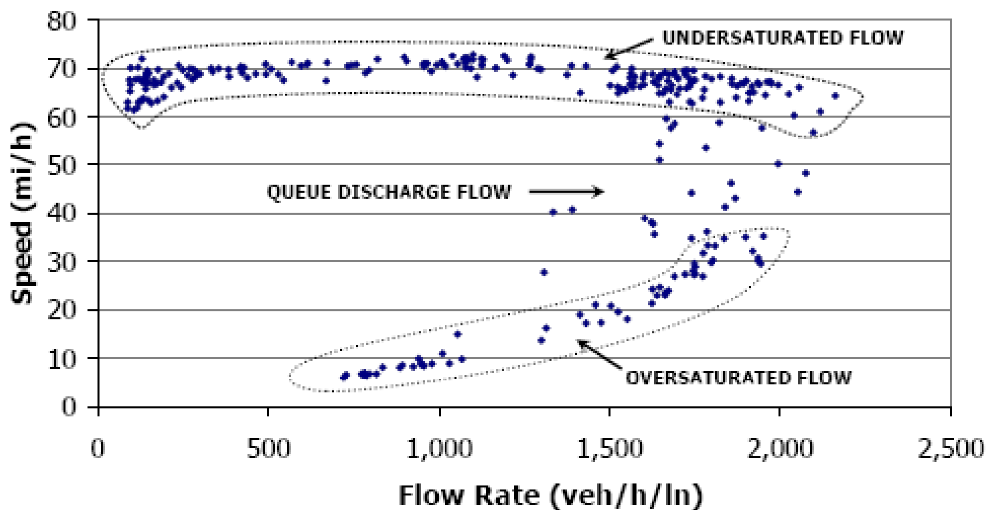


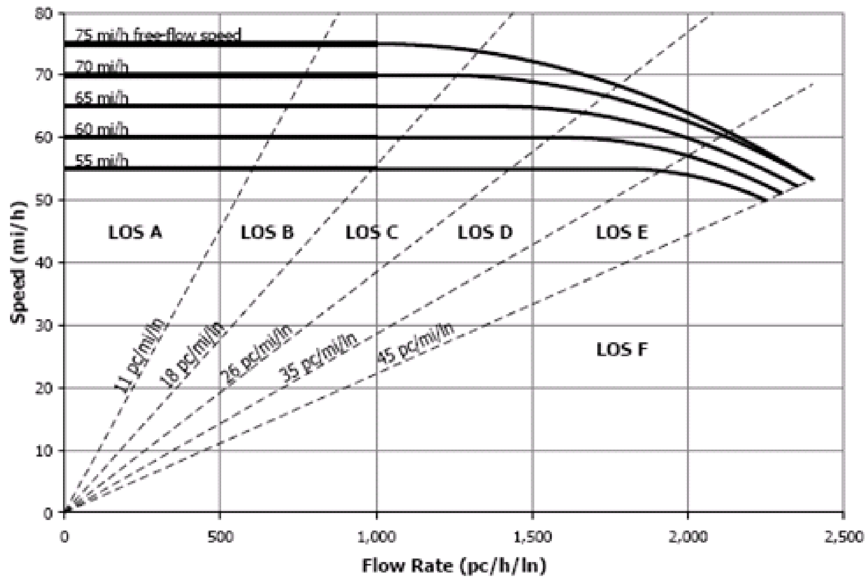
Figure 6. Typical speed-flow relationship presented in the HCM

Recognizing this, the speed-flow curves are presented in the basic freeway segments chapter of the 2010 HCM as shown in Figure 7. Note that these relationships do not extend past the point of capacity. The free-flow speed is an important parameter in this relationship. Lower free-flow speeds result in lower maximum flow rates.

From the fundamental relationship (Q=K*U), it is possible to compute the density at any point on the curves and to represent the density graphically as the slope of the radius vector as illustrated previously. Density is important because it is the measure used to determine the level of service on a basic freeway segment. The density thresholds for each level of service are shown in this manner in Figure 7. Note that the speed-flow graphics terminate for each value of free-flow speed at a density of 45 veh/mi/ln. This density level defines the capacity of the segment.

One important difference between the Greenshields fundamental diagram and the 2010 HCM is the speed at which the capacity is determined to occur. Because of the symmetry in the Greenshields parabolic relationship, the capacity occurs at a speed equal to half of the free flow speed. As indicated in Figure 7, the 2010 HCM places the capacity at speeds between 50 and 54 mph.

Figure 7. Speed-flow relationships for basic freeway segments in the HCM



Considering an example of 75 mph FFS, the Greenshields relationship would place the maximum flow rate at 37.5 mph, while the 2010 HCM would place it at 54 mph. The difference is probably due to a combination of changes in car following behavior in the last 75 years and advances in the modeling of traffic flow.

The HCM flow-density model is also of interest, especially when dealing with oversaturated operation. The segment-oriented chapters of the HCM (basic freeway segments, weaving segments and ramps) do not deal with oversaturation explicitly. In these chapters, Level of Service F is declared whenever the demand exceeds the capacity.

The freeway facilities chapter does recognize oversaturation and queues are propagated upstream from bottlenecks and released when the bottleneck situations are cleared. Figure 8 shows the assumed flow-density relationship in the HCM for undersaturated and oversaturated conditions. Note that this relationship conforms generally to the shape shown in Figure 5, except that there is no region of unstable flow shown here. The specific values indicated in Figure 8 apply to a free flow speed of 75 mph. The left (undersaturated) side conforms to the 75 mph curve presented in Figure 7.

The right (oversaturated) side assumes a linear decline in flow between the capacity of 2400 veh/ln/h at a density of 45 veh/mi/ln and the jam density, which defaults in the HCM procedure to 190 veh/mi/ln. The value of 190 veh/mi/ln is equivalent to a spacing of 27.8 feet (5280/190) from front bumper to front bumper. Assuming an average vehicle length of 16 feet, the space gap between vehicles (front bumper to rear bumper) would be slightly less than 12 feet at jam density.

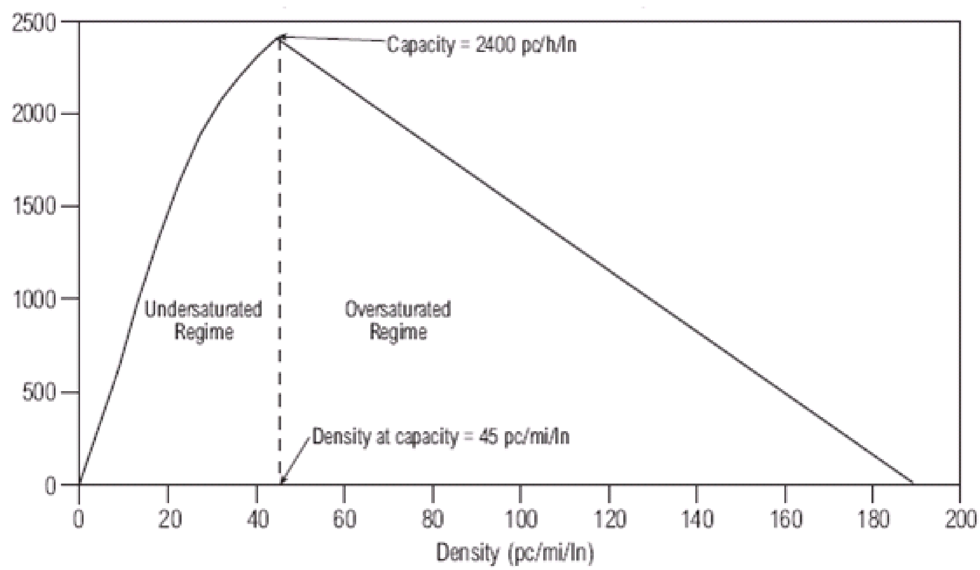


Figure 8. HCM flow-density relationship for basic freeway segments

The relationships shown in Figure 8 are used to project queues in both directions from bottlenecks. The linear relationship between speed and flow introduces the very convenient approximation of a constant backward wave speed, independent of the operating point. For the given assumptions, the calculated speed of the backward wave would be

$$\frac{2000}{(190-45)} = 16.5mph$$

3.2 Level of Service (LOS)

Level of service is defined by the HCM in terms of a six-letter grade system (A through F) with "A" representing free-flow conditions and "F" representing conditions in which demand exceeds capacity. The LOS criteria are specified in the HCM for each type of facility, based on threshold values of a selected performance measure. The selected performance measure for a freeway segment is the average density within the segment. The LOS thresholds for density are given in Table 7. (Source Exhibit 11-3 of the 2010 HCM)

Table 7: Level of Service Thresholds in the 2010 HCM

Level of Service	**Density (pc/mi/ln)**
A	≤11
B	> 11 ≤ 18
C	> 18 - ≤ 26
D	> 26 - ≤ 35
E	> 35 - ≤ 45
F	Demand exceeds capacity (> 45)

Density is not an explicit measure provided by the archive data; however, reasonable approximations can be obtained from the speed-flow-density equation by dividing the flow rate by the speed. Thus, it is possible to estimate the level of service for each freeway segment represented in the archive data.

3.3 Platoon Propagation

All traffic-flow models and theories must satisfy the law of conservation of the number of vehicles on the road. Assuming that the vehicles are flowing from left to right, the continuity equation can be written as

$$\frac{\partial k(x,t)}{\partial t} + \frac{\partial q(x,t)}{\partial x} = 0$$

where x denotes the spatial coordinate in the direction of traffic flow, t is the time, k is the density and q denotes the flow. This relationship offers some research potential for archived data applications to congestion modeling

3.4 Maximum Flow Rates

A very high-flow rate (e.g., greater than 2,400 vph in any lane) could be an indication of a detector calibration problem. The maximum flow rate for any interval was one of the QC criteria mentioned previously.

3.5 Effective Vehicle Lengths

The effective vehicle length is defined as the length of the vehicle plus the length of the detection zone. It may be calculated from the volume, speed and occupancy values for each time interval. The consistency of effective vehicle length provides another quality assessment indicator that will be discussed later in this report.

3.6 Lane Volume Balance Ratio

The lane volume balance ratio (LVBR) is expressed as the ratio of the highest to lowest lane volume at each station. If all lane volumes at a given station were identical, then the lane balance value would be 1.0. During periods of light flow, the LVBR is essentially a reflection of driver preference. On the other hand, during periods of moderately heavy flow, LVBR values above 1.5 might indicate detection problems unless a reasonable explanation, such as a downstream lane closure, can be found. In some cases, traffic patterns can result in non-balanced lane volumes. For example, imbalances could occur ahead of major diversion points or in cases were a queue backs up from a signalized intersection to the freeway resulting in a reduction of the exit lane capacity.

3.7 Input/Output (I/O) Volume Balance

The total volume entering and leaving each link in the system, including freeway and ramp inputs and outputs should balance, except for short intervals in which congestion is either building or dissipating. Over reasonable time periods, an unbalance between inputs and outputs would suggest volume-counting errors unless there are entrance or exit ramps without detectors.

4 CENTRAL DATA WAREHOUSE REQUIREMENTS

4.1 Functional Requirements

The first step in the development of an information storage and retrieval system is to determine what the system must do. To this end, it was determined in consultation with various stakeholders that the system must provide TMC managers, District ITS program managers, traffic operation engineers, and management with the following useful functions:

- Identify detector malfunctions
- Provide calibration guidance for detectors
- Perform quality assessment data reliability tests on data
- Provide daily performance measures for system and statewide performance measures
- Facilitate periodic reporting requirements
- Provide data for research and special studies

A summary of the required functions and data flow is presented in Figure 10.

4.1.1 Raw SunGuide Archive Data

The Traffic Sensor Subsystem (TSS) data are stored in comma-delimited flat files, with each file representing a 24-hour day. Zipped versions of these files are provided daily by the TMC staff by a process described in Section 5 of this report. The TSS data file naming convention is "TSS-mmddyyyy--1.dat." Each record in the file represents the volume, speed and occupancy data from one lane over a single 20-second period. An example of the format is shown in Figure 9.

Timestamp	Detector_id	Lane_id	Speed	Volume	Occupancy
00.00.04	RTMS 95N003	R95N003_01Lane_01	55	1	1
00.00.09	RTMS 95S004A	R95S004A_01Lane_03	55	1	1
00.00.09	RTMS 95N006	R95N006_01Ramp_01	0	0	0
00.00.09	RTMS 95N026	R95N026_01LaneN_01	0	0	0
00.00.09	RTMS 95N026	R95N026_04LaneS_01	55	2	6

Figure 9. Example of the raw data from the SunGuide TSS archive

The *detector_id* and *lane_id* fields contain the station and lane detector names assigned by the TMC. Each district uses its own naming convention.

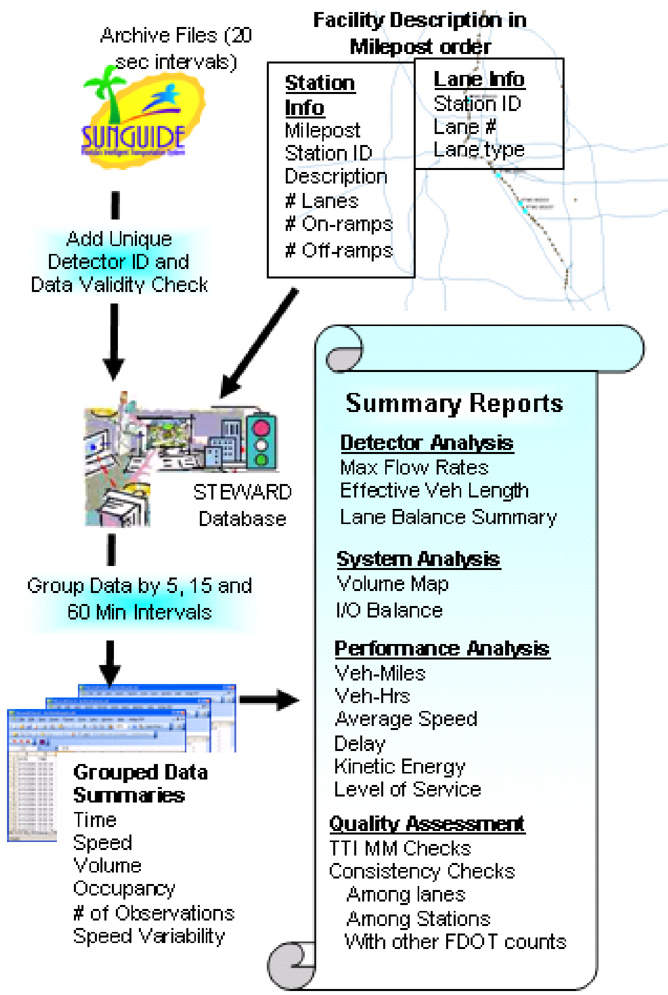


Figure 10. Summary of required functions and data flow

4.1.2 TMC Configuration Requirements

Not all of the information required to convert the raw data to the STEWARD database is contained in the raw data. Some additional information is required for three purposes:

1. To ensure that each record in the STEWARD database represents a globally unique time and location.
2. To support the analysis and reporting of system-based measures and quality assessment.
3. To relate the measures obtained from a specific location to other forms of data, such as RCI, Statistics Office counts, crash records etc.

Two facility information databases must be created for each facility to be included in the STEWARD database. This information must be presented in two Excel spreadsheets.

The Station Data Spreadsheet

The station data spreadsheet must include the following fields for each station on the facility:

- *Station_Index:* This is a number assigned sequentially to all stations in a facility. It is required for internal processing purposes and does not appear in the database or the reports.
- *Stationcdw_Num*: This is a 5 character of the form dfnnn, where
    - d represents the district number
    - f represents the facility number within the district (0-9)
    - nn represents the station number within the facility (0-999)

  An example of a station number in District 2 would be "20001."
- *Description:* A physical description of the station (Example: I-95 NB at Forest St). Some districts embed the description in their station ID and lane ID.
- *Status:* This indicates the known status of the station (0 = Normal, 1 = Offline, 2 = Undetected). The offline stations will not be reported as defective. The undetected station locations are required for the input/output analysis to indicate that the inputs and outputs for a specific link should not be expected to balance.
- *Road:* This is the name given to the facility (Either I-95N or I-95S in District 2).
- *Longitude and Latitude:* These are expressed in degrees and decimal degrees.
- *State_Milepost:* This is required for sequential ordering of stations (Example 351.451).
- *Roadway_Id:* This is required for correlation with RCI and crash data and for identifying the county number for generating traffic count files (Example 72020000).
- *Roadway_Milepost:* This is required for correlation with RCI and crash data and to identify the county number for generating traffic counts (Example 2.7).

The Lane Data Spreadsheet

The lane data spreadsheet must include the following fields for each detected lane on the facility:

- *CDWStation:* This is the same 5-digit station number as in the station data spreadsheet. It is used as a key to relate the station and lane data (Example 20001).
- *Lane:* The lane number reference in the STEWARD database (Example: 20001131). The compositors of the lane description are
    - CDW Station number (5 characters)
    - Direction (1 character)
    - Function code (1 character)
        1. Left entrance ramp
        2. Left exit ramp
        3. Freeway main lane
        4. Right entrance ramp
        5. Right exit ramp
        6. Auxiliary lane
        7. HOV Lane
    - Lane number, starting from the left side

- *Tmc_Id:* The lane ID used by the archive file generated by the TMC (Example: R95N001_01Lane_01). Note that this must match the lane_id field in the archive data file. Archive data records in which the lane_id is not found in the lane data spreadsheet are reported as orphan lanes. Records in the lane data spreadsheet that have no matches in the archive data file are reported as Null Lanes.

- *Det_Type:* Always "RTMS" for District 2. This is not used in any analysis at present, but is provided for future use.
- *Direction:* The direction of the traffic detected on this lane (1 = Increasing mileposts, 2 = Decreasing mileposts)
- *Status:* This indicates the known status of the station (0 = Normal, 1 = Offline, 2 = Undetected). The offline stations will not be reported as defective. The undetected station locations are required for the input/output analysis to indicate that the inputs and outputs for a specific link should not be expected to balance.
- *Roadway_Id:* Required for correlation with RCI and crash data (Example 72020000). Also required to obtain the county number for generating traffic count files compatible with the FDOT Statistics Office files.
- *Roadway_Milepost:* Required for correlation with RCI and crash data. Note that these fields are also in the station data file. They are required here because stations that detect traffic in both directions may have different roadways assigned.
- *Max_Speed:* Normally the speed limit. This information is required for travel time reliability reporting because the notion of on time arrival is based on the speed limit.
- *Count_Station:* The number assigned by the FDOT Statistics Office or District Planning Office for generating traffic count data files from the SunGuide detectors.

Steps in Configuring the TSS Data

The following steps are involved in configuring the TSS data for STEWARD:

- Develop a list of stations and lanes from sample TSS archive files. This is done automatically by a special utility program called TSSBuilder, which reads the archive files and compiles a list of all station and lane IDs found in the files.
- Assign each station to a facility or geographical subsystem. Each district may have up to 10 facilities, numbered 0 through 9. District personnel must carry out the facility assignment.
- Assign station ID numbers to each station. When the facility numbers have been assigned, the station number can be added sequentially. The order is not especially important as long as each number represents a unique station within the facility.
- Establish the position of each station on the facility. The coordinates, state milepost, RCI road number and county milepost must be determined. Most districts have this information compiled in separate records.
- Establish the station status. The station status is now assigned as (0 = normal, 1 = 0ffline). Other values might occur in the future.
- Assign lane ID numbers to each lane. This is the most detailed part of the process. The station number determined in Step 3 provides the first four characters of the lane ID. Three more characters are needed to complete the lane ID:
    - The direction of traffic in the lane (1 = increasing mileposts, 2 = decreasing mileposts)
    - The function of the lane:
      1 = left entrance ramp
      2 = left exit ramp
      3 = normal freeway mainline
      4 = right entrance ramp
      5 = right exit ramp

6 = auxiliary lane
7 = HOV lane

- The lane number (left to right)

- Assign lane detector operating parameters. The operating parameters for each detector include the status (same definition as the station status), the detector type (RTMS, Loop, etc) and the speed limit.

4.1.3 ETL Requirements

The ETL process must accept the raw archive data, combine it with the facility data that describes the properties of each detector in the system and load the combined data into the CDW database. Three summary intervals will be required:

- 5 minutes, for compatibility with the analysis of short term phenomena and perturbations
- 15 minutes, for compatibility with general traffic engineering analyses
- 60 minutes, for compatibility with statewide traffic counting program data

In addition, it will be necessary to summarize the data by 1-minute intervals to provide a resource for researchers. The 1-minute data will be stored on a separate medium and will not be included in the database.

4.2 Data Transfer Automation Requirements

Productive operation of the CDW requires that the daily archive files be transferred automatically from the TMCs to be loaded into the STEWARD Database. The key to this scheme is the implementation of a scheduled task by the districts in their SunGuide systems. An overview of the data flow for the required task that transfers the daily archives to an FTP site on the UF Campus and loads it to the STEWARD database is illustrated in Figure 11. Separate FTP sites have been established for each district.

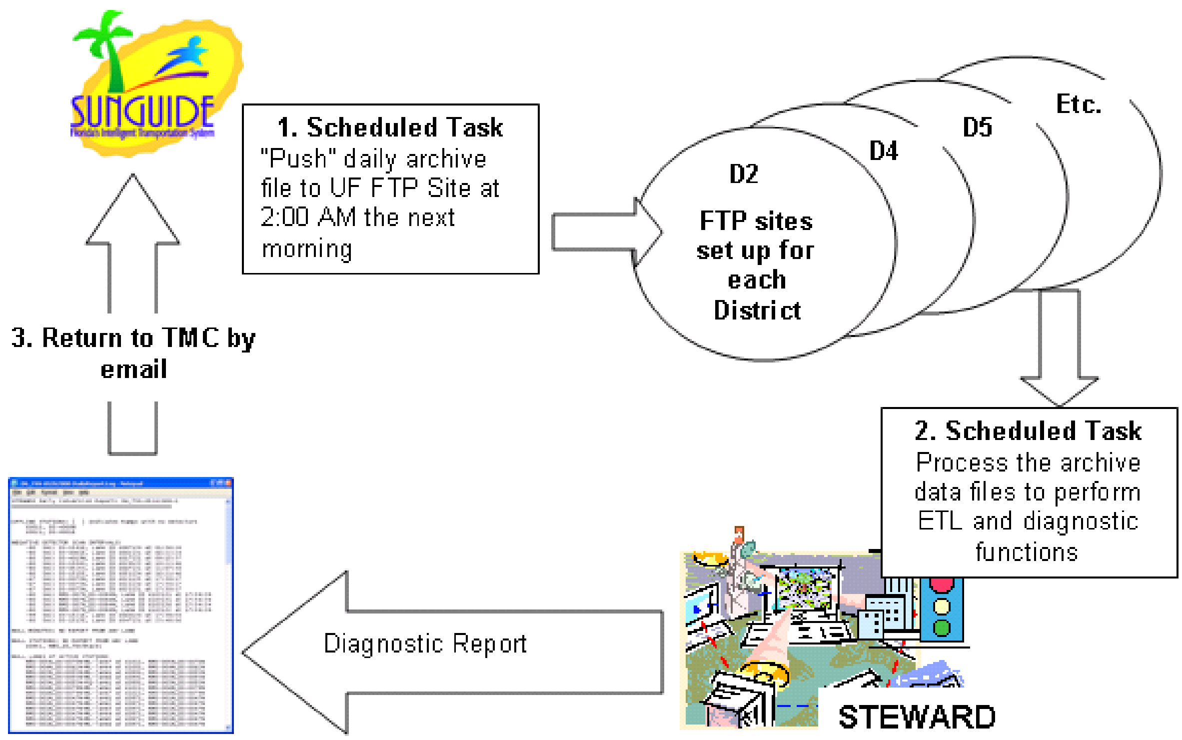


Figure 11. Automated data flow diagram for the SunGuide archive data

4.3 Reporting Requirements

The following reports will be required from the CDW. Each report represents commonly used performance measures that can be derived from the basic traffic flow theory relationships described earlier:

- *Vehicle Miles of Travel (VMT):* This is a measure of productivity of the freeway, typically accrued over a peak period or longer.
- *Vehicle Hours of Travel Time (VHTT):* This is the accumulated travel time of all vehicles in the system over the analysis period.
- *Average speed:* This is a measure representing the average speed of all vehicles in the system is computed as VMT / VHTT.
- *Delay:* There are several definitions of delay, each with its own method of computation. For a freeway system, the most appropriate delay measure is obtained by subtracting the VHTT that would have accrued at some desired speed from the measured VHTT. The result is expressed in vehicle hours of delay.
- *Kinetic Energy:* Kinetic energy is proportional to the product of speed and volume. Higher values of kinetic energy are obtained when heavy volumes are carried at high speeds. For this reason, kinetic energy has been suggested as the "bottom line" performance indicator for a freeway facility. It has also been suggested that high values of kinetic energy could be associated with safety hazards. This measure is produced to support future research.

Three performance measures derived from the travel times will be investigated in this report:

- *Congestion Delay:* based on a travel time index of 1.5. The travel time index is defined as the ratio of the actual travel time to the travel time at the free flow speed. The speed limit will be used to represent the free flow speed. The unit of measurement is accumulated minutes of delay.
- *On Time Delay:* referenced to a travel speed of 10 mph below the speed limit. This threshold has been specified for purposes of travel time reliability reporting in Florida. The unit of measurement is also accumulated minutes of delay.
- *Percent of on-time trips*: defined as the percent of trips made at a speed no less than 10 mph below the speed limit.

5 CENTRAL DATA WAREHOUSE DEVELOPMENT

This section summarizes the main features of the system that was developed under this project to meet the requirements outlined in the previous section. It provides a high-level overview of the system description, operation and Internet access features and the ETL process by which the archived data from regional traffic management centers is processed and incorporated into the STEWARD database. Considerably more detail on each of these topics is presented in technical appendices, which are published in a separate volume of this report.

5.1 STEWARD System Description

STEWARD consists of three main elements, including the ETL process, the database (DB) and the web-user interface. This section will describe the system architecture and implementation with respect to those elements.

STEWARD was developed using a variety of tools to design, deploy and maintain the system efficiently. The Oracle database was selected as a basic requirement from FDOT at the beginning of the project. The Windows 2003 Server and Microsoft Internet Information Services were selected as the operating system and web server, respectively. Based on this decision, the Oracle Warehouse Builder 10g2, Oracle Enterprise Manager and ASP/JavaScript were selected for the integrated ETL processes, the database management and the web development.

5.1.1 System Overview

Figure 12 shows the overall STEWARD configuration. The front end is the FTP server, which collects the traffic data from each district, processes it and archives it into the backup storage. The STEWARD DB server retrieves and loads these data files, which are then archived into the STEWARD database and used to update the materialized views. STEWARD users can access the data via the Web site or retrieve the data from the data back-up on request.

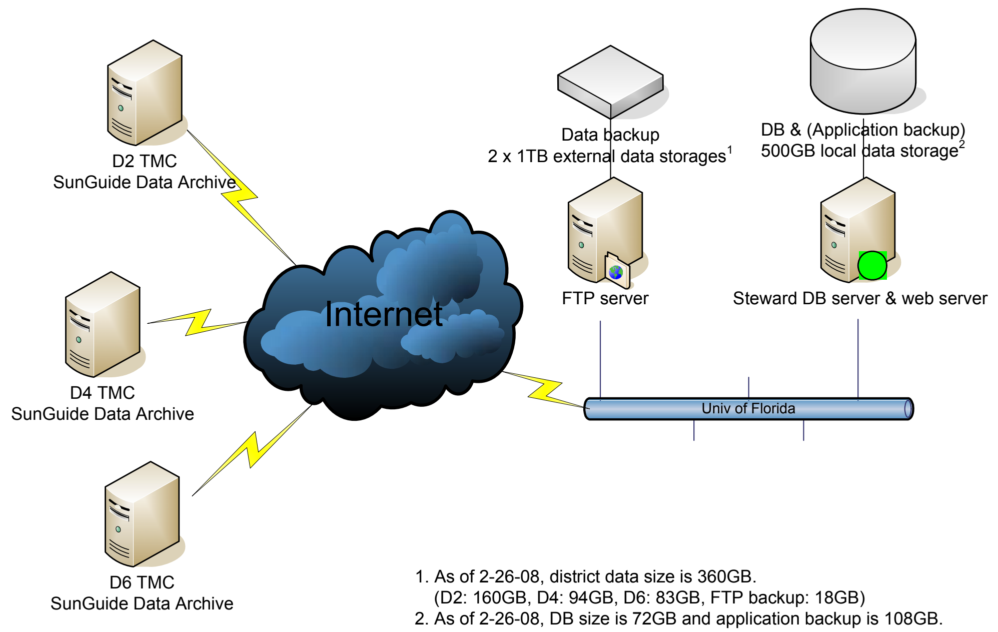


Figure 12. STEWARD system configuration

5.1.2 Database Design and Architecture

The STEWARD database design, development and management were carried out using the Oracle Warehouse Builder program, which is an integrated tool with a graphic user interface. This program includes predefined rules that are generally required in the warehouse design. The database consists of several types of tables, the functions of which are integrated by the database manager. The following table types are involved:

- External tables
- Dimensions (Dimension tables)
- Cubes (Fact tables)
- Materialized views
- Functions

The relationships between these tables are quite complex. The entire database schema is described in detail in Appendix 1.

5.1.3 Data Flow

Compressed (zipped) archive data files are obtained daily from each of the SunGuide TMCs by FTP file transfer. The required ETL functions previously summarized in Figure 11 are performed and the data are added to the STEWARD database. The database may then be queried through the Internet to select locations by station, section and facility to produce several reports that will be described later in this document.

5.2 STEWARD Operation

The STEWARD operation has been documented in detail in Appendix 2, which was developed for personnel who must install, operate and maintain the system. The topics include Oracle database program installation, STEWARD deployment, the STEWARD Web site installation and the STEWARD Web site management. A working knowledge of the Oracle database manager and Internet site management is assumed in the discussion contained in Appendix 2.

5.2.1 Oracle Database Program Installation

Oracle 10gR2 and the Oracle Workflow Server 2.6.4 are required for the STEWARD database installation. The Oracle Workflow Server is included in the Oracle Database 10g companion CD. Detailed instructions are provided for installing and configuring the following components:

- Oracle 10g Release 2
- Oracle Workflow
- Install Oracle Warehouse Builder
- Oracle Database Configuration Assistant
- Oracle Net Configuration Assistant
- Oracle Enterprise Manager

The instructions are presented as a step-by-step process with screen captures displayed to describe each step. A sample screen capture from the documentation of the installation process is included as Figure 13. A total of 139 screen images of this type are presented in Appendix 2 to guide the reader smoothly through the process.

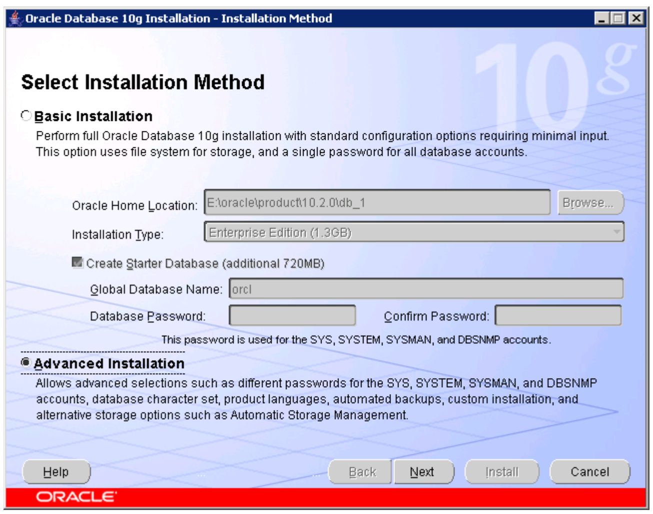


Figure 13: Example screen capture from the installation process

5.2.2 STEWARD Deployment

Deployment instructions are also presented in detail in Appendix 2. The topics Include:

- First step: Login
- Prerequisites: Create a target user
- Prerequisites: Uploading files
- Importing metadata
- Registration of the control center manager
- Data deployment process
- Data loading process

5.2.3 STEWARD Web Installation

To communicate with the Oracle database, the net configuration for Oracle data needs to be set up. This is accomplished through the Net Configuration Assistant for Oracle. A section of Appendix 2 guides the installer through the following steps:

- Net Configuration Assistant setup
- STEWARD web program installation
- System configuration
- Firewall setting
- Permission for file sharing
- Web program configuration

5.2.4 STEWARD Management

STEWARD receives archive data from SunGuide systems in each district every day. All data are processed and loaded into the STEWARD database for users to access the various reports. The process by which this operation is managed is detailed in Appendix 2. The topics include:

- Data transfer from district SunGuide systems
- Data backup and transformation in the STEWARD FTP server
- Data loading into the STEWARD database
- Refresh configuration for materialized views in the STEWARD database
- Updating the materialized views in the STEWARD database
- Backup plan and procedure for the STEWARD database and web
- Adding a new district or facility

5.3 Internet Access Features

While some use of STEWARD will be made within FDOT by accessing the databases directly, most users in the future will gain access to the archived data via the Internet. This section summarizes the Internet-based features of STEWARD from the perspective of a user who seeks to query the database and produce reports via the Internet. A more detailed description of the various reports and features and instructions for Web site use is provided in Appendix 3. The appendix material is also incorporated into a user manual, which is accessible from the Web site. The following topics are covered:

5.3.1 Overview of the STEWARD Web Interface

The STEWARD Web site has been developed for an audience of general users to access and retrieve the data. The web interface allows users to access the database remotely, to retrieve the specific data easily and to download the data to the local computer for further analysis. All data are downloaded in comma-delimited (CSV) format to facilitate presentation with office productivity software. At this time, the Web site can be accessed from the following Internet address:
http://cdwserver.ce.ufl.edu/steward/index.html

5.3.2 STEWARD Web Architecture

The STEWARD Web site consists of four main categories: overview, resources, maps and reports. The overall architecture is shown in the site map of Figure 14.

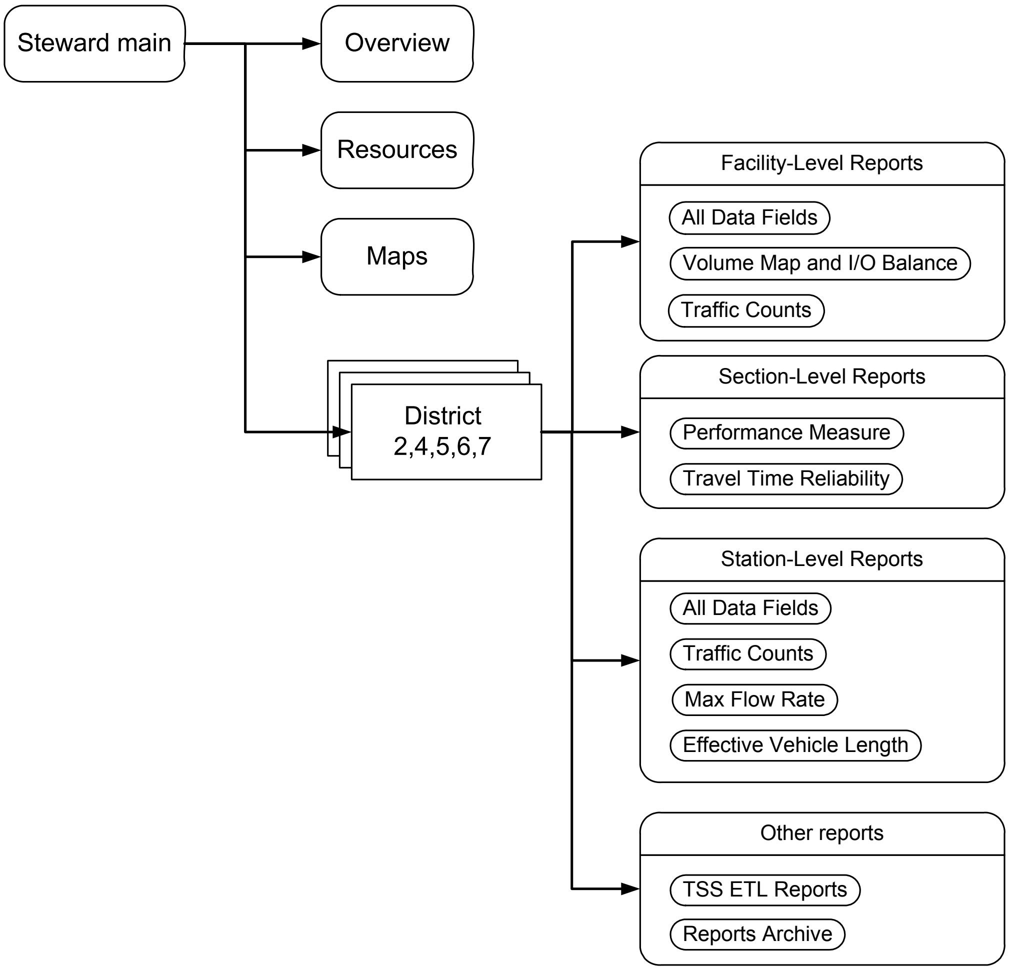


Figure 14: STEWARD web architecture

The overview page, which is shown in Figure 15, provides a general description of the STEWARD project. This page includes two panes as shown in the figure. The right pane displays a brief description of the STEWARD project, objectives and tasks. The left pane is used to navigate to the STEWARD overview, resources, maps and District data/reports sections.

The resource page provides access to reports, desktop utilities and traffic volume summaries. The report section includes the Phase II final report, progress reports and presentation materials. The utility section includes several utility programs: SunETLUtility, MPConverter, ITSCounts, SunVol, Hotter, SimTMC, and FTP Scripts. The traffic volume report section has links to the traffic volume reports for all the detectors in 2008. The utility programs are described in detail in Appendix 4.

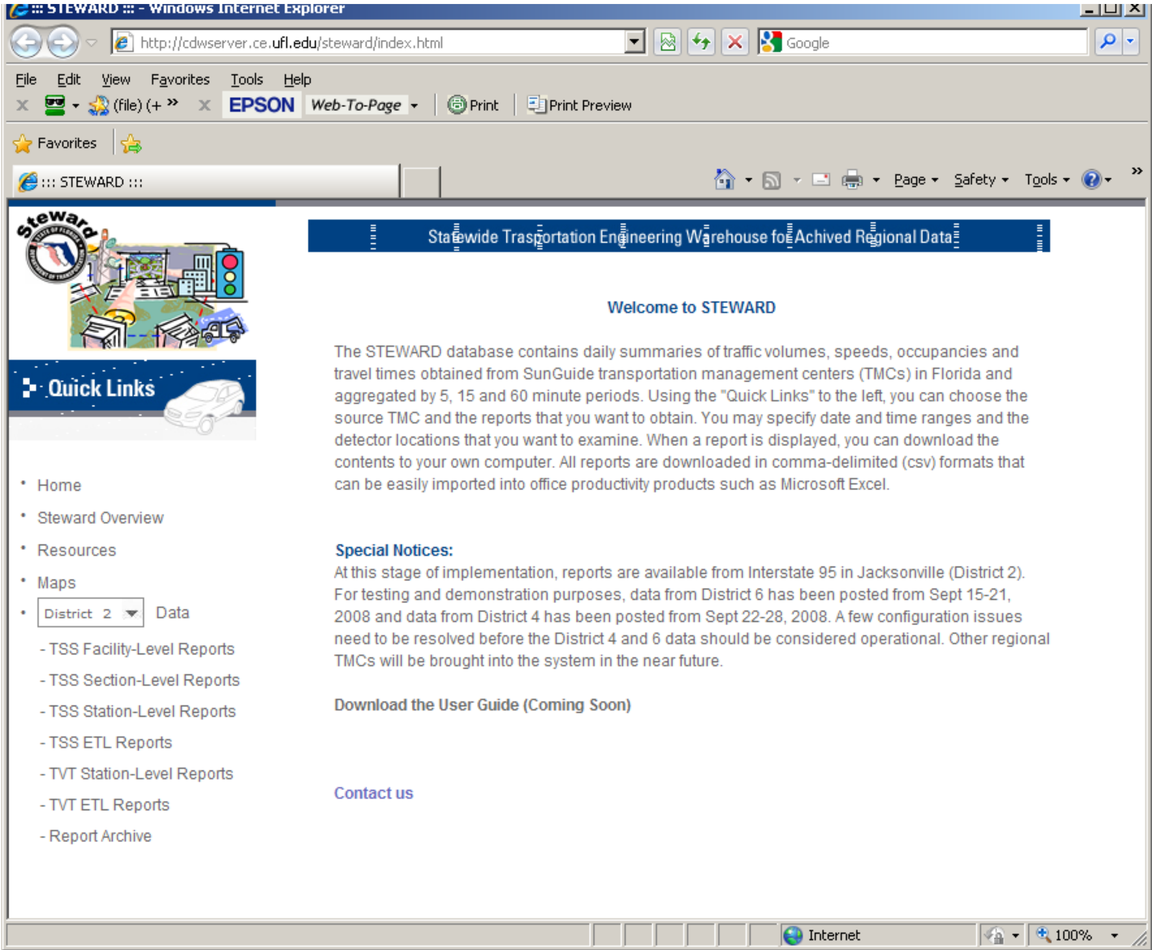


Figure 15. STEWARD overview page

5.3.3 *Maps*

Two graphics-based maps can be accessed for each district. The first, as shown in Figure 16, presents an interactive map superimposed on a Google Maps satellite photo. The second, as shown in Figure 17, presents an overview of the facilities in the district with detector locations shown on a GIS map.

5.3.4 *Report Levels*

Reports are available at the facility, section and station levels. Facility level reports apply to the entire facility, covering all stations. Section level reports apply to a user-defined section that includes all stations between a specified beginning and ending point. Station level reports apply to a single station. At all levels it is possible to specify the following selection criteria:

- The facility and direction within the district
- A date and time range
- Day of week or combination of days
- The desired aggregation level (five minutes, 15 minutes or one hour)

The selected report will be downloaded in coma delimited (CSV) format conforming to these selection criteria.

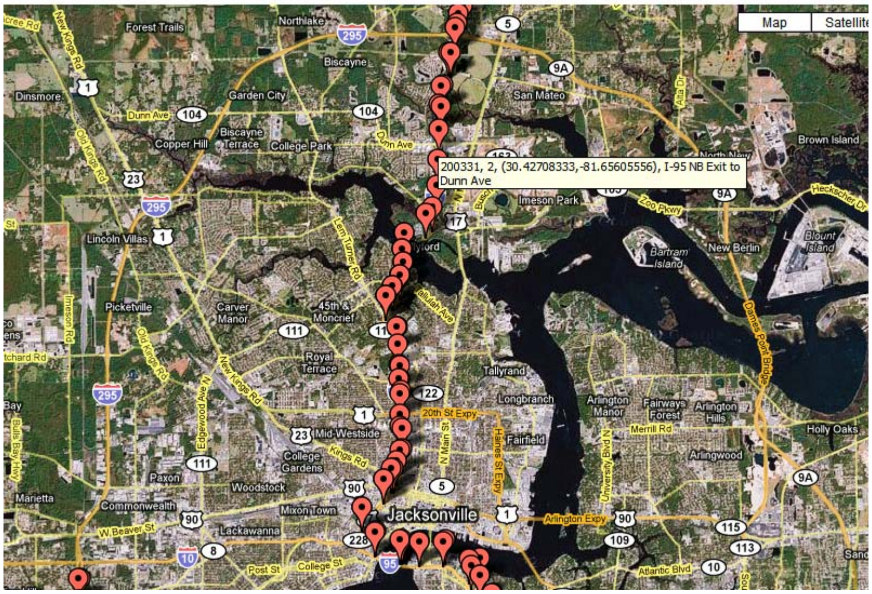


Figure 16. Example of an interactive satellite photo map

Figure 17. GIS map example

5.3.5 Report Types

Several report types are available from STEWARD:

- All Data Fields in the TSS Facility Level Report
- Volume Map and I/O Balance in the TSS Facility Level Report
- Traffic Counts in the TSS Facility Level Report
- Performance Measure in the TSS Section Level Report
- Travel Time Reliability in the TSS Section Level Report
- All Data Fields in TSS Station the Level Report
- Traffic Counts in TSS Station the Level Report
- Maximum Flow Rates in TSS Station Level Report
- Effective Vehicle Lengths in the TSS Station Level Report

Appendix 3 describes the content of these reports in detail and provides examples of specific report selection.

5.4 ETL Operations

The ETL process must accept the raw archive data, combine it with the facility data that describes the properties of each detector in the system and load the combined data into the CDW database. All of these operations are accomplished by a specially developed utility program, called Sun ETL Utility. This program is described in Appendix 4. Figure 18 illustrates the flow of data involved in the ETL process. The elements of Figure 18 focus on the ETL Utility program, which uses two types of data input:

1. The SunGuide archive data files, which are received as raw data input by the ETL Utility. These files are eventually discarded from the ETL process. They are kept as raw data on separate media to be furnished to researchers who require the raw data.
2. The facility data files, which are developed as a part of the configuration process for each facility. The facility configuration process was described in Section 4.1.2.

The ETL Utility program produces three output files:

1. Daily reports, which may be used by the facility operators to assess problems with the detector system. A sample of a daily report is presented in Figure 19. The following terminology applies to each system element (lane, station or ramp):
    - "Offline" identifies elements that are not currently functioning, usually due to construction.
    - "Null" refers to an element that is present in the configuration data but reported no volumes during the entire day. Null minutes indicate intervals in which no report was received from any system element, suggesting that the system was down during that interval.
    - "Orphan" refers to an element that reported data but was not identified in the configuration file. Orphan elements are usually the result of new additions or misspellings in the configuration file. Elimination of orphans is an important maintenance task.

2. Detector data, which is summarized by five, 15 and 60 minute periods for each detected lane in the facility.
3. Station data, which is accumulated from the detectors assigned to each station on the facility. A station consists of one or more detected lanes that carry traffic in the same direction on the same roadway.

The station data for each day of operation are loaded into the STEWARD database. A combination of the station and lane data is used to produce the QA reports described elsewhere in this document.

5.5 Current Status of the System

STEWARD receives TSS archive data from District 2, 4, 5, 6 and 7 daily from 1,200 stations. Most detectors are radar/video detector types, which cover up to eight lanes at one location. STEWARD stations cover approximately 4,200 lanes.

Table 8 shows information on the facilities and detector stations as of 9-30-2009.

Table 8. Status of STEWARD Facilities and Stations

District	Facility	Number of Stations
2	I-95, I-295	192
4	I-95, I-595, SR869	334
5	I-4, I-95	452
6	I-75, I-95, I-195, SR-826, US-1	233
7	I-4, I-275	150

Table 9 shows the data available from STEWARD for each district as of Sept. 30, 2009.

Table 9: TSS Data Availability in STEWARD

District	Data available on the STEWARD Web site
2	6-28-07 to current
4	5-1-08 to current
5	4-7-09 to current
6	5-26-08 to current
7	1-8-09 to current

STEWARD provides traffic data and reports through its Web site. Table 10 shows visitor statistics for three months. Average visitors were more that 60 per day and more than 240 different computers per month accessed the STEWARD web pages. The principal users are located in Gainesville, Miami and Tallahassee, Fla.

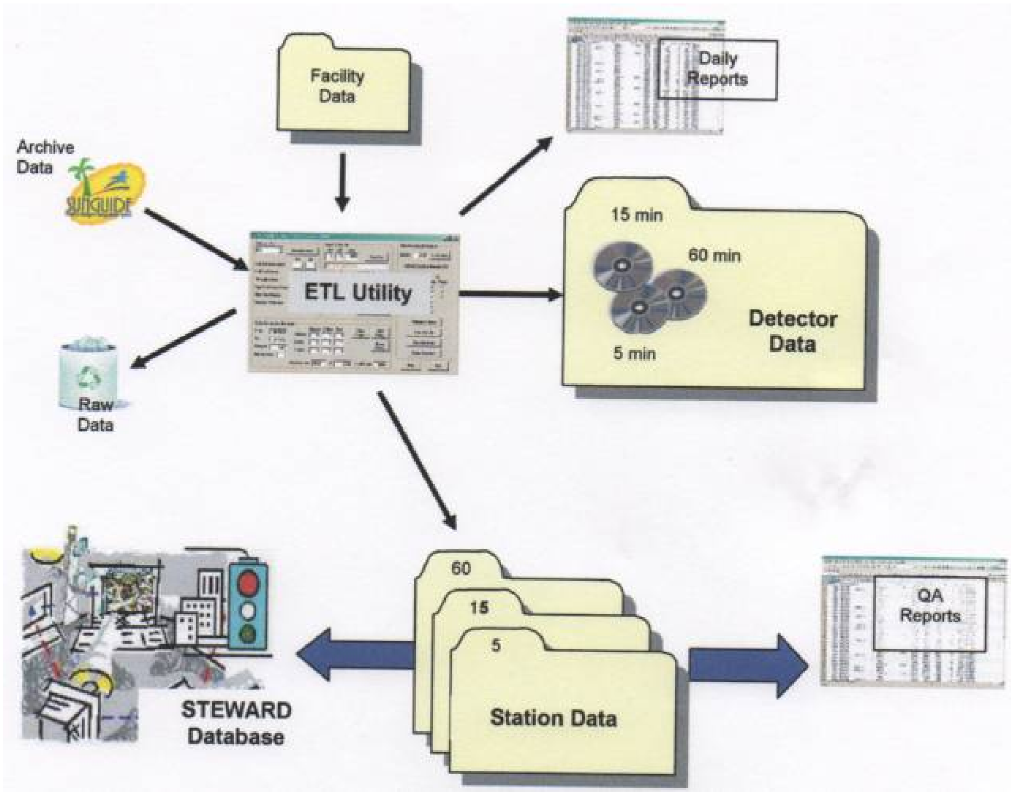


Figure 18: ETL Utility data flow

```
00:00:39   RTMS 95N006: Ramp R95N006_01Ramp_01 is not in the TSS Map.
00:00:39   RTMS 95N024: Lane R95N024_05LaneS-02 is not in the TSS Map.
Offline Station #  1 : RTMS 95N001, I-95 NB at Forest St
Offline Lane: R95N001_01Lane_01 at RTMS 95N001, I-95 NB at Forest St
Offline Lane: R95N001_02Lane_02 at RTMS 95N001, I-95 NB at Forest St
Offline Lane: R95N001_03Lane_03 at RTMS 95N001, I-95 NB at Forest St
Offline Lane: R95N001_04Lane_04 at RTMS 95N001, I-95 NB at Forest St
Offline Station #  2 : RTMS 95N002, I-95 SB at Forest St
Offline Lane: R95N002_01Lane_01 at RTMS 95N002, I-95 SB at Forest St
Offline Lane: R95N002_02Lane_02 at RTMS 95N002, I-95 SB at Forest St
Offline Lane: R95N002_03Lane_03 at RTMS 95N002, I-95 SB at Forest St
Offline Lane: R95N002_04Lane_04 at RTMS 95N002, I-95 SB at Forest St
Offline Lane: R95N002_05Lane_05 at RTMS 95N002, I-95 SB at Forest St
Null Ramp: R95N014A_01Ramp_01 at #  14 , RTMS 95N014, I-95 NB Exit to MLK Blvd
Null Ramp: R95N022A_01Ramp_01 at #  22 , RTMS 95N022, I-95 NB Exit to Lem Turner Blvd
Null Lane: R95N024_05LaneS_02 at #  24 , RTMS 95N024, I-95 NB Entrance from Lem Turner Blvd
Null Ramp: R95S054A_01Ramp_01 at #  104 , RTMS 95S054, I-95 SB Entrance from Baymeadows Rd
Null Ramp: R95S054A_02Ramp_02 at #  104 , RTMS 95S054, I-95 SB Entrance from Baymeadows Rd

Conversion Summary for D2_TSS-02152006-1
From 00:00:39
To 23:59:39
Elapsed Time 1440
Null Intervals  0
Orphan  Stations  0
      Lanes     1
      Ramps     1
 Null  Stations  0
      Lanes     1
      Ramps     4
Offline Stations  2
      Lanes     9
      Ramps     0
 Max flow rate    5400
         at    5400
 No. > 2400vph    8098
```

Figure 19: Sample daily report from the ETL process

Table 10: STEWARD Web Statistics for June, July and August 2009

		Jun. 2009	*Jul. 2009*	***Aug. 2009***
Visitors	Total Visitors	2,176	2,572	1,452
	Average Visitors per Day	70	80	64
	Total Unique IPs	243	280	293
	IP location of top downloader	Gainesville, FL (17GB)	Miami, FL (23GB)	Tallahassee, FL (72GB)
Page Views	Total Page Views	10,784	12,566	11,525
	Average Page Views per Visitor	4.96	4.89	5.57
Bandwidth	Total Bandwidth	41.30 GB	119.17 GB	172.90 GB
	Average Bandwidth per Day	1.33 GB	3.72 GB	5.40 GB
	Average Bandwidth per Hit	1.19 MB	3.66 MB	5.42 MB
	Average Bandwidth per Visitor:	19.44 MB	47.45 MB	85.13 MB
Hits	Total Hits	35,570	33,337	32,651
	Average Hits per Day	1,147	1,041	1,020
	Average Hits per Visitor	16.35	12.96	15.77

6 DEVELOPMENT OF QUALITY ASSURANCE PROCEDURES

The literature review presented in Section 2 includes a description of the quality assurance (QA) procedures that are commonly applied in U.S. cities. This section will apply these procedures to the data in the STEWARD archives. It will also explore the potential for expanding the procedures to include additional QA tests and rules that are made possible by the comprehensive nature of the STEWARD data. The current QA procedures are applied to each detected lane. Because the STEWARD database is organized geographically, it is possible to create additional tests that examine the consistency of data among the individual lanes at a station and among contiguous stations along a facility.

These are four levels of quality assurance procedures for STEWARD:

- Level 1: Completeness test
- Level 2: Data validity test
- Level 3: Station level tests
- Level 4: System level tests

Each category focuses on different aspects of the traffic data. The first two levels apply the current QA procedures to individual lanes. The Level 1 data completeness test checks the detector malfunctions, communication failures, archive errors, etc. The Level 2 validity test checks that the traffic data are within the operational data range such as the maximum or minimum allowable values.

The last two levels examine consistency among groups of lanes. The Level 3 station data validation examines the variation between traffic conditions in the lanes that comprise a station. The Level 4 system data validation examines the variation between traffic conditions between adjacent stations. For these two additional levels, the measures that should be expected to show consistency are identified and examples of consistent and inconsistent data are presented.

6.1 Level 1 Completeness Test

The Level 1 completeness test verifies that the traffic data collection, transfer and archiving system is functioning properly. Therefore, it will identify the system hardware or data communication issues observed in the archived data. For example, the traffic data are produced from the freeway detectors every 20 or 30 seconds and delivered into the district SunGuide systems. Within the SunGuide systems, these data are processed and archived into the database. The daily traffic data items are retrieved, formatted and transferred into STEWARD. During these operations, several problems could arise:

- Detector malfunctions
- Communication errors between detectors and SunGuide system
- Data processing errors, such as duplicate traffic data on the same timestamp at the same location
- Communication errors between SunGuide and STEWARD

To verify these problems, following items will be tested and verified:
- Availability of the district data
- Missing detector data
- All-zero or stuck detector data
- Duplicate or negative-scan data

6.1.1 *Availability of the District Data*

The availability check for the district data is simple but critical to the overall system performance. STEWARD receives the traffic data everyday from each district. This test will verify that STEWARD receives valid data files from each district on time. If the data files are not available, all data for that day would be missing for that district.

The district data availability for STEWARD is shown here: Three, ten and 14 days of traffic data are missing from Districts 2, 4 and 6 during the second half of 2008.

Availability of the District 2,4, and 6 data (7/1/08~12/31/08)

	Number of Missing Days	Percent Unavailable
District 2	3	3/184 = 1.6%
District 4	10	10/184 = 5.4%
District 6	14	14/184 = 7.6%

In addition, the data file for Dec. 2, 2008 from District 4 had a mechanical error in the archived file format and was therefore rejected. This case would be another example of missing dates from the district.

6.1.2 *Missing Detector Data*

The availability test for the detector data checks all of the detector data in the traffic data file. If there are any problems in detector or communication errors during the data transfer, all of the detector data would not be available. This test covers all of the traffic data that are collected and archived.

Completeness is defined as *the degree to which data values are present in the attributes that require them*. This is a percentage value calculated from the available number of data values as a percent of the number of total expected data values.

$$Percent\ Complete\ (\%) = \frac{n_{available values}}{n_{total expected}} \times 100$$

Where

$n_{\text{available values}}$ = the number of records or rows with available values present
$n_{\text{total expected}}$ = the total number of records or rows expected

In this calculation, completeness is defined to verify the availability rather than the validity.

The completeness of traffic data from District 2, District 4 and District 6 were examined for the month of October 2008. The results are as follows:

$$\text{Completeness}_{\text{District 2}} = \text{(total number of traffic data values during the period)} / \text{(total number of detected lanes * 31 days * (24 hours / 20 sec))}$$
$$= 44{,}696{,}974 / (681 * 31 * (24 * 60 * 60) / 20)$$
$$= 0.49$$

$$\text{Completeness}_{\text{District 4}} = 70{,}357{,}757/75{,}129{,}120 = 0.94$$

$$\text{Completeness}_{\text{District 6}} = 56{,}819{,}105/92{,}404{,}800 = 0.61$$

If a station produces no data (including all zero-volume data) for a day, it is defined as a *null station* for that day. In District 2, an average of 74 out of 190 stations were null stations in October 2008. Null stations (74/190 = 38.9 percent) impact the completeness of District 2 data. In District 6, 54 out of 233 stations are null stations. Null stations in these districts are the result of the system implementation schedule. They should eventually be eliminated as the system implementation progresses.

If a lane produces no data (including all zero-volume data) for a day, it is defined as a *null lane* for that day. The occurrence of null lanes at non-null stations was minimal in all districts. For the period examined, it is clear that there were problems in the District 2 data collection and archiving systems. These problems have been resolved since the period of the analysis.

6.1.3 All Zero or Stuck Detectors

This data test checks the variation of the data values for the traffic detector during a time period. These data could be all-zero or one-fixed value for a time period. The time periods for the all-zero or stuck data test are suggested as 8 consecutive identical values from an FHWA report on monitoring urban freeways [11].

A threshold of five consecutive minutes has been used in this study to check if the detector is in the all-zero or stuck condition. The sampling rates for District 2, 4, and 6 are 20-second, so 15 data observations are tested during each five-minute period. These rules were applied to detector data from Districts 2, 4 and 6 for the month of October 2008. The results were as follows:

	Number of all zero data (V/S/O) for 5 min (1)	Number of stuck data (V/S/O) for 5 min (2)	Total expected number of traffic for 5 min (3)	Percent ratio (1+2)/3
District 2	45,254	230	6,079,968	0.8%
District 4	24,478	250	5,008,608	0.5%
District 6	144,222	2241	6,160,320	2.4%

Most of the detectors in Districts 2, 4 and 6 are RTMS detectors. Given the limited scope of the data collection and the fact that the systems were in various stages of implementation, it is difficult to draw general conclusions on the relative significance of these results. Most detectors exhibit all-zero problems rather than stuck detector problems.

6.1.4 Duplicate or Negative-Scan Data

Duplicate (zero-scan) or negative-scan data problems are specific issues in the SunGuide system. Duplicate-data records are defined by multiple records from the same detector with the same time stamp. Therefore, two or more detector records are archived in the traffic data file with the same detector/lane ID with no time intervals (zero scan intervals). The SunGuide data archive system is designed to log the traffic data into the file in chronological order. Negative scan intervals are defined by records in which a time stamp indicates an earlier time than the preceding record. District 2 and District 6 had these problems at one point and more than 10 percent of the daily traffic data records were reported with zero or negative scan intervals. It appears that the SunGuide contractor has resolved these problems. The suspicious data were not archived into the STEWARD database.

Duplicate or Negative Scan Records in Districts 2, 4, and 6 (10/1/08~10/30/08)

	Duplicate or negative-scan data	Total traffic data	Percent ratio
District 2	571	5,681,9105	0.001%
District 4	4,091	70,357,757	0.006%
District 6	489,779	56,819,105	0.9%

6.1.5 Level 1 Test Summary

The level 1 test focuses on the system operation related traffic data. As STEWARD is not involved in the data generation, communication, archive and transfer, it needs to verify that the delivered traffic data files have all the expected data items with predefined formats. The causes and results at this level are summarized as follows:

Causes	Results
Detector malfunctions Detector turn-off/lane closure Detector-SunGuide communication problems	No detector data All-zero or Stuck data values
SunGuide data archiving and retrieving problems	Duplicate or negative time scan Parity error on the archive files
Delivery problems	Missing all district data for one or more days

Most of problems in this level are mechanical in nature could be resolved by each district.

6.2 Level 2 Data Validity Test

Data validity tests in this level check that traffic data are in an acceptable operational range. Most of the quality control methods offered in the FHWA report on monitoring urban freeways were used for the test.

Eight validation rules were set up and applied to the District 2, 4 and 6 data. The rules are based on the following measures:

- Maximum volume
- Maximum occupancy
- Maximum speed
- Multivariate consistency (zero speed with non-zero volume)
- Multivariate consistency (zero volume with non-zero speed)
- Multivariate consistency (zero volume, zero speed with non-zero occupancy)
- Truncated occupancy values of zero
- Maximum estimated density

6.2.1 Maximum Volume Test

Maximum allowable volume for each lane data should be less than

- 17 vehicles for 20-second data
- 25 vehicles for 30-second data
- 250 vehicles for five-minute data
- 3000 vehicles for 1-hour data

This rule was applied to the 20-second data from District 2, 4 and 6 during October 2008. The percent of total observations that failed the test in each district was as follows:

	District2	District4	District6
Maximum volume	0.13%	0.28%	0.66%

Note that less than 1 percent of the observations failed this test.

6.2.2 Maximum Occupancy Test

Maximum allowable occupancy for each lane data should be less than

- 95 percent for 20- or 30-second data
- 80 percent for one- to 5-minute data

This rule was applied to the 20-second data from District 2, 4 and 6 during October 2008. The percent of total observations that failed the test in each district was as follows:

	District2	District4	District6
Maximum occupancy	0.00%	0.01%	0.25%

Note that less than 1 percent of the observations failed this test. It is also observed that District 2 had a zero failure rate.

6.2.3 Minimum and Maximum Speed Tests

The minimum allowable speed for each lane data should be higher than 5 mph and the maximum allowable speed for each lane data should be less than

- 100mph for 20- or 30-second data
- 80mph for one- to 5-minute data

These rules were applied to the 20-second data from District 2, 4 and 6 during October 2008. The percent of total observations that failed the test in each district was as follows:

	District2	District4	District6
Minimum speed	0.03%	0.02%	3.32%
Maximum speed	0.00%	0.00%	0.05%

In most cases, the failure rates were below 0.05 percent. The only exception is the minimum speed test in District 6, which showed a failure rate of 3.32 percent. This suggests that some attention to the calibration of certain detectors in District 6 might be desirable.

6.2.4 Multivariate Consistency Test

There are three cases that are related with multiple variables:

- Zero speed and non-zero volume case
- Zero volume and non-zero speed case
- Zero speed an volume and non-zero occupancy case

These rules were applied to 20-second data from District 2, 4 and 6 during Oct. 2008. The percent of total observations that failed the test in each district was as follows:

	District2	District4	District6
Zero speed and non-zero volume	0.01%	0.02%	3.22%
Zero volume and non-zero speed case	0.01%	10.24%	3.43%
Zero speed an volume and non-zero occupancy case	0.01%	0.00%	0.44%

It was observed that District 4 detectors tended to generate non-zero speed values with zero volume when the traffic volumes are very low in early morning. The reason for this is not known.

6.2.5 Truncated Occupancy Values Test

Older detectors on the roadway have a lower resolution in occupancy data. With very low volumes, the occupancy might fall below 1 percent and be truncated to a zero value. This test was applied to 20-second data from District 2, 4 and 6 during Oct. 2008. The percent of total observations that failed the test in each district was as follows:

	District2	District4	District6
Zero occupancy and non-zero volume	0.00%	0.01%	0.03%

Failure rates for this test were very low.

6.2.6 Maximum Estimated Density

For this test, the maximum allowable density for each lane data should be less than:

Estimated density ((VOLUME*(3600/NOM_POLL))/SPEED) < 220

This test was applied to the 20-second data from District 2, 4 and 6 during October 2008. The percent of total observations that failed the test in each district was as follows:

	District2	District4	District6
Maximum estimated density	0.05%	0.00%	0.18%

6.2.7 *Summary of Level 2 Tests*

These eight criteria were implemented in the STEWARD ETL processing. Error codes are defined for these criteria as follows. If one of the data items fails two or more criteria, the sum of the error codes will be recorded. Data that fail any of these criteria are marked with the appropriate error code and are not used in the aggregated STEWARD reports.

Error type	Error code
Maximum volume	1
Maximum occupancy	2
Minimum speed	4
Maximum speed	8
Zero speed and non-zero volume	16
Zero volume and non-zero speed case	32
Zero speed an volume and non-zero occupancy case	64
Zero occupancy and non-zero volume	128
Maximum estimated density	256

The results of applying the eight criteria to the data from District 2, 4 and 6 during Oct. 2008 are summarized here. The percentage of records that failed one or more of the eight quality check criteria is summarized as follows:

	District2	District4	District6
Eight QC rules	0.24%	10.58%	11.58%

6.3 Level 3 Station Data Validation

The Level 1 and Level 2 tests are applied to the data from individual lanes to identify problems with the detectors. The Level 3 station data validation is applied to the aggregated data for all lanes at the station. This procedure uses traffic flow principles to identify inconsistencies among the lane-specific data. The relationship between volume, speed and occupancy and other performance measures from the following STEWARD reports are used for this purpose:

- Maximum flow rates for the station
- Effective vehicle length (EVL)
- Lane balance
- Daily volume variation
- Annual volume variation

To set up the criteria, a set of stations was chosen with locations and time limits that were known to be free of reported problems associated with system malfunctions, construction, etc. that would generally be detected by the lower level tests. The test sample had the following characteristics:

- Facility: District 2 I-95 northbound, south of I-10
- Date: October 2008, weekdays (23 days)
- Time: Morning peak (7 a.m. — 10 a.m.) rush-hour traffic

The following four stations were selected from 23 active stations in this area:

- I-95 NB South of Butler Blvd
- I-95 NB North of Baymeadows Rd
- I-95 NB Entrance from Baymeadows Rd
- I-95 NB Entrance from Philips Hwy

6.3.1 Maximum Flow Rates

Maximum flow rates can be used to identify the stations that produce excessive traffic volumes and therefore might need calibration or other maintenance attention. STEWARD provides the maximum flow rate per day from all stations. These rates are calculated from five-minute, 15-minute or one-hour traffic volumes. Excessive maximum flow rates could be the result of detector overcounting. In the Level 2 tests, the threshold for maximum flow rate from an individual lane was set at 3,000 veh/ln/h. For purposes of this test, a station level threshold value of 2,400 veh/ln/h, which is more in line with the HCM capacity estimates, will be used.

Flow rate histograms for the four selected stations were created as shown in Figure 20 from 15-minute traffic volume data. From the cumulative percentages, 99.3 percent of flow rates are less than 2,400 veh/ln/h, which was selected as the threshold criterion.

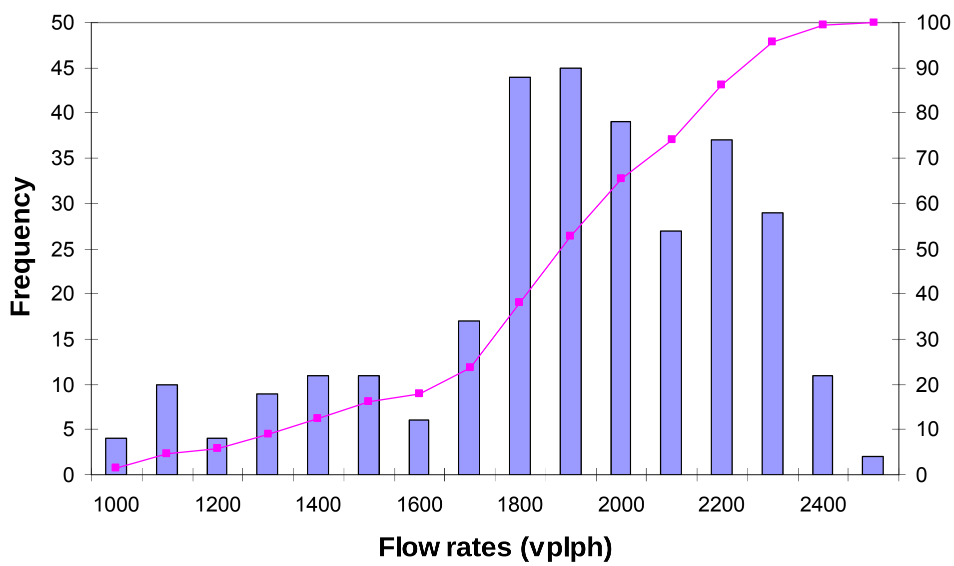


Figure 20. Flow rate histogram from four selected stations

To examine a more comprehensive sample, maximum flow rate reports were created from the STEWARD web site by applying the following selection criteria:

- District 2, I-95
- Northbound and southbound between I-10 and I-295
- All weekdays in Oct. 2008

Table 11 shows the frequency of maximum flow rates from this data report. It is observed that 95.2% of the observations have less than 2400 veh/ln/h. The other 5% are producing volumes that exceed the upper limits of capacity indicated by the HCM procedures for analysis of basic freeway segments. While it is not impossible that higher flow rates could occur, there might be a need to examine the operation of these stations in more detail.

Maximum flow rates	Frequency	Cumulative %
100	18	0.36%
200	6	0.48%
300	38	1.23%
400	53	2.29%
500	65	3.58%
600	125	6.07%
700	93	7.92%
800	46	8.83%
900	67	10.17%
1,000	94	12.04%
1,100	118	14.39%
1,200	128	16.93%
1,300	129	19.50%
1,400	229	24.05%
1,500	247	28.97%
1,600	267	34.28%
1,700	346	41.17%
1,800	378	48.69%
1,900	414	56.92%
2,000	521	67.29%
2,100	515	77.54%
2,200	378	85.06%
2,300	287	90.77%
2,400	225	95.24%
2,500	98	97.19%
2,600	41	98.01%
2,700	26	98.53%
2,800	20	98.93%
2,900	21	99.34%
3,000	16	99.66%
3,100	8	99.82%
3,200	3	99.88%
3,300	4	99.96%
3,400	2	100.00%

Table 11. Maximum Flow Rate Frequencies

The flow-rate observations in excess of 2,400 veh/ln/h were examined in more detail to determine whether the threshold value for this test should be raised, keeping in mind that the value of 2,400 represents the typical capacity of a freeway segment reported in the HCM. It could be argued that observations that exceed a typical value would not necessarily indicate a data quality problem.

As indicated in the cumulative distribution plot of Figure 21, the observations that exceeded 2,400 veh/ln/h were all below 3,400 veh/ln/h. The cumulative distribution reached the 99 percent level at a flow rate of approximately 2,900 veh/ln/h. Raising the threshold to 2,900, which is approximately 20 percent higher than the HCM's value of 2,400, will increase the usefulness of the maximum flow rate test without an unreasonable departure from the HCM results. Therefore a value of 2,900 veh/ln/h will be applied.

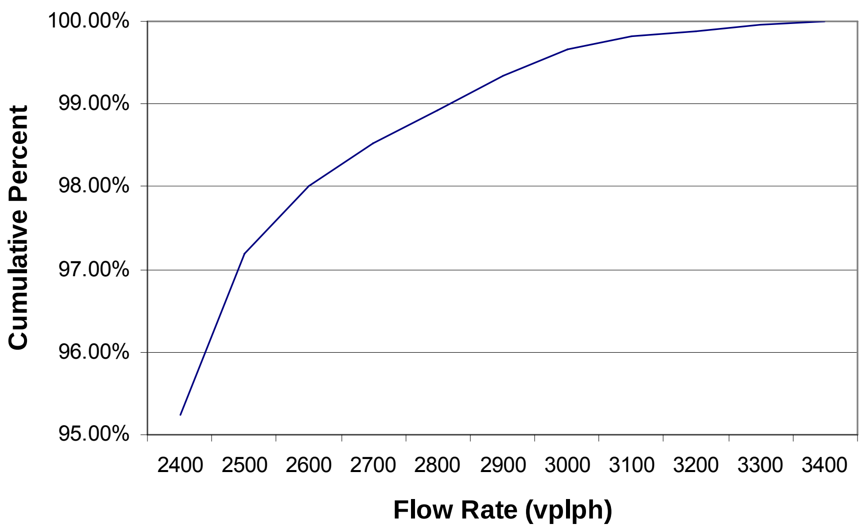


Figure 21. Cumulative percentage more than 2,400 veh/ln/h from I-95 stations

From Table 11, the maximum flow rates exceed 2,400vplph in 464 cases but they were produced from 39 lanes or 27 stations. This suggests that these stations produce excessive maximum flow rates repeatedly for this time period.

6.3.2 Speed Flow Relationships

The basic speed-flow relationship was presented previously in this report in connection with the principles of traffic flow discussed Section 3. Figure 6 in that section illustrated the traditional shape of the speed-flow curve. Detector stations that are functioning properly would be expected to produce speed-flow relationships that generally conform to the same shape. Most of the STEWARD detectors conformed generally to the basic traffic flow relationships but some still appear to over count the volume. Station 210371 is an example. This station produced 46 occurrences of maximum flow rates over 2,400 vphpl. Figure 22, Figure 23 and Figure 24 show the speed-flow rate relationships for the three lanes of this station. All of the relationships appear to be normal, suggesting the need to check the detectors for over counting.

It also noted that the over counting appears to take place in Lane 1, as evident from the peak volumes of 3,500 vphpl. Lanes 2 and 3 show smaller peaks in the range of 2,300 to 2,600 vphpl. This example shows how shape of the speed-flow relationship can be used to provide more insight into the operation of a detector in any given lane.

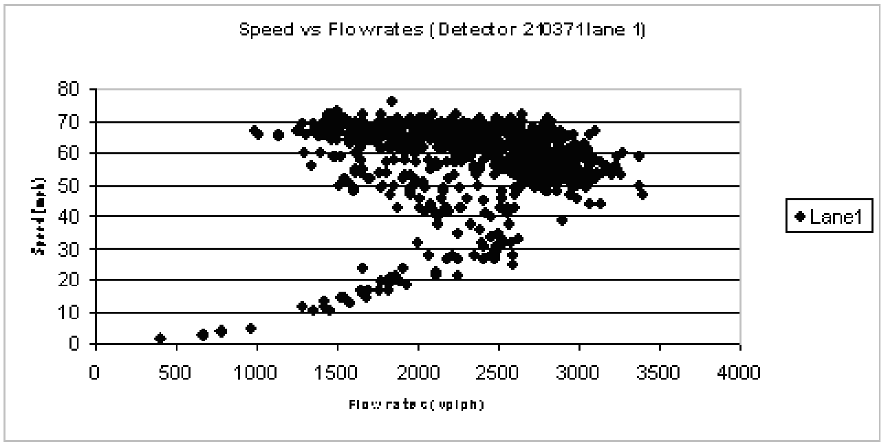


Figure 22. Speed vs Flow rates at station 210371, Lane 1

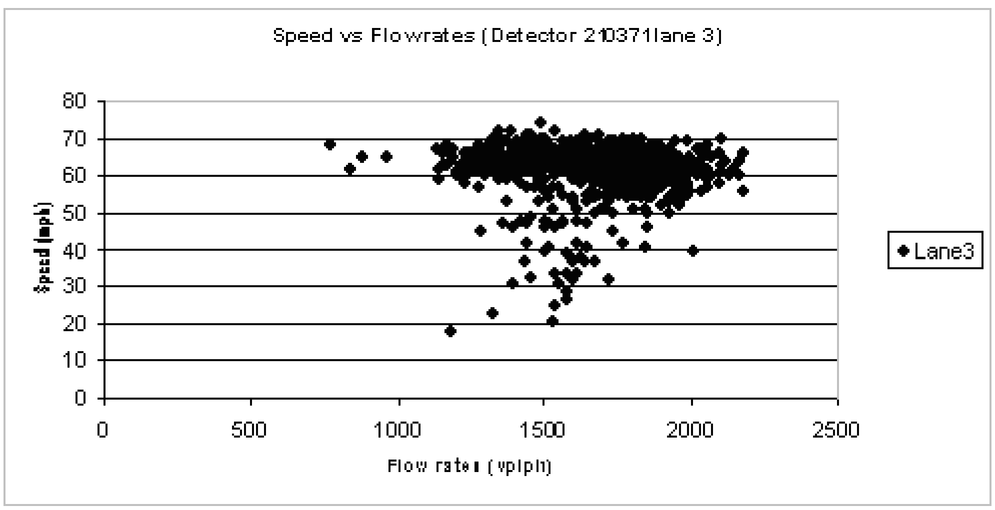


Figure 23. Speed *vs*. Flow rates at station 210371, Lane 2

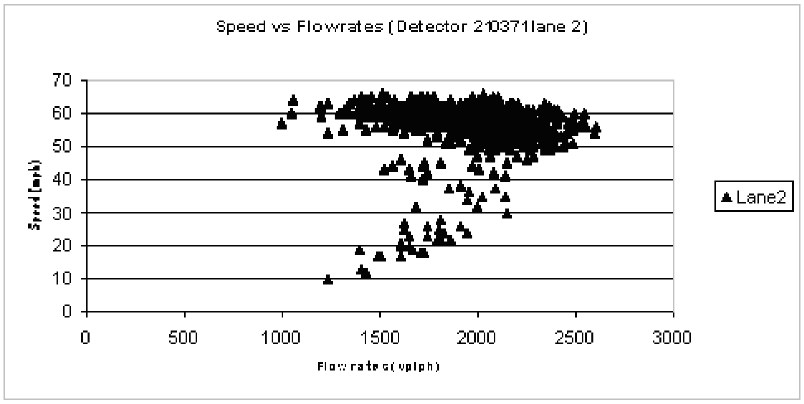


Figure 24. Speed vs Flow rates at station 210371, Lane 3

6.3.3 Effective Vehicle Lengths

The effective vehicle length (EVL) is defined as the length of the vehicle plus the length of the detection zone because a vehicle will be detected as long as it any part of it remains within the detection zone. While the EVL can be calculated at the individual lane level (i.e., Level 2), it has been considered as a Level 3 characteristic for purposes of this discussion because it has not been included in the Level 2 tests described in the literature.

The EVL is calculated using the relationship from the following basic traffic measures. There are three parameters that are related by a simple equation:

$$Q=K*U$$

$$\text{Occupancy} = K * (EVL)$$

$$U \approx V$$

Where Q = Flow rate (Vehicles per hour)
K = Density (Vehicles per mile)
U = Space mean speed (Miles per hour)
V = Time mean speed (Miles per hour)

Then

$$\text{(Effective vehicle length)} = V * \text{Occupancy} / Q$$

The first equation is a simple flow-density-speed relationship. The second equation uses the assumption that the speeds are constant over the link, which makes the space mean speed the same as the measured time mean speed. The third equation comes from the density definition and previous two equations.

From this relationship, the EVLs were estimated using the four selected detectors. Figure 25 shows the EVLs for Station 210471. The EVL shows relatively constant values at medium to high flow rates. But as the flow rate decreases below this level, high EVLs begin to occur.

Some of the higher values are associated with congested operation, with low flow rates and high density (occupancy). Under this condition, successive vehicles can be counted as a single long vehicle.

The average EVLs for the four selected stations are calculated as follows. During the calculation, high occupancy data (occupancy >18 percent) were excluded to avoid the oversaturated region.

- Station 210471: 21.4ft
- Station 210511: 21.5ft
- Station 210513: 20.1ft
- Station 210513: 24.2ft

The average EVL should be in the order of 21 feet, which is the sum of average passenger vehicle length (15ft) and average detector length (6 feet). Note that determination of an EVL requires a reasonable sample size. EVLs computed for individual polling intervals tend to be very erratic. Five minutes should be considered as the minimum aggregation period.

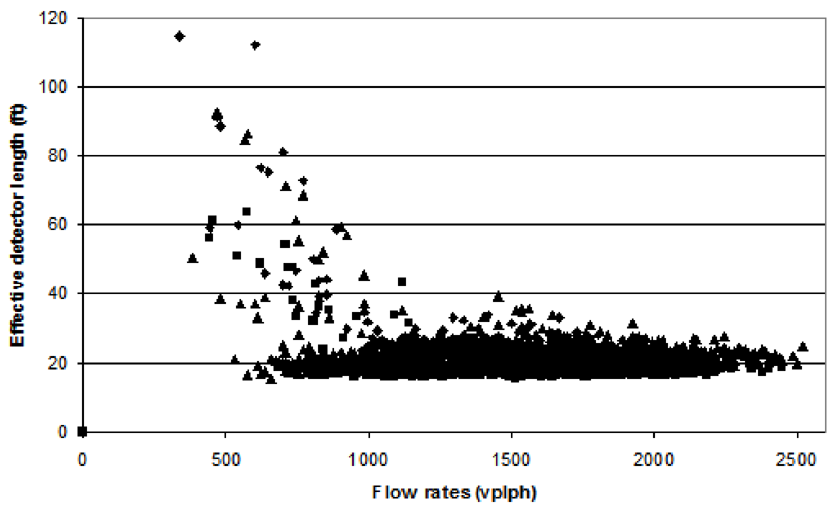


Figure 25. Effect of flow rate on EVL at a selected station

The flow rate histograms for the four selected stations were created as shown in Figure 26 from five-minute data. From the cumulative percentages, 96.7 percent of the EVLs are less than 30 feet, which was selected as the threshold criterion for purposes of this test.

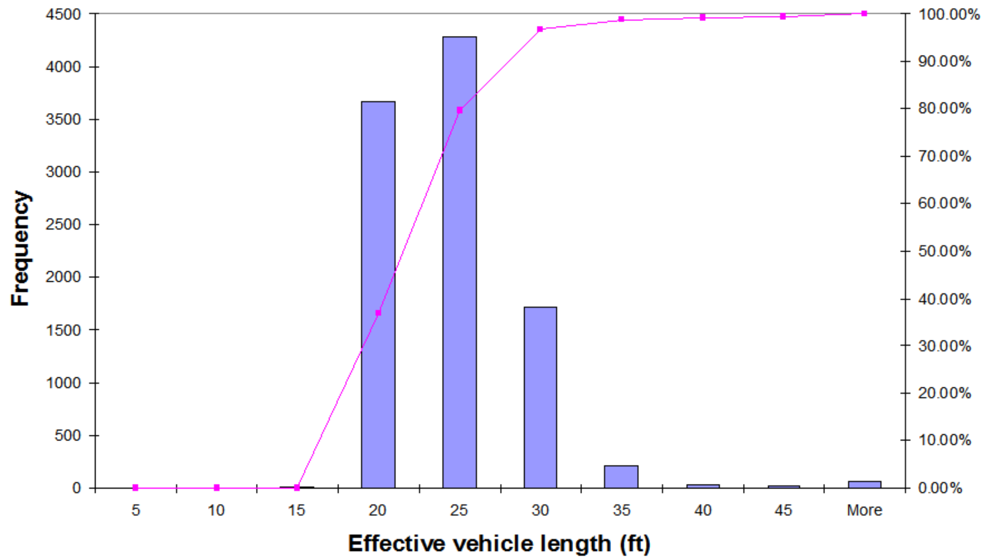


Figure 26. EVL histogram from four selected stations

To provide a more comprehensive example, the EVL report was obtained from the STEWARD Web site by applying the following selection criteria:

- All stations on I-95 NB and SB in District 2
- All weekdays in October 2008
- Five-minute aggregation of data

Table 12 shows the histogram of the effective vehicle length from this report. As this table shows, the distributions of effective vehicle lengths are different between Lane 1 and Lane 2 from all the stations during the study period.

Figure 27 shows that effective vehicle lengths from Lane 2 are more skewed to right than the Lane1. The average effective vehicle length from Lane 2 is 22.1 feet and Lane 1 is 16.0 feet. It might be expected that the average vehicle length would differ slightly between lanes because of different truck percentages. However, large differences between lanes might suggest that some calibration of the detectors should be considered. In this example, approximately 20 percent and 5 percent of the intervals showed lengths less than 15 feet in Lanes 1 and 2, respectively. Approximately 3 percent and 8 percent of the intervals showed lengths greater than 30 feet in Lanes 1 and 2, respectively.

Figure 28 and **Figure 29** show the occupancy-flow relationships for low-effective vehicle length cases (< 10 fee) and normal cases (>10 feet). Note that the average occupancy in the low effective vehicle length cases is much smaller than the normal case and the slope of the flow-occupancy graph for the small effective vehicle length case is much steeper than that for the normal case. Since effective vehicle length is proportional to (occupancy / flow), it would be anticipated that low effective vehicle lengths would be associated with low occupancy values. This example illustrates how the effective vehicle lengths may be used to provide additional insight into the operation of a SunGuide detector station. More specific examples of the use of this parameter are provided in Section 6.4.1.

	Lane1		*Lane2*	
EVL:(ft)	*Frequency*	*Cumulative %*	*Frequency*	*Cumulative %*
1	0	0.00%	1	0.00%
2	0	0.00%	0	0.00%
3	1	0.00%	3	0.01%
5	382	0.73%	333	0.64%
10	10505	20.71%	2064	4.57%
15	15194	49.60%	3735	11.67%
20	14412	77.01%	12903	36.20%
25	7972	92.17%	19892	74.03%
30	2723	97.35%	9619	92.32%
35	784	98.84%	2359	96.81%
40	186	99.20%	834	98.40%
45	98	99.38%	325	99.01%
50	58	99.49%	153	99.30%
55	43	99.58%	94	99.48%
60	31	99.63%	55	99.59%
65	26	99.68%	40	99.66%
70	24	99.73%	17	99.70%
75	16	99.76%	21	99.74%
80	21	99.80%	17	99.77%
85	14	99.83%	14	99.79%
90	13	99.85%	12	99.82%
95	3	99.86%	7	99.83%
100	3	99.86%	10	99.85%
More	72	100.00%	79	100.00%

Table 12. Histogram of the Effective Vehicle Lengths

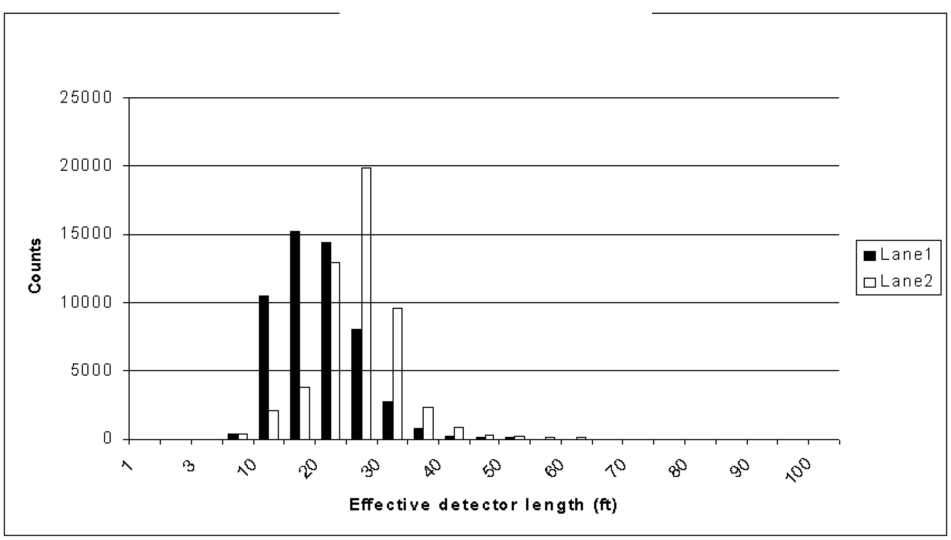


Figure 27: Lane-by-lane effective vehicle length distribution

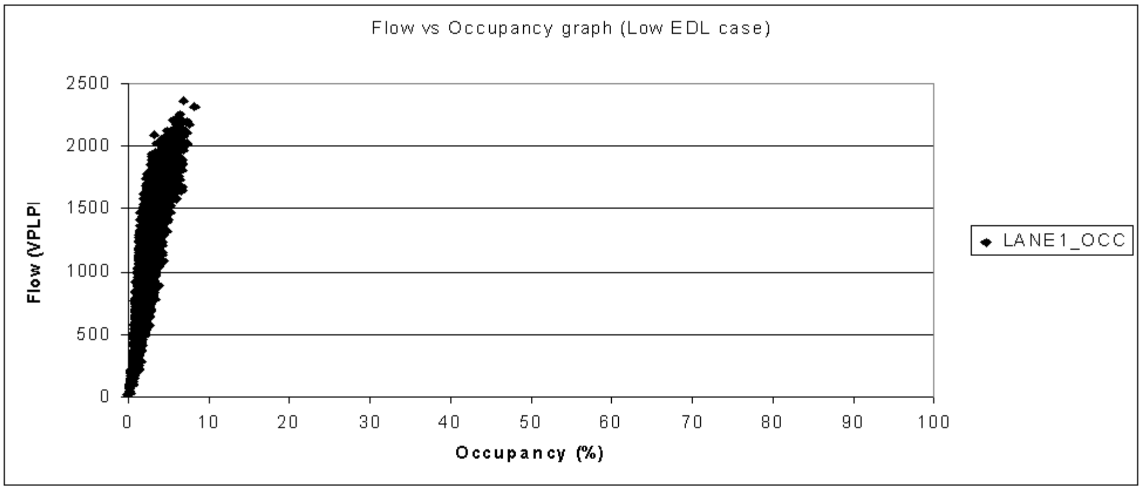


Figure 28. Flow-occupancy relationship for low effective vehicle lengths

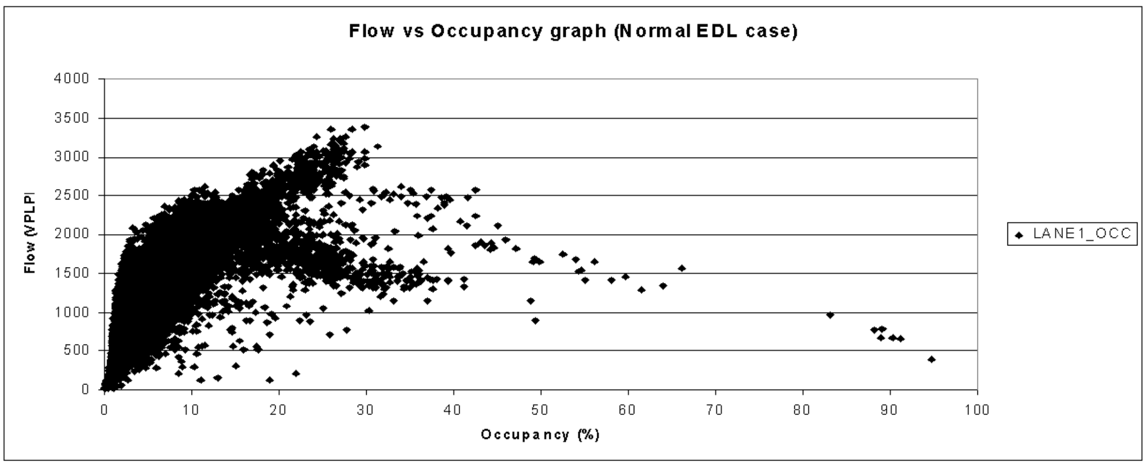


Figure 29. Flow-occupancy relationship for normal effective vehicle lengths

6.3.4 Lane Volume Balance Ratio

As indicated in Section 3, the lane volume balance ratio (LVBR) is expressed as the ratio of the highest to lowest lane volume at each station. This measure is calculated from the lane-by-lane volumes with five-minute, 15-minute and one-hour aggregation levels.

From this relationship, the LVBR was estimated using the four selected detectors. Figure 30 shows the LVBR for station 210511. It is observed that the value converges to a level near 1.0 at high-flow rates. This would be expected because, as the flow rates increase, the lane volume for all lanes should be similar. The maximum number is less than 2.0. Figure 31 shows the histogram of the lane volume balance ratio from all four stations. Note that that 96.1 percent of lane LVBR values are less than 2. During the estimation, the observations with the large occupancy values (>18 percent) were excluded to avoid downstream situations that might affect the natural balance. For example, when the lanes are partially closed during an incident, the lane volume balance ratio could increase substantially.

To provide a more comprehensive example, the LVBR report was obtained from the STEWARD Web site by applying the following selection criteria:

- All detectors on I-95 NB and SB in District 2
- All weekdays of October 2008
- Five-minute aggregation

Table 13 shows the histogram of the lane volume balance from all stations in the STEWARD report. The maximum LVBR value is limited to 99 during the ETL process to avoid meaningless results. Therefore, if the highest lane volume is 99 times or more than the lowest volume, the ratio will be archived as 99. Note that 82 percent of lane volume balance observations are less than 2, and approximately 5 percent of the observations are greater than 10.

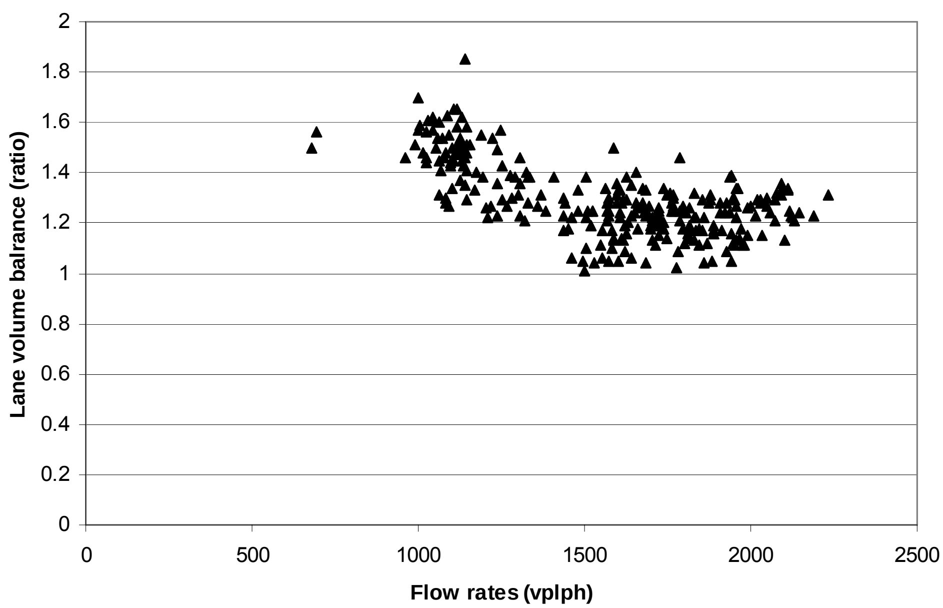


Figure 30. Effect of flow rate on lane volume balance ratio at a selected station

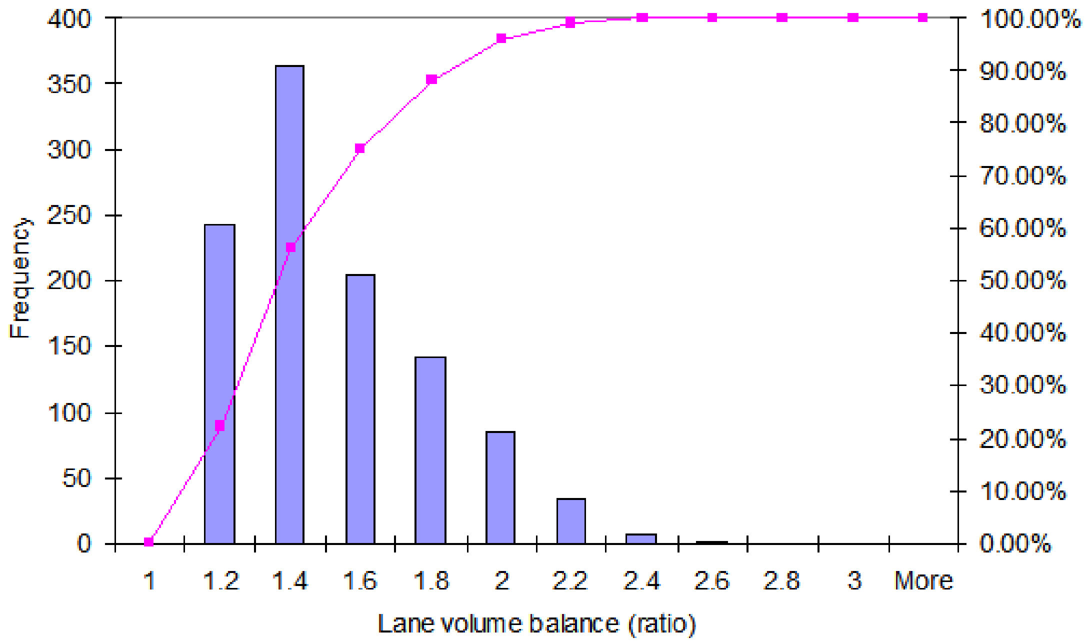


Figure 31. Lane volume balance ratio histogram from four selected stations

Table 13: Histogram of Lane Volume Balance Ratio

Lane Volume Balance Ratio	Frequency	Cumulative %
1	3,348	6.17%
2	32,226	65.57%
3	8,917	82.01%
4	3,086	87.70%
5	1,557	90.57%
10	2,583	95.33%
15	1,061	97.29%
20	548	98.30%
25	292	98.84%
30	183	99.17%
35	110	99.38%
40	65	99.49%
45	47	99.58%
50	51	99.68%
55	44	99.76%
60	13	99.78%
65	8	99.80%
70	10	99.81%
75	14	99.84%
80	25	99.89%
85	15	99.91%
90	5	99.92%
95	6	99.93%
100	36	100.00%

There are several scenarios that might explain cases with high LVBR values:

- Detector configuration problems
- Lane configuration problems
- Installation and calibration issues
- Incidents
- Downstream origin/destination issues

Detector configuration problems are the most systematic of all lane volume balance ratio problems and therefore tend to produce substantially higher values. As an example of a possible detector configuration problem, all of the stations with lane volume balance ratios of 10 or more are shown in

Station ID	*Frequency*
210032	761
200201	698
210711	280
200082	269
210412	230
200112	118
210122	51
210562	49
200192	45
200281	23
210681	11
210422	7
210041	6
210192	5
210642	4
200031	4
200141	2
210702	1
210162	1

Table 14. Several stations generate multiple cases of lane volume balance over 10.0 during the data period. For example, it is observed that station 210032 has 761 occurrences of excessive LVBR, accounting for 11.5 percent of the entire time period. Stations 210032, 210711, 200081, 210122 and 210702 also have an excessive maximum-flow rate problem in addition to a lane volume balance problem.

Station ID	*Frequency*
210032	761
200201	698
210711	280
200082	269
210412	230
200112	118
210122	51

210562	49
200192	45
200281	23
210681	11
210422	7
210041	6
210192	5
210642	4
200031	4
200141	2
210702	1
210162	1

Table 14. Frequency of LVBR Greater than 10.0

All of these cases were verified at a lane-by-lane level. All of the stations have a similar pattern in which one of the lanes has a much lower average flow rate than other lanes. This situation could be explained by either a detector configuration problem or a physical anomaly in the lane configuration.

Detector configuration problem Example
Station 210032 appears to have a detector configuration problem. Average flow rates for the lanes from Station 210032 are as follows.

	Lane1	*Lane2*	*Lane3*
Average flow rate (VPLPH)	1452	2277	140

Note that Lane 2 has relatively high flow rates for the sampling period and Lane 3 has minimal flow rates.

This station is located on the Fuller Warren Bridge and covers I-95 SB as Figure 32. Location of Station 210032shows. Lane 3 is a diverging lane for I-95 and the exit ramp and is wider than other lanes. This detector might need calibration to cover Lane 3 more precisely.

Figure 32. Location of Station 210032

Lane Configuration Problem Example
Station 200201 would appear to have a lane configuration problem. Average flow rates for three lanes as follows. It is evident here that the Lane 1 volume is very small compared to the other lanes.

	Lane1	*Lane2*	*Lane3*
Average flow rate (VPLPH)	78	872	1166

As Figure 33 shows, this station is located at the south of Norwood Avenue and covers two NB through lanes. This lane configuration and the facility data from the traffic data do not match and need to be verified. .

Figure 33. Location of Station 200201

6.3.5 Hourly and Daily Volume Variation

Hourly and daily volume variation can be used to verify the station data quality. Figure 34 shows the hourly volume variation at station 210471. The station is located at the south end of Jacksonville's downtown. The traffic is northbound with three through lanes. The x-axis shows the time of day and the y-axis shows the station flow rates (veh/hr). The average flow rates are shown on this figure along with upper and lower 95 percent confidence bounds based on the computed standard deviation. The highest flow rates are observed during the morning peak around 7:30 a.m.

The volume levels and peaking characteristics are typical of what would be expected at a freeway traffic monitoring station. Unexplained discrepancies from these typical characteristics could suggest a problem with the detectors.

Figure 35 shows the total 24-hour weekday volume at station 210471 for f 2008. This graph excludes all of the weekend data but not holidays. Note that some days show unreasonably low volumes indicating possible mechanical problems. Most of the District 2 detectors exhibit periods of missing data from several days to months. Many of the problems were associated with construction activities. More recent data suggest that these problems have been resolved.

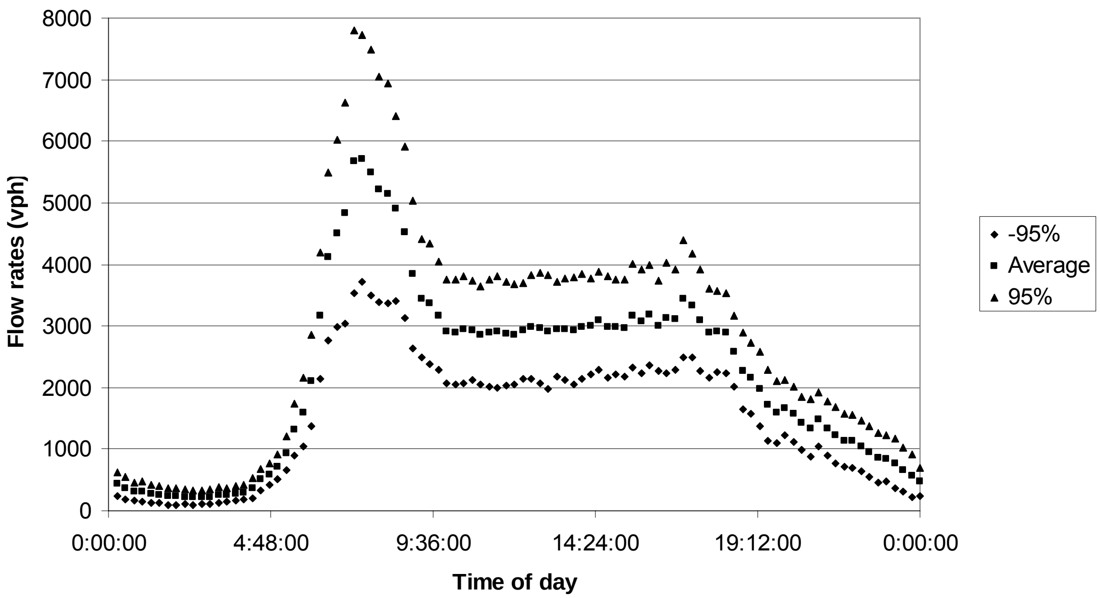


Figure 34. Average flow rates by time of day for a selected station

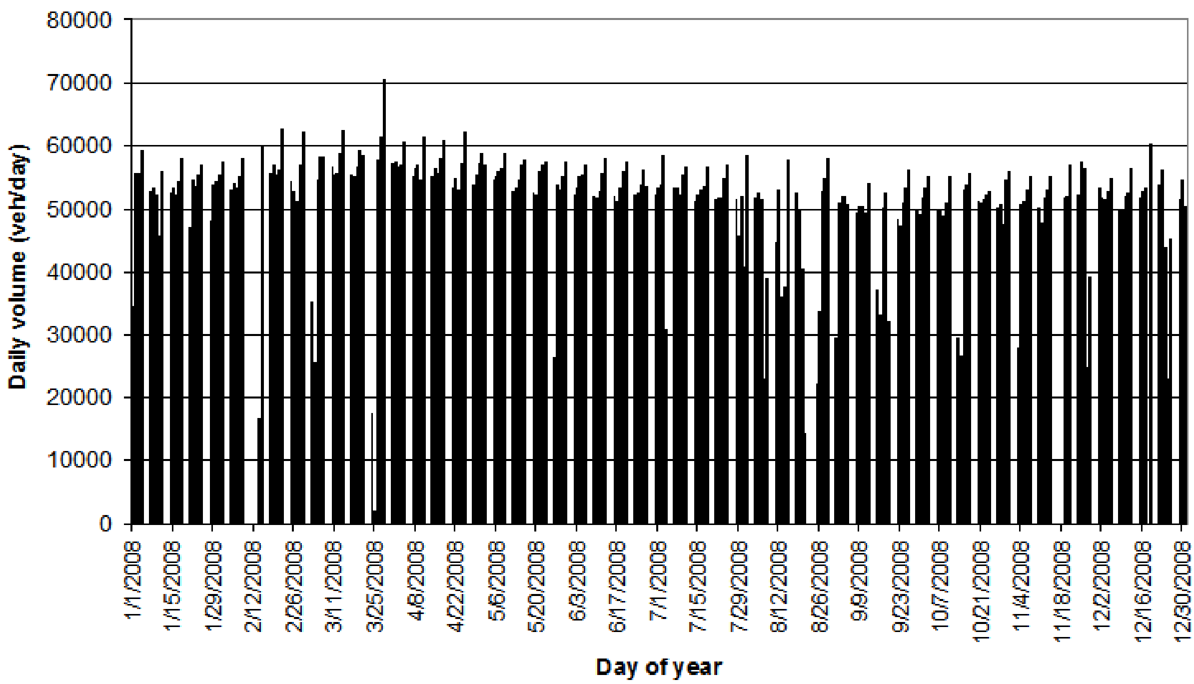


Figure 35. Total 24-hour weekday volume for one year at a selected station

6.4 Level 4: System Level Tests

The level 4 system data validation uses the continuity of traffic flow to check the quality of the traffic data. Both the traffic volumes and the EVL can be expected to be continuous within the system. To set up the criteria, a set of stations were chosen within spatial and temporal limits that were free of reported incidents:

- Facility: District 2 I-95 northbound, south of I-10
- Date: Oct. 6, 2008, Monday
- Time: Morning peak (7 a.m. — 10 a.m.) rush-hour traffic
- Eighteen active stations

6.4.1 Continuity of EVL

Figure 36 shows the EVL over the system. As described in the previous sections, EVLs are expected to have an average of 21 ft and a maximum value of less than 30 feet.

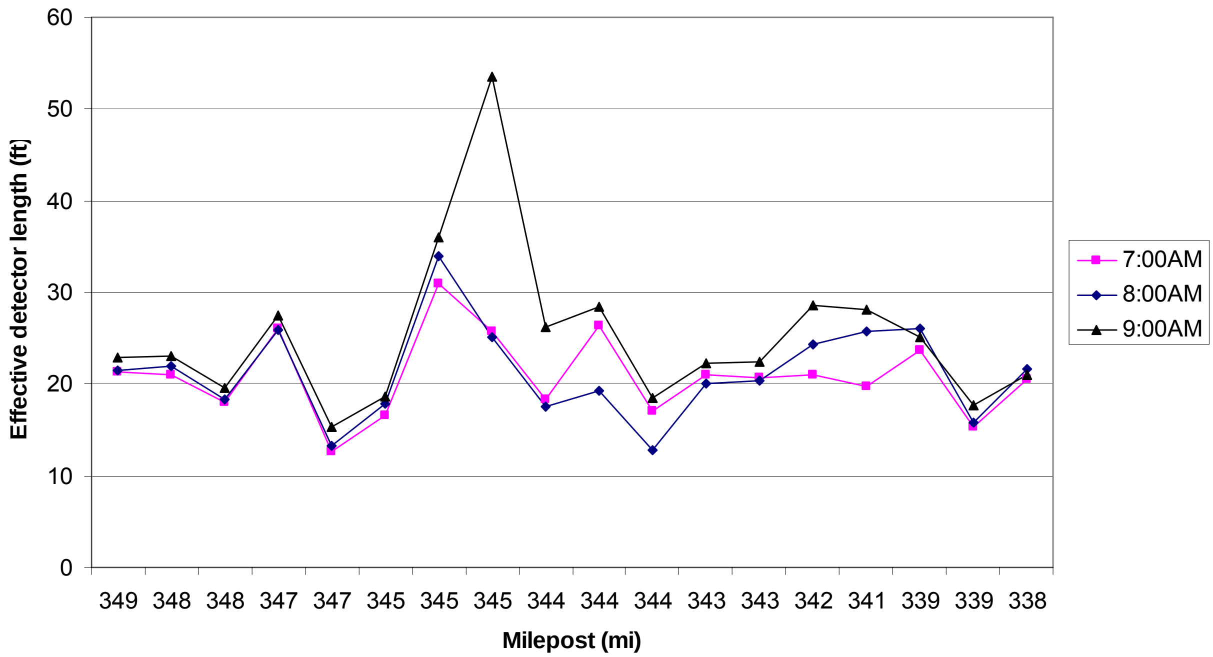


Figure 36: EVLs at stations on a section of I-95

The EVLs were fairly consistent over the three-hour period at most stations. Only one station showed values that were consistently above 30 feet. One station exhibited a peak value above 50 feet, suggesting a possible unreported incident. This example was presented to illustrate the potential use of EVL as a diagnostic tool. Additional investigation beyond the scope of this project would be required to draw develop definitive guidelines.

6.4.2 Continuity of Volume

Traffic volumes should also be continuous over time within the system, taking into account the entrance and exit ramp volumes. Unexplained discontinuities in traffic volumes over reasonably long periods, such as one hour, could be taken as a sign of a potential detector problem. An example of volume continuity analysis will be presented here.

The freeway section for this example is 11.26 miles long, with 18 stations in the northbound direction. There are nine interchanges with 15 on/off ramps. Nine ramps are covered to count the exit/entry volumes. The number of lanes starts from four lanes and decreases to three from milepost 340mi to 349.4mi. Figure 37 shows the hourly traffic volume by freeway milepost.

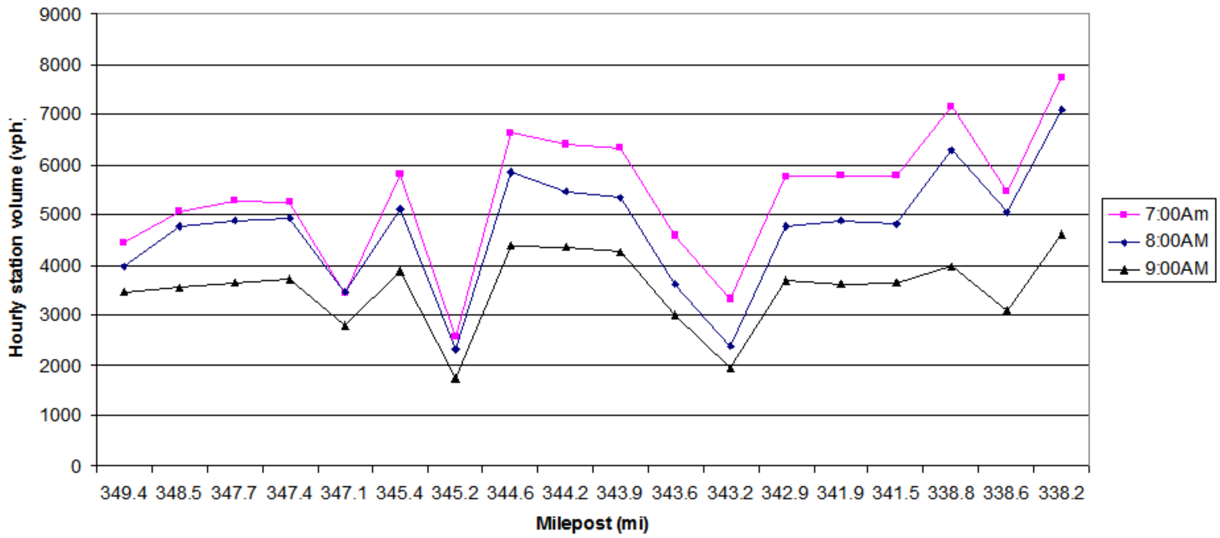


Figure 37. Hourly traffic volumes at stations on a section of I-95

Several discontinuities are observed in this figure, which shows three stations that appear to be consistently undercounting at mileposts 347, 345 and 343. These stations have 30 percent or more hourly volume differences compared to their upstream stations. The first two stations do not have exit or entry ramps between them and their upstream stations. Therefore, these detectors might need to be recalibrated. On the other hand, the third station has neighboring exit ramps, so the difference in volumes cannot be attributed entirely to counting accuracy. Without these two undercounting stations, the volume differences between consecutive detectors are relatively minimal.

System level analyses should be able to identify data quality problems that would be missed by lower-level tests. Reliability could be improved with segment-specific tests involving some knowledge of the facility configuration. Maximum reliability could be obtained by establishing historical benchmark values for comparison with the daily measures. The task of creating benchmark data will be possible when the system has been in operation for a few years, but that task is clearly beyond the scope of this project.

7 OPERATIONAL APPLICATIONS FOR ARCHIVED DATA

The data available from STEWARD can be used for both operational and research purposes. The two purposes are discussed in separate sections of this document because they tend to have different requirements. This section deals with operational applications.

7.1 Summary of Available Reports

The detailed reports that are available from STEWARD at various levels will be discussed first. Each report type will be described in terms of its field definitions and methods of computation.

7.1.1 Diagnostic Procedures

As each day's archive data file is processed, a log is created showing all of the problems encountered, with a summary at the end. A record is also added to the conversion history file to summarize the results for that day. The following diagnostic items are reported:

- *File Name*: The date is embedded in the file name.
- *From*: The time at which the first record was received.
- *To*: The time at which the last record was received.
- *Elapsed Minutes*: Should be 1,440 if the system ran for the whole day.
- *Null Minutes*: Number of minutes in which no report was received from any detector: This should be zero unless the system was off-line for a portion of the delay.
- *Total Records*: The number of records processed: Should be consistent from day to day.
- *Total Count*: The sum of all of the volumes reported: Some variation is expected, especially by day of week.
- *Missed Scans:* Number of times a report was not received within the specified polling interval.
- *Negative/Zero Scans*: Negative scans indicate that the time stamp for a report was before the time stamp for the previous report. Zero scans indicate duplicate reports with the same time stamp. There should be no negative or zero scans.
- *Orphan Stations, Lanes and Ramps:* Stations, lanes and ramps that are in the daily archive file but not in the configuration file.
- *Null Stations, Lanes and Ramps*: Stations, lanes and ramps that are in the configuration file but reported no data for the day.
- *Offline Stations, Lanes and Ramps:* Stations, lanes and ramps that are flagged as offline to avoid showing up as nulls in the conversion report.

7.1.2 Station Level Reports

Four reports are provided at the station level. The *All Data Fields Report* presents the station level traffic data (speed, volume and occupancy) and their statistics. The details of each column are described in Table 15.

Table 15. Column Description for the *All data fields* Report

Measure	*Unit*	*Description*
DAY	N/A	Date
TIME	N/A	Time
STATION_ID	N/A	A statewide-unique station identifier in the format DFnnnS, where: • D is the district number • F is the facility number within the district • nnn is the sequence number of the station within the facility S is direction (1 = increasing mileposts, 2 = decreasing mileposts)
FWY_SPD	Mi/Hr	Volume-weighted average thru speed
FWY_VOL	Veh	Sum of thru volume
FWY_OCC	Percentage	Average thru occupancy
SPD_CV	Veh	Coefficient of variation for speed
VOL_RATIO	Ratio	Ratio of max volume lane to min volume lane
ENTRY_VOL	Veh	Sum of entry ramp volume
EXIT_VOL	Veh	Sum of exit ramp volume
FWY_QA	Ratio	Percentage of freeway volume observation hit rate 100 * (Thru Volume Observed)/(Thru Volume observation expected)
ENTRY_QA	Ratio	Percentage of on ramp volume observation hit rate 100*(on ramp volume observed)/(on ramp volume observation expected)
EXIT_QA	Ratio	Percentage of off ramp volume observation hit rate 100*(off ramp volume observed)/(off ramp volume observation expected)
HOV_VOL	Veh	Sum of HOV lane volume
HOV_SPD	Mi/Hr	Volume-weighted average HOV speed
HOV_OCC	Percentage	Average HOV occupancy
HOV_QA	N/A	Percentage of HOV volume observation hit rate 100*(HOV volume observed)/(HOV volume observation expected)

The T*raffic Counts Report* presents the lane volume data and their statistics. The details of each column are described in Table 16.

Table 16: Column Descriptions for the *Station Traffic Counts R*eport

Measure	*Unit*	*Description*
DATE	N/A	Date
TIME	N/A	Time
FACILITY		Facility ID District 2 0: North of I-95, 1: South of I-95, 2: I-295 District 4 1: I-95, 2: I-595 District 6 0: I-95, 1: SR 826, 3: US-1, 4: I-195, 5: 1-75
STATION_ID	N/A	A statewide-unique station identifier in the format DFnnnS, where: • D is the district number • F is the facility number within the district • nnn is the sequence number of the station within the facility S is direction (1=increasing mileposts, 2=decreasing mileposts)
NUM_OF_LANES	N/A	Number of lanes
DIRECTION	N/A	Direction (N/S/E/W)
STATION_DESC	Mile	Station description
STATION_MP	N/A	Station milepost
COUNT_STATION	N/A	Count station ID provided by FDOT statistical office
TOTAL	Veh	Total thru volume LANE1_VOL + LANE2_VOL+ LANE3_VOL+ LANE4_VOL + LANE5_VOL + LANE6_VOL
LANE1_VOL	Veh	Lane1 volume
LANE2_VOL	Veh	Lane2 volume
LANE3_VOL	Veh	Lane3 volume
LANE4_VOL	Veh	Lane4 volume
LANE5_VOL	Veh	Lane5 volume
LANE6_VOL	Veh	Lane6 volume
BALANCE	Ratio	Ratio of max volume lane to min volume lane
FWY_QA	Percentage	Percentage of freeway volume observation hit rate 100 * (Thru Volume Observed)/(Thru Volume observation expected)
ON_RAMP1	Veh	On ramp1 volume
ON_RAMP2	Veh	On ramp2 volume
ON_RAMP3	Veh	On ramp3 volume
ON_RAMP_QA	Percentage	Percentage of on ramp volume observation hit rate 100*(on ramp volume observed)/(on ramp volume observation expected)
OFF_RAMP1	Veh	Off ramp1 volume
OFF_RAMP2	Veh	Off ramp2 volume
OFF_RAMP3	Veh	Off ramp3 volume
OFF_RAMP_QA	Percentage	Percentage of off ramp volume observation hit rate 100*(off ramp volume observed)/(off ramp volume observation expected)
COUNTY	N/A	County ID provided by FDOT statistical office

The *Maximum Flow Report* presents the maximum hourly flow rate for the station. The details of each column are described in Table 17.

Table 17: Column Description for the *Maximum Flow R*eport

Measure	*Unit*	*Description*
DATE	N/A	Date
FACILITY	N/A	Facility ID District 2 0: North of I-95, 1: South of I-95, 2: I-295 District 4 1: I-95, 2: I-595 District 6 0: I-95, 1: SR 826, 3: US-1, 4: I-195, 5: 1-75
STATION_ID	N/A	A statewide-unique station identifier in the format DFnnnS, where: • D is the district number • F is the facility number within the district • nnn is the sequence number of the station within the facility S is direction (1=increasing mileposts, 2=decreasing mileposts)
DIRECTION	N/A	Direction 1: NB or EB (increasing mileposts,) 2: SB or WB (decreasing mileposts)
STATION_DESC	N/A	Station description
STATION_MP	Mi	Station milepost
LANE_NUM	N/A	Number of lanes
MAX_FLOW	Vphpl	Max hourly flow rate at the selected stations
MAX_TIME	N/A	Timestamp when Max_Flow occurred

The *Effective Vehicle Length Report* presents the effective vehicle length data and its statistics. The details of each column are described in Table 18.

7.1.3 Section Level Reports

Two new section level reports have been developed. The *Performance Measures Report* presents the important performance measures for each segment within the section and provides totals for the section as a whole. Table 19 describes the details of the columns for each segment and Table 20 describes the baseline. See Table 21 for the sample report.

Table 18. Column Description for the *Effective Vehicle Length R*eport

Measure	*Unit*	*Description*
DATE TIME	N/A	Date
FACILITY	N/A	Facility ID for each District • District 2 0: North of I-95, 1: South of I-95, 2: I-295 • District 4 1: I-95, 2: I-595 • District 6 0: I-95, 1: SR 826, 3: US-1, 4: I-195, 5: 1-75
STATION_ID	N/A	A statewide-unique station identifier in the format DFnnnS, where: • D is the district number • F is the facility number within the district • nnn is the sequence number of the station within the facility • S is direction (1=increasing mileposts, 2=decreasing mileposts)
DIRECTION	N/A	Direction 1: NB or EB (increasing mileposts,) 2: SB or WB (decreasing mileposts)
STATION_DESC	N/A	Station description
STATION_MP	Mi	Station milepost
LANE1_VOL	Veh	Lane1 volume
LANE1_SPD	mi/hr	Lane1 volume-weighted average speed
LANE1_OCC	percentage	Lane1 average occupancy
LANE1_EFF_DET_LENGTH	Ft	Lane1 effective vehicle length EVL = Speed * Occupancy/ Flow = (LANE1_SPD * 5280ft/mi) * (LANE1_OCC /100) / (LANE1_VOL * 12)
LANE2_VOL	Veh	Lane2 volume
LANE2_SPD	mi/hr	Lane2 volume-weighted average speed
LANE2_OCC	percentage	Lane2 average occupancy
LANE2_EFF_DET_LENGTH	Ft	Lane2 effective vehicle length
LANE3_VOL	Veh	Lane3 volume
LANE3_SPD	mi/hr	Lane3 volume-weighted average speed
LANE3_OCC	percentage	Lane3 average occupancy
LANE3_EFF_DET_LENGTH	Ft	Lane3 effective vehicle length
LANE4_VOL	Veh	Lane4 volume
LANE4_SPD	mi/hr	Lane4 volume-weighted average speed
LANE4_OCC	percentage	Lane4 average occupancy
LANE4_EFF_DET_LENGTH	Ft	Lane4 effective vehicle length
LANE5_VOL	Veh	Lane5 volume
LANE5_SPD	mi/hr	Lane5 volume-weighted average speed
LANE5_OCC	percentage	Lane5 average occupancy
LANE5_EFF_DET_LENGTH	Ft	Lane5 effective vehicle length
LANE6_VOL	Veh	Lane6 volume
LANE6_SPD	mi/hr	Lane6 volume-weighted average speed
LANE6_OCC	percentage	Lane6 average occupancy
LANE6_EFF_DET_LENGTH	Ft	Lane6 effective vehicle length

Table 19: Column Descriptions for the *Performance Measures R*eport

Measure	*Unit*	*Description*
SEGMENT	N/A	Segment is defined as the link between the current station (downnode) and the upnode station in the upper row
MP	Mile	Milepost of current station
LENGTH	Mile	Segment length = *(upnode MP – downnodeMP)*
AVERAGE VOLUME	Veh	Daily average of total link volume *Sum(((Upnode thru volume + upnode entry volume)* *+ (downnode thru volume + upnode exit volume)) /2) /(Number of days)*
LANES	N/A	Number of thru lanes in downnode
VOL per LANE	Vphpl	Average per-lane hourly flow rate *Average of (AVERAGE_VOLUME/LANES/hours)*
VEH-MILES	Veh-Mi	Daily average of Veh-Miles = *AVERAGE_VOLUME * LENGTH*
VEH-HOURS	Veh-Hr	Daily average of Veh-Hours *Link volume * LENGTH / Link Speed* *= Sum (((Upnode thru volume + upnode entry volume)* *+ (downnode thru volume + upnode exit volume)) /2* ** LENGTH / downnode speed) / (Number of days)*
SPEED	Mi/Hr	Volume-weighted downnode thru speed *Sum(downnode speed* ** ((Upnode thru volume + upnode entry volume)* *+ (downnode thru volume + upnode exit volume)) /2)* */Sum(((Upnode thru volume + upnode entry volume)* *+ (downnode thru volume + upnode exit volume)) /2)*
DELAY	Veh-Hr	Congestion delay: When downnode thru speed is less than the reference speed, delay is calculated as sum of the differences between the link travel time measured and the reference link travel time. Reference speed is defined as 2/3 of the speed limit. *Sum(((Upnode thru volume + upnode entry volume)* *+ (downnode thru volume + upnode exit volume)) /2* ** LENGTH * (1/downnode speed -1.5/ downnode speed limit)))*
KINETIC ENERGY (10^6)	Veh-Mi /Hr	Daily average of Kinetic energy *Sum (((Upnode thru volume + upnode entry volume)* *+ (downnode thru volume + upnode exit volume)) /2* ** downnode speed) / (Number of days)*
PERCENT OBSERVATIONS	%	Percentage of downnode data observation hit rate 100*(downnode volume observed)/((distinct numbers of days that downnode volume observed)*(number of observations expected per day))
DENSITY	Veh/Mi/ Ln	Max density using 15min data. *Max (Hourly flow rate / downnode speed/ Number of lanes)* *= Max (4 * Sum((Upnode thru volume + upnode entry volume)* *+ (downnode thru volume + upnode exit volume) /2) during 15min* */downnode speed /Number of lanes)*
V/C RATIO	%	Max volume/capacity ratio using 15min data *Max (Hourly flow rate / Lane capacity / Number of lanes)* *= Max (Sum((Upnode thru volume + upnode entry volume)* *+ (downnode thru volume + upnode exit volume) /2) during 15min* */Lane Capacity (2,200 vphpl) / Number of lanes)*
LOS	N/A	Level of service *If V/C ratio > 100% LOS = F* *Else if Density > 35 LOS = E* *Else if Density > 26 LOS = D* *Else if Density > 18 LOS = C* *Else if Density > 11 LOS = B* *Else LOS = A*

Table 20. Column Descriptions for the *Performance Measures* Report Baseline

Measure	*Unit*	*Description*
SEGMENT	N/A	N/A
MP	N/A	N/A
LENGTH	Mi	Total segment length
AVERAGE VOLUME	N/A	N/A
LANES	N/A	N/A
VOL per LANE	N/A	N/A
VEH-MILES	Veh-Mi	Daily Veh-Miles for the total segments Sum of (*VEH-MILES*)
VEH-HOURS	Veh-Hr	Daily Veh-Hours for the total segments Sum of (*VEH-HOURS*)
SPEED	Mi/Hr	Segment average speed *VEH-MILES / VEH-HOURS*
DELAY (Veh-Hr)	Veh-Hr	Sum of Congestion delay Sum of (*Delay*)
KINETIC ENERGY	Veh-Mi/Hr	Daily average of total kinetic energy Sum of (*Kinetic Energy*)
PERCENT OBSERVATIONS	N/A	N/A
DENSITY	N/A	N/A
V/C RATIO	N/A	N/A
LOS	N/A	N/A

Table 21: Sample Performance Measures Report

SEGMENT	MP	LENGTH	AVERAGE VOLUME	LANES	VOL per LANE	VEH-MILES	VEH-HOURS	SPEED	DELAY (Veh-Hr)	KINETIC ENERGY	PERCENT OBSERVATIONS	DENSITY	V/C RATIO	LOS
I-95 SB Entrance from Bowden Rd														
I-95 SB South of Bowden Rd	344.56	0.29	290	3	386	84	1	58.6	0	0.02	1	3.6	0.1	A
I-95 SB North of Butler Blvd	344.21	0.35	62393	3	885	21650	351	62.6	0	3.91	97.9	36.1	0.9	E
I-95 SB Exit to Butler Blvd	343.85	0.36	60068	3	852	21684	320	67.9	0	4.08	97.9	27.6	0.9	D
I-95 SB Entrance from Butler Blvd WB	343.67	0.18	38534	3	547	6975	105	66.7	0	2.57	97.9	18	0.6	B
I-95 SB Entrance from Butler Blvd EB	343.26	0.41	46300	3	676	19122	295	67.9	12	3.14	95.1	39.4	0.8	E
I-95 SB South of Butler Blvd	342.9	0.35	49024	3	716	17306	257	67.3	0	3.3	95.1	26.2	0.8	D
I-95 SB Between Butler and Baymeadows	342.48	0.42	52768	3	748	22163	313	70.9	0	3.74	97.9	27.2	0.9	D
I-95 SB North of Baymeadows Rd	341.94	0.55	53128	3	754	29168	435	67.2	0	3.57	97.9	27.8	0.9	D
I-95 SB Entrance from Baymeadows Rd	341.11	0.38	142	3	190	54	1	63.8	0	0.01	1	1.6	0	A
Totals:		3.3				138206	2078	66.5	12	24.33				

The *Travel Time Reliability Report* presents the measures that are used nationally in travel time reliability assessment. It also presents two measures (Percent on-time arrivals and on-time delay) that are specific to Florida. Table 22 describes the details of the columns for each segment. Table 23 describes the baseline and Table 24 describes frequency table for the travel time. See Table 25 for the sample report.

Table 22: Column Description for the *Travel Time Reliability* Report

Measure	*Unit*	*Description*
SEGMENT	N/A	Segment is defined as the link between the current station (downnode) and the upnode station in the upper row
MP	Mi	Milepost of current station
LENGTH	Mi	Total segment length
AVERAGE VOLUME	N/A	Compensated downnode daily thru volume *Sum (Downnode thru volume / ((thru volume observed)/(thru volume observation expected))) / (Number of days)*
LANES	N/A	Number of thru lanes in downnode
SPEED	Mi/hr	Volume-weighted Downnode thru speed Sum*(Downnode thru seed * Compensated volume)* */* Sum*(Compensated volume)* *= Sum (Downnode thru seed * (Downnode thru volume* */ ((thru volume observed)/(thru volume observation expected)))* */ sum (Downnode thru volume* */ ((thru volume observed)/(thru volume observation expected))*
AV_TT	Min/Veh	Average travel time in minutes *Average (60* LENGTH/SPEED)*
TT INDEX	Ratio	Average TT / Reference TT *Average(Downnode speed limit / Downnode speed)*
PCNT ONTIME	Percent	Percent of 5 min intervals in which AvTT =< (TT with 10mi below the speed limit)
95% TT	Min	95 %ile value of travel time in this segment
BUFFER INDEX	Ratio	Buffer index *(95% TT – AV_TT) / AV_TT*
ONTIME DELAY	Min/Veh	On time delay per vehicle: When downnode thru speed is less than the reference speed, delay is calculated as sum of the differences between the link travel time measured and the reference link travel time devided by the number of vehicles. Reference speed is defined as 10 mph below the speed limit. *60* LENGTH * Compensated volume *(1/(downnode speed)-1/(downnode reference speed))/ Compensated volume* *= 60* LENGTH * (1/(downnode speed)-1/(downnode reference speed))* *= 60* LENGTH* ** (1/(downnode speed)-1/((downnode speed limit)-10))*
CONGESTION DELAY	Min/Veh	Congestion delay per vehicle: When downnode thru speed is less than the reference speed, delay is calculated as sum of the differences between the link travel time measured and the reference link travel time. Reference speed is defined as two-thirds of the speed limit. *60 * LENGTH * Compensated volume *(1/(downnode speed)-1/(downnode reference speed))/ Compensated volume* *= 60 * LENGTH *(1/(downnode speed)-1/(downnode reference speed))* *= 60* LENGTH* ** (1/downnode speed -1.5/ downnode speed limit)))*

Table 23. Column Descriptions for the *Travel Time Reliability* Report Baseline

Measure	Unit	Description
SEGMENT	N/A	N/A
MP	N/A	N/A
LENGTH	Mi	Segment length (distance between upnode and downnode)
AVERAGE VOLUME	Veh	N/A
LANES	N/A	N/A
SPEED	Mi/hr	Average SPEED *LENGTH / (AV_TT/60)*
AV_TT	Min/Veh	Sum of AV_TT in the segment
TT INDEX	Ratio	Average TT / Reference TT *Average(Downnode speed limit / (Length-weighted average speed))*
PCNT ONTIME	Percent	Percent of 5 min intervals in which (section AvTT) =< (section TT with 10mi below the speed limit)
95% TT	Min	95 %ile value of travel time in this section
BUFFER INDEX	Ratio	Section buffer index *(95% TT – AV_TT) / AV_TT*
ONTIME DELAY	Min/Veh	Sum of ontime delay per vehicle
CONGESTION DELAY	Min/Veh	Sum of congestion delay per vehicle

Table 24/ Travel Time Reliability Frequency Table Column Descriptions

Measure	Unit	Description
INDEX	N/A	Row index
LOWER LIMIT(min)	Min	Lower limit of the histogram is the minimum value of section travel time. LOWER LIMIT is calculated as follows: LOWER LIMIT = Min (Travel Time) +(Max (Travel Time) - Min (Travel Time)) /NUMBER_OF_BINS * (INDEX-1)
UPPER LIMIT(min)	Min	UPPER LIMIT = Min (Travel Time) +(Max (Travel Time) - Min (Travel Time)) /NUMBER_OF_BINS * INDEX
COUNT	N/A	Number of section travel time in this bin
CUMULATIVE PERCENT	Percentage	Cumulative percentage of travel time

Table 25: Sample Performance Measures Report

SEGMENT	MP	LENGTH	AVERAGE VOLUME	LANES	SPEED	AV_TT	TT INDEX	PCNT ONTIME	95% TT	BUFFER INDEX	ONTIME DELAY (Min/Veh)	CONGESTION DELAY (Min/Veh)
I-95 SB Entrance from Bowden Rd	344.85											
I-95 SB South of Bowden Rd	344.56	0.29	64450	3	54.91	0.31	1.17	87.94	0.47	0.48	0.03	0.02
I-95 SB North of Butler Blvd	344.21	0.35	62881	3	62.71	0.33	1.02	89.01	0.43	0.3	0.01	0
I-95 SB Exit to Butler Blvd	343.85	0.36	41343	3	67.81	0.32	0.97	100	0.35	0.08	0	0
I-95 SB Entrance from Butler Blvd WB	343.67	0.18	37859	3	66.48	0.17	1.02	90.78	0.21	0.23	0	0
I-95 SB Entrance from Butler Blvd EB	343.26	0.41	49212	3	68.51	0.37	0.98	97.06	0.4	0.08	0	0
I-95 SB South of Butler Blvd	342.9	0.35	53041	3	67.3	0.32	0.97	100	0.34	0.06	0	0
I-95 SB Between Butler and Baymeadows	342.48	0.42	54747	3	70.81	0.36	0.93	100	0.39	0.08	0	0
I-95 SB North of Baymeadows Rd	341.94	0.55	53657	3	67.15	0.5	0.99	99.65	0.55	0.1	0	0
I-95 SB Entrance from Baymeadows Rd	341.11	0.38	39461	3	63.23	0.37	1.04	95.74	0.41	0.11	0	0
Totals:		3.3			64.68	3.47	1.14	99.29	3.78	0.09	0.05	0.02

Frequency Table

INDEX	LOWER LIMIT(min)	UPPER LIMIT(min)	COUNT	CUMULATIVE PERCENT
1	6.23	6.33	19	6.74
2	6.33	6.44	44	22.34
3	6.44	6.55	64	45.04
4	6.55	6.66	42	59.93
5	6.66	6.77	21	67.38
6	6.77	6.88	27	76.95
7	6.88	6.99	19	83.69
8	6.99	7.1	17	89.72
9	7.1	7.21	12	93.97
10	7.21	7.32	6	96.1
11	7.32	7.43	7	98.58
12	7.43	7.54	2	99.29
16	7.87	7.98	1	99.65
20	8.31	8.42	1	100

7.1.4 Facility Level Reports

Three reports are provided at the facility level. The *All Data Fields Report* presents the facility level traffic data (speed, volume and occupancy) and their statistics. The details of each column are described in Table 26.

Table 26. Column Descriptions for the *All Data Fields* Report

Measure	*Unit*	*Description*
DATE	N/A	Date
TIME	N/A	Time
STATION_ID	N/A	A statewide-unique station identifier in the format DFnnnS, where: • D is the district number • F is the facility number within the district • nnn is the sequence number of the station within the facility S is direction (1=increasing mileposts, 2=decreasing mileposts)
STATION_DESC	N/A	Station description
STATION_MP	Mile	Station milepost
LANE1_VOL	Veh	Lane1 volume
LANE1_SPD	mi/hr	Lane1 volume-weighted average speed
LANE1_OCC	percentage	Lane1 average occupancy
LANE2_VOL	Veh	Lane2 volume
LANE2_SPD	mi/hr	Lane2 volume-weighted average speed
LANE2_OCC	percentage	Lane2 average occupancy
LANE3_VOL	Veh	Lane3 volume
LANE3_SPD	mi/hr	Lane3 volume-weighted average speed
LANE3_OCC	percentage	Lane3 average occupancy
LANE4_VOL	Veh	Lane4 volume
LANE4_SPD	mi/hr	Lane4 volume-weighted average speed
LANE4_OCC	percentage	Lane4 average occupancy
LANE5_VOL	Veh	Lane5 volume
LANE5_SPD	mi/hr	Lane5 volume-weighted average speed
LANE5_OCC	percentage	Lane5 average occupancy
LANE6_VOL	Veh	Lane6 volume
LANE6_SPD	mi/hr	Lane6 volume-weighted average speed
LANE6_OCC	percentage	Lane6 average occupancy
ONRAMP1_VOL	Veh	On ramp1 volume
ONRAMP1_SPD	mi/hr	On ramp1 volume-weighted average speed
ONRAMP1_OCC	percentage	On ramp1 average occupancy
ONRAMP2_VOL	Veh	On ramp2 volume
ONRAMP2_SPD	mi/hr	On ramp2 volume-weighted average speed
ONRAMP2_OCC	percentage	On ramp2 average occupancy
ONRAMP3_VOL	Veh	On ramp3 volume
ONRAMP3_SPD	mi/hr	On ramp3 volume-weighted average speed
ONRAMP3_OCC	percentage	On ramp3 average occupancy
OFFRAMP1_VOL	Veh	Off ramp1 volume
OFFRAMP1_SPD	mi/hr	Off ramp1 volume-weighted average speed
OFFRAMP1_OCC	percentage	Off ramp1 average occupancy
OFFRAMP2_VOL	Veh	Off ramp2 volume
OFFRAMP2_SPD	mi/hr	Off ramp2 volume-weighted average speed
OFFRAMP2_OCC	percentage	Off ramp2 average occupancy
OFFRAMP3_VOL	Veh	Off ramp3 volume
OFFRAMP3_SPD	mi/hr	Off ramp3 volume-weighted average speed
OFFRAMP3_OCC	percentage	Off ramp3 average occupancy

The *Volume map and I/O balance Report* presents a tabular map of the entry, exit and mainline volumes and indicates discrepancies in the balance of input to and output from each segment. The details for each column are described in Table 27.

Table 27. Column Descriptions for the *Volume Map and I/O Balance* Report

Measure	*Unit*	*Description*
DATE	N/A	Date
TIME	N/A	Time
STATION_ID	N/A	A statewide-unique station identifier in the format DFnnnS, where: • D is the district number • F is the facility number within the district • nnn is the sequence number of the station within the facility S is direction (1=increasing mileposts, 2=decreasing mileposts)
STATION_MP	Mile	Station milepost
UPNODE_ID	N/A	Upnode ID. Has the same format as Station_ID. Upnode is defined as the first upward station from the current station with STATION_ID
ENTRY_VOLUME	Veh	Station entry ramp volume
FWY_VOLUME	Veh	Station thru volume
EXIT_VOLUME	Veh	Station exit ramp volume
LINK_INPUT	Veh	Link input volume. Link is defined between the upnode and the current node. *Upnode entry ramp volume + upnode thru volume*
LINK_OUTPUT	Veh	Link output volume. Link is defined between the upnode and the current node. *Downnode thru volume + downnode exit ramp volume*
DIFFERENCE	Veh	Difference between the link input and output volumes *Link_input – Link_output*
PCNT_DIFF	Percentage	Percentage difference of link input and output volumes *100 * Difference / ((Link_input + Link_output)/2)*

The *Traffic Counts Report* presents the lane volume data for all stations in the facility. The details of each column are described in Table 28.

Table 28. Column Descriptions for the *Facility Traffic Counts* Report

Measure	*Unit*	*Description*
DATE	N/A	Date
TIME	N/A	Time
STATION_ID	N/A	A statewide-unique station identifier in the format DFnnnS, where: • D is the district number • F is the facility number within the district • nnn is the sequence number of the station within the facility S is direction (1 = increasing mileposts, 2 = decreasing mileposts)
STATION_DESC	N/A	Station description
STATION_MP	Mile	Station milepost
COUNT_STATION	N/A	Count station ID from Florida DOT statistical office
TOTAL	Veh	Total thru volume *LANE1_VOL + LANE2_VOL+ LANE3_VOL+ LANE4_VOL + LANE5_VOL + LANE6_VOL*
LANE1_VOL	Veh	Lane1 volume
LANE2_VOL	Veh	Lane2 volume
LANE3_VOL	Veh	Lane3 volume
LANE4_VOL	Veh	Lane4 volume
LANE5_VOL	Veh	Lane5 volume
LANE6_VOL	Veh	Lane6 volume
FWY_QA	Percentage	Percentage of freeway volume observation hit rate *100 * (Thru Volume Observed)/(Thru Volume observation expected)*
ON_RAMP1	Veh	On ramp1 volume
ON_RAMP2	Veh	On ramp2 volume
ON_RAMP3	Veh	On ramp3 volume
ON_RAMP_QA	Percentage	Percentage of on ramp volume observation hit rate *100*(on ramp volume observed)/(on ramp volume observation expected)*
OFF_RAMP1	Veh	Off ramp1 volume
OFF_RAMP2	Veh	Off ramp2 volume
OFF_RAMP3	Veh	Off ramp3 volume
OFF_RAMP_QA	Percentage	Percentage of off ramp volume observation hit rate *100*(off ramp volume observed)/(off ramp volume observation expected)*
COUNTY	N/A	County ID

7.2 Traffic Volume Data for Traffic Counting Programs

The FDOT Statistics Office maintains several continuous telemetered traffic count stations on Florida highways. Three permanent traffic counters are located on I-95 within the District 2 SunGuide system. With the cooperation of the Statistics Office, the research team was able to compare the data from one count station to the archived counts generated by SunGuide and stored in the STEWARD database. The permanent count station was located in the southbound lanes of Interstate 95 between Emerson Street and University Blvd. The two adjacent SunGuide detector stations were located approximately 1,000 feet north and 700 feet south of the permanent count station.

Figure 38 shows an example comparison between the hourly counts from the permanent count station and the two SunGuide detectors. Note that a near perfect agreement is apparent here. This will not always be the case and comparison of data from the two sources could potentially improve the accuracy of both sources. There is clearly a potential benefit that could be derived from a mutual exchange of traffic count data between the ITS centers and the Statistics Office. The Statistics Office data could provide an important reference for calibrating the ITS detectors, most of which are microwave based. The ITS data could provide a useful supplement to the statewide traffic count coverage now in place.

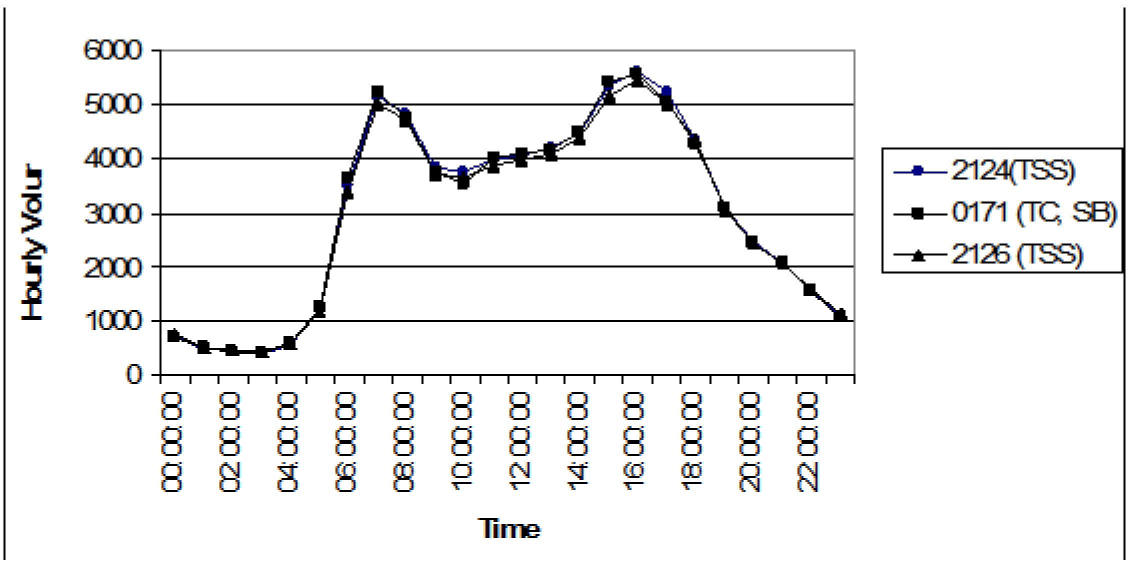


Figure 38. Example SunGuide TSS and Statistics Office count comparison

7.2.1 Converting ITS Data to FDOT Counts

The traffic counts in the SunGuide archives have essentially the same content as the FDOT Statistics Office and District Planning Office traffic count files. A desktop utility program has been developed to convert the count data in the SunGuide data archive to either of the FDOT count formats. This program, called ITSCounts, is summarized in Appendix 4b. An overview of the data flow is shown in Figure 39. The STEWARD database is accessed via the Internet to download traffic count data in CSV format. The ITSCounts utility program accepts the CSV formatted files as input and converts these file to either the FDOT Central or District Office count formats. A separately formatted spreadsheet file is also produced to facilitate plotting of the results.

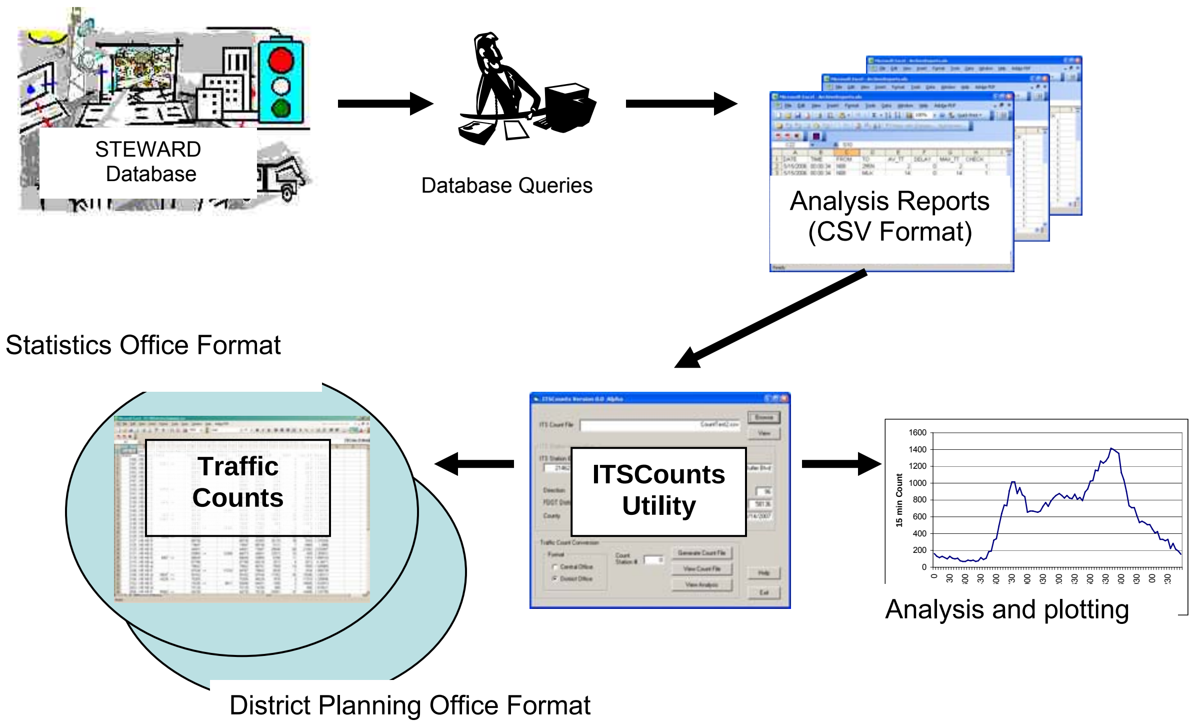
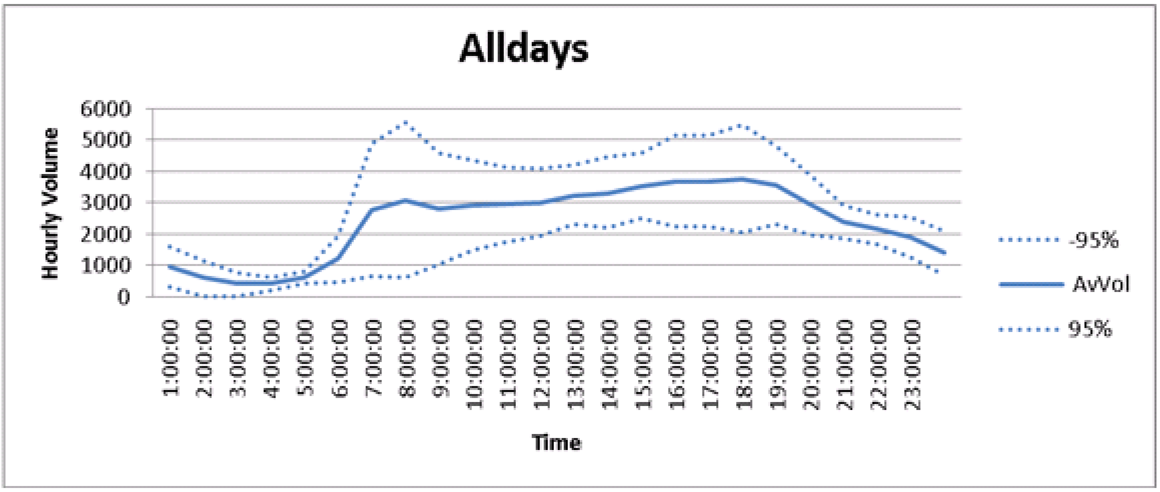


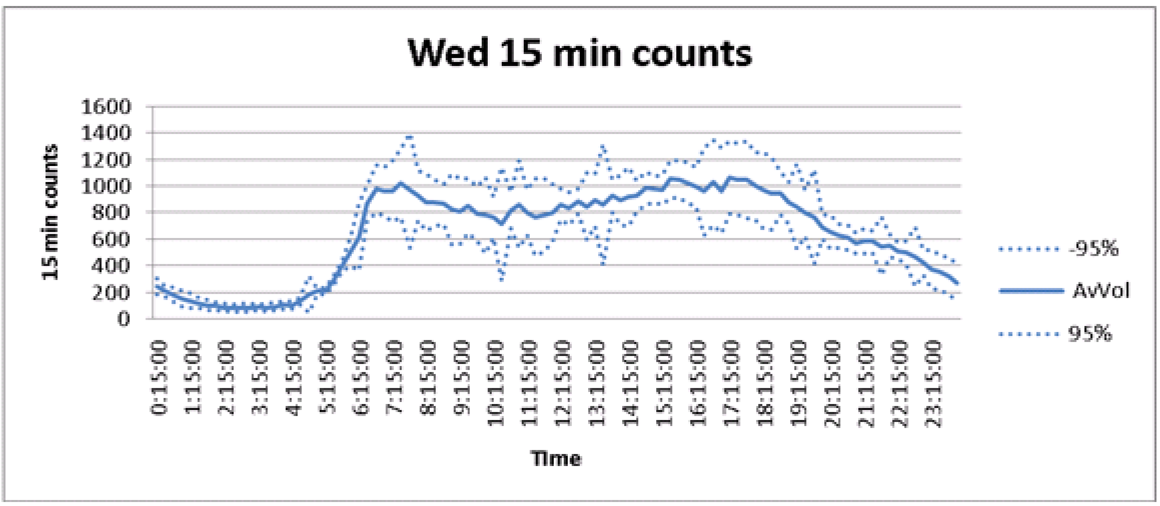
Figure 39. Overview of the ITSCounts data flow

7.2.2 Reporting of Traffic Volume Trends

Another desktop utility program called SunVol was developed to analyze traffic volume trends over a full year to examine the variability of data from day to day and to identify questionable days. Sample outputs from this utility program are illustrated in **Figure 40**. The full program documentation is included in Appendix 4c.

Plots similar to **Figure 40** have been prepared for all SunGuide sensor stations in Districts 2, 4 and 6. They may be accessed from the resources page of the STEWARD Web site.





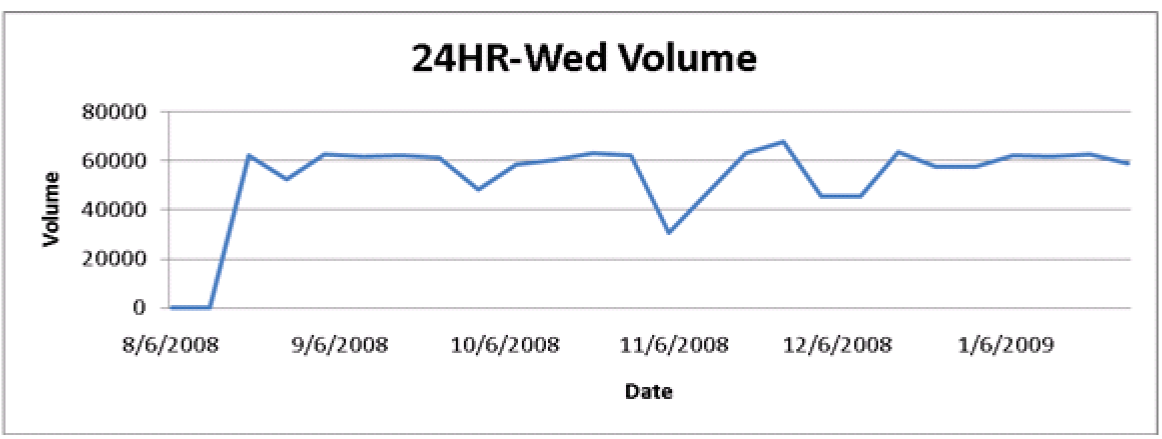


Figure 40: Sample page from a facility count analysis report

7.3 Integration with Statewide Crash Data Records

The Florida Department of Transportation maintains a crash database known as CARS, which is implemented in their mainframe computer. The CARS database includes crash information dating back to 2001. The project team was given remote access to the CARS system to retrieve the detailed output as a comma-delimited text file. Each crash record has the following 38 fields:

- Crash Report Number
- Crash Date
- Time of Crash
- DOT County Number
- Section Number
- Subsection Number
- Located Mile-point
- Nearest Node Number
- Located Route Id
- DOT Site Location
- Side of Road
- Lane of Accident
- Road Surface Condition
- Lighting Condition
- Weather Condition
- Traffic Control
- Road Conditions at Time of Crash
- Crash Rate Class Category
- Average Daily Traffic
- Crash-Level Alcohol Involved Code
- 1st Harmful Event for At-Fault Vehicle
- Vehicle Type for At-Fault Vehicle
- Vehicle Use Code for At-Fault Vehicle
- First Point of Impact for At-Fault Vehicle
- Vehicle Movement Code for At-Fault Vehicle
- Direction of Travel for At-Fault Vehicle
- 1st Contributing Cause Driver/Pedestrian for At-Fault Section
- Driver/ Pedestrian Age for At-Fault Section
- Vehicle Type for Next Vehicle
- Vehicle Use Code for Next Vehicle
- 1st Point of Impact for Next Vehicle
- Vehicle Movement Code for Next Vehicle
- Direction of Travel for Next Vehicle
- Contributing Cause Driver/ Pedestrian for Next Section
- Driver/ Pedestrian Age for Next Section
- Total Number of Vehicles in Crash
- Total Number of Traffic Fatalities in Crash
- Total Number of Injuries in Crash

The archived data offers an excellent potential for integration with the crash records. An example of the analysis of a selected crash will be presented in Section 9 of this report.

7.4 Integration with the Roadway Characteristics Inventory

FDOT maintains a comprehensive roadway characteristics inventory (RCI) database containing several descriptive fields for each roadway segment. For example, the RCI data for I-95 in Jacksonville indicates that there are 106 segments, including 3 mainline segments, 9 one-way segments and 94 ramps.

The integration of the RCI data with STEWARD is limited at this point to the provision of a field in the STEWARD facility data to indicate the RCI roadway segment for each TSS station. This will facilitate access within FDOT to the RCI Data for any SunGuide detector station. There are some possibilities for the creation of more automated links to this information. This question should be explored with STEWARD users when the user base has expanded sufficiently.

7.5 General Support for Periodic Reporting Requirements

As indicated previously in this section of the report, several performance measure reports are generated from the TSS data. When the STEWARD user base is expanded, these reports will serve to facilitate the periodic reporting requirements for the districts. Changes to the performance report content and format to meet the district expectations will probably be required.

Instead of adhering to a rigid format, all reports are now generated as CSV files that may be directly imported into office productivity programs such as Microsoft Excel. This will allow the districts to modify the actual presentation formats to meet their individual preferences.

7.6 Diagnostic Support for TMC Detector Operation and Maintenance

As indicated previously in this section of the report, several diagnostic reports are generated from the TSS data. In the initial phase of development, the data were obtained from District 2 on a more or less monthly basis. Now that the acquisition of the archived data has been streamlined and more districts have been brought on board, the diagnostic reports will be able to be generated on a schedule that will give more timely feedback to the personnel at the SunGuide TMCs. This feedback should provide useful support for the maintenance of their detector systems and communications facilities

7.7 Other Applications for the STEWARD Reports

The balance of this section will describe some current and potential real-world applications for the archive data reports. The projects covered in this section were carried out by others and are not a product of this project. They are mentioned here as a demonstration of the usefulness of the STEWARD data, keeping in mind that demonstrating the value of STEWARD was one of the stated objectives of the project.

7.7.1 Work-Zone Crash Analysis

The primary focus of this research, which is sponsored by the Southeastern Transportation Center, is to determine the impact of reduced capacity on crashes in work zone queues in Jacksonville. The University of Florida, in cooperation with the FDOT Jacksonville Traffic Management Center and the Florida Highway Patrol, is performing the study. The study focuses on the I-95 Trout River Bridge reconstruction project and the Interstate-10/Interstate-95 interchange project in Jacksonville.

Crash data were obtained from the Florida Highway Patrol and those crashes occurring in the vicinity of the identified work zones were isolated from the larger crash data set. STEWARD data were used to confirm the traffic impacts that are caused by incidents near the work area. The dates used for this project were from June 2007 to December 2007. The STEWARD data included 15-minute aggregations of traffic volumes and speeds from the stations closest to the work zone.

7.7.2 Support for Identification of Recurring Congestion

The consulting firm of RS&H is currently conducting a "Bus in Shoulder" study for the Jacksonville Transportation Authority. One of their tasks is to identify recurring congestion on I-95 in Jacksonville. Congestion has been designated in terms of speeds below 35 mph. Their initial request was for monthly station-level and lane-level, volume and speed data on I-95 in the Jacksonville area. The STEWARD Web site and documentation were provided to access and retrieve the traffic data via the STEWARD web pages.

This is an ongoing activity. RS&H has made some constructive suggestions regarding possible improvement of the data report formats to facilitate their use. It is anticipated that similar studies will be conducted in other districts.

7.7.3 Travel Time Reliability Reporting

As part of Strategic Intermodal System (SIS) management, two research projects on travel time reliability models were developed for freeways travel time reliability. The first project used data from Philadelphia, Penn., and the second project is evaluating the feasibility of using truck travel time data collected by the FHWA and the American Transportation Research Institute (ATRI) to estimate travel times and determine the travel time reliability for freeways in Florida.

The UF research team for that project obtained data for the I-95 freeway in Florida from the STEWARD Web site and are using it for model development. In addition to supporting the specific study, a continuing involvement with the supply of travel time reliability data in support of ongoing research projects and FDOT's periodic data reporting requirements is anticipated.

8 RESEARCH APPLICATIONS FOR ARCHIVED DATA

The previous section of this document dealt with operational applications for the STEWARD data. Research applications will be covered in this section. Research applications differ from operational applications in two respects:

- They tend to require data at a finer level of granularity than operational applications. Therefore they often require custom data with one-minute aggregations. Data at this aggregation level are created as a part of the ETL process but they are not stored in the STEWARD database because of storage space and access time requirements.
- They frequently need to refer to data from other sources to develop relationships with external factors such as roadway characteristics, incidents, etc.

This section considers some current examples and potential research applications for archived freeway data. The activities described in this section have been conducted in connection with other projects and do not reflect the accomplishments of the project described in this report. They are included here primarily as a demonstration of the ability of STEWARD to perform useful functions. Developing this ability was mentioned as one of the principal challenges to be addressed by the project.

The previous section introduced the various reports and materialized views available as STEWARD resources. Those resources apply equally to this section.

8.1 Analysis of Breakdown at a Freeway Ramp

The objective of National Cooperative Highway Research Program (NCHRP) Project 3-87, which started in October 2006, is to develop procedures for selecting ramp-management strategies for a freeway section under the threat of flow breakdown. These procedures will be evaluated using simulation in conjunction with field data. One of the current sites in the data collection plan will be within the District 2 SunGuide facility on Interstate 95. The archived volume, speed and occupancy data is well-suited to that project's data needs. This project provides a good example of a research application that will use short interval aggregations to model the breakdown of traffic flow on a freeway in the vicinity of an entrance ramp.

8.2 Simulation Support for SunGuide

A SunGuide simulation support project is being carried out in District 6 by a team from Florida International University under FDOT Research Project BDK80 Task Order No. 977-3. The goal of that project is to explore the development of microsimulation methods and tools to support the SunGuide system implementation, operation, testing and evaluation. STEWARD is associated with this project in two ways:

- One of the objectives of that project is to provide support for the future development and testing of STEWARD by producing data in the SunGuide archive file format based on simulation outputs.
- One of the project tasks involves the development of simulation modeling applications for SunGuide using data downloaded from the STEWARD Web site. The research team has already made extensive use of the STEWARD data for that purpose.

The STEWARD-SunGuide interface included the following components:

- A data quality check
- Daily pattern identification
- Period segmentation
- Spatial conciliation and missing volume estimation
- Free-flow speed estimation.

The FIU project team has enumerated the benefits of the STEWARD interface in the following terms:

- Significant improvement in modeling and analysis of traffic
- Lower the cost of simulation and other analysis
- Much more details in time and space
- Provide an important and a new source of data for planning and traffic analyses.

They have suggested that the interface modules can be used by a TMC for the following purposes:

- Assess normal day performance
- Assess incident performance
- Segment time of the day into intervals
- Estimate system demands

The FIU team has already developed procedures to fine-tune simulation model parameters to produce throughputs/capacities, volumes, speeds and occupancy close to those of detector data in STEWARD.

In addition to the simulation support project, researchers from Florida International University are conducting two projects that make use of data from STEWARD:

- Decision Support Tools to Support the Operations of Traffic Management Centers, sponsored by FDOT Research Center (Period: July 2008 — July 2010)
- Use of Advanced Analysis Tools to Support Freeway Corridor Freight Management, sponsored by FDOT Research Center (Period: August 2008 — March 2010)

8.3 Other TRC Research Applications

The following current, past and future projects related to STEWARD have been carried out by University of Florida researchers:

8.3.1 Modeling the Location of Crashes within Work Zones

Dr. Siva Srinivasan, PI
The objective of this study is to model the location of crashes within work zones as a function of the lengths of the different work-zone segments, traffic volume, weather and other exogenous factors. Data from crash reports were augmented with spatial attributes by using geographic information systems. The results from a multinomial logit model were used to construct the crash probabilities per lane-mile for the different work-zone segments.

8.3.2 A Case Study in Spatial Misclassification of Work Zone Crashes

Dr. Siva Srinivasan, PI
Studies associated with work zone crashes are often based on law enforcement traffic crash reports. Work zone crashes are typically segregated from larger, statewide, crash data sets by special coding within reports that describe the crash as being "in" or "near" roadway construction. The assumption is that crash report coding for the "work zone" variable are accurate, however this case study of 388 crashes in a Florida work zone finds that such an assumption may be flawed. CDW information was used to match traffic information with the work zone crash data.

8.3.3 Analyzing the Effectiveness of Enhanced Penalty Zones

Dr. Siva Srinivasan, PI
The full title of this project is "Analyzing the Effectiveness of Enhanced Penalty Zones and Police Enforcement as Freeway Speed-Control Measures." The objective is to examine the simultaneous impacts of police enforcement and increased penalties on freeway speeds and crash characteristics. The project will analyze crash, traffic enforcement and roadway traffic data from the CDW to reach the objective. (Anticipated start date January 2010).

8.3.4 Capacity of Florida Freeways, FDOT Project BDK-75-977-08

Dr. Scott Washburn, PI
CDW data are being used in an FDOT-funded project to assess the capacity of Florida freeways. In this project, speed and flow data from Jacksonville, Miami, Ft. Lauderdale, Orlando and Tampa are being used to develop speed-flow relationships and capacity distributions for a variety of basic freeway segments.

8.3.5 Travel Time Reliability

Dr. Lily Elefteriadou, PI
CDW data were used in several FDOT-funded projects on travel time reliability (BD-545-70, BD-545-75, and BDK-77-977-02). The data were used to identify areas with congestion around the Jacksonville area, and to extract speed, flow and travel time information from those locations.

8.3.6 Freeway Work Zone Capacity

Dr Lily Elefteriadou, PI
CDW data were also used in an FDOT-funded project to assess the capacity of freeway work zones (BD 545-82). In this project, speed and flow data were obtained for a work zone along I-95 in Jacksonville.

8.3.7 NCHRP 3-87, Proactive Ramp Management

Dr Lily Elefteriadou, PI
The full title of this project is "Proactive Ramp Management under the Threat of Freeway Flow Breakdown." CDW data were used during the initial stages of this project for site selection. Researchers were examining various sites around the country for obtaining speed and flow data to develop probability of breakdown models.

8.3.8 Doctoral Dissertation Project

Dr. Alexandra Kondyli
CDW data were used in Dr. Alexandra Kondyli's dissertation to identify suitable data collection locations around Jacksonville.

8.4 Assessment of Turbulence as a Predictor of Incidents

This subject was addressed as a part of the project, but it was not possible to obtain results worth reporting within the time and resource constraints. Material on this topic has been relocated to an item in the recommendations for future research.

8.5 Other STEWARD Users

The following additional current and potential users that have communicated with the STEWARD project team:

- Identification of Recurring Congestion (RS&H)
- Mobility Monitoring Program (Cambridge Systematics)
- Evaluation of DMS effectiveness for diversion (HNTB)

9 DATA ANALYSIS EXAMPLES

The use of the STEWARD data for the development of a variety of reports presenting useful performance measures was described in Section 7. This section presents examples of the use of the STEWARD data for more specific investigations. Four examples will be presented. The first compares the speed-flow-density relationships obtained from a selected station with the relationships found in the literature. The second examines the effect of a selected incident on the performance of the facility. The third deals with the extraction of measures that could be used to evaluate the performance of a managed lane on the freeway. The fourth deals with travel time reliability reporting.

9.1 Speed-Flow-Density Relationships

The literature on the relationships between speed, flow rate and density as the macroscopic descriptors of traffic flow was summarized in Section 3. The purpose of this example is to illustrate how the archived data from STEWARD may be used to estimate these descriptors and to evaluate how well the results matched those in the literature. Archived data from a selected station over a period of one month at a five-minute aggregation level will be used for this example.

The raw data from the detectors was obtained at 20-second polling intervals. The data items include a count of the number of vehicles that passed the detector during the interval, the average speed of all vehicles during the interval and the proportion of time that the detector was occupied by a vehicle. The speed, flow rate and density can be estimated from these data items as follows:

9.1.1 Speed

A distinction between time-mean speed and space-mean speed must be made at the outset. The time-mean speed is represented by a simple arithmetic mean of the individual vehicle speeds. The space-mean speed is calculated using the harmonic mean from the following equation:

$$U_s = \frac{1}{\frac{1}{N} \times \sum_i \frac{1}{U_i}}$$

Where Us is the space-mean speed
U_i is the speed of vehicle i
N is the number of vehicles observed

The space-mean speed is the appropriate choice for all computations involving the speed-flow density relationships. The speed values from the radar detectors at this location are produced by a proprietary algorithm and the exact definition is not clear. Furthermore, the individual vehicle speeds required for computation of the space-mean speed are not available. Therefore, the speed values included in the raw data must be used as the best available estimator of the space-mean speed.

9.1.2 Density

The density estimation is more complex. From the definition, occupancy is the fraction of time that vehicles are over the detector and could be described as follows:

$$Occupancy = \frac{\sum_i \frac{Li + d}{Ui}}{T}$$

Where Li and Ui are vehicle length and speed
D is detector length
T is the time interval

Using the basic macroscopic relationship

$$q = k \times Us$$

Where q is flow rate
k is density
U is the space-mean speed

The occupancy equation can be simplified as follows:

$$Occupancy = \frac{\sum_i \frac{L_i + d}{Ui}}{T} = (L + d) \times k / 5280 = C_k \times k / 5280$$

Where L is average vehicle length
d: detector length
k: density
C_k: sum of the average vehicle length and detector length

C_k is the effective vehicle length (EVL) described in the previous section. Therefore, the density could be estimated from the measured occupancy with the following equation.

$$K = \frac{5280 * Occupancy}{EVL}$$

To illustrate this point with a numerical example, let us assume that the EVL is 17.3 feet and the occupancy is 0.1. The density would then be computed as (5,280 x 0.1) / 17.3 = 30.5 veh/lane/mi.

Since density is expressed in units of vehicles per lane per mile, this is clearly a spatial measure that applies to an entire segment, whereas the field data represent the conditions at a single point. The point measurement is the only information available to support density estimates. This measure uses the same units as density and is generally referred to as “concentration.” It is common for traffic surveillance systems to use concentration as a direct estimate of density and to refer to the computed results as density.

9.1.3 Flow Rate

A distinction must be made between volume and flow rate at this point. Both measures are expressed in units of vehicles per unit of time, with a period of one hour generally used to represent the unit of time. Volume refers to the actual number of vehicles that pass the point during the hour. The flow rate term is applied to periods of less than one hour and represents a normalized value that would be obtained if the conditions persisted for an hour. So, for example, the flow rate over a five-minute period would be determined by multiplying the accumulated five-minute vehicle count by 12. The result may be applied to an individual lane or to all lanes at a station. Since the density is expressed on a per-lane basis it is necessary to express the flow rate in the same manner to preserve the speed-flow-density relationships.

9.1.4 Examples of Relationships

The selected detector station is located on Interstate 95 North of Baymeadows Road in Jacksonville. The freeway at this station carried northbound traffic in three lanes. The analysis period included all of October 2008 during the morning peak (7 a.m. — 10 a.m.). The figures presented in this section show the speed-flow-density relationships estimated from the STEWARD data at the selected location.

The relationship between the speed and flow rate is shown in Figure 41. This figure shows a uniform free flow speed around 68 mph in the undersaturated area. In the oversaturated region, the flow rates drop with speed. The capacity is reached at a flow rate of 2,236 veh/ln/h and a speed of 65.6mi/h. The density at capacity estimated by dividing the flow rate by the speed is computed as 2,236 / 65.6 = 34.1 veh/mi/lane.

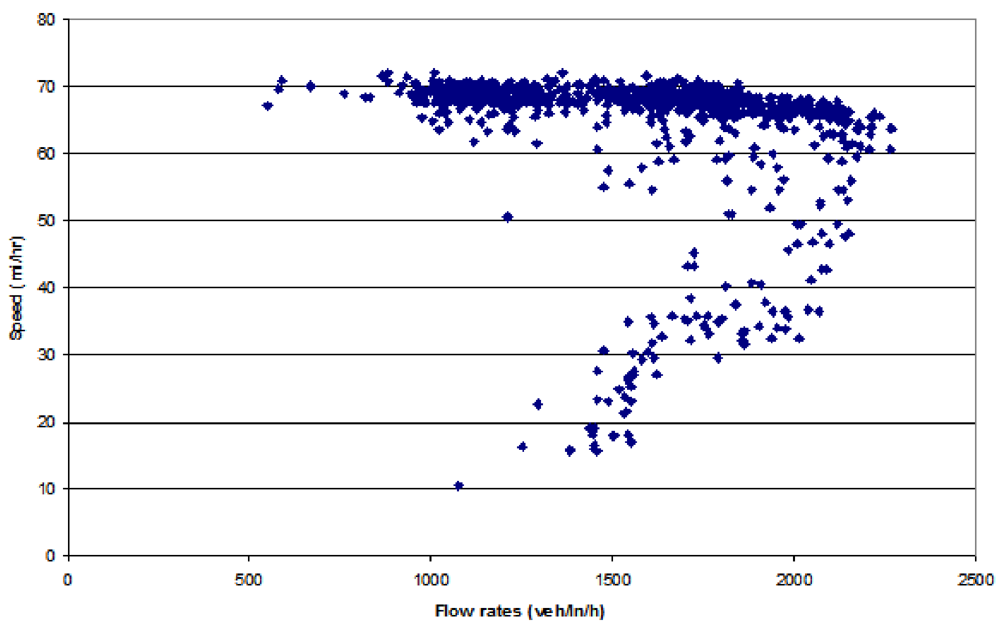


Figure 41. Example of flow rate vs. speed

Graphically, this figure compares very well with the corresponding figure from the HCM, presented previously in Figure 6. The more quantitative HCM figure presented previously in Figure 7 shows that the capacity is reached at a somewhat lower speed and density. The HCM procedure for basic freeway segment analysis suggests that the capacity of a segment corresponds to a density of 45 veh/mi/lane.

The density was calculated from the occupancy using an assumed effective vehicle length (EVL) of 17.3 feet. The relationship between flow rate and density is shown Figure 42. The trend line in the undersaturated area shows that the flow rates increase with density below 43 veh/mi/lane. The slope of the trend line represents the speed of a backward wave at a given density. The slope at a density of zero indicates the free-flow speed.

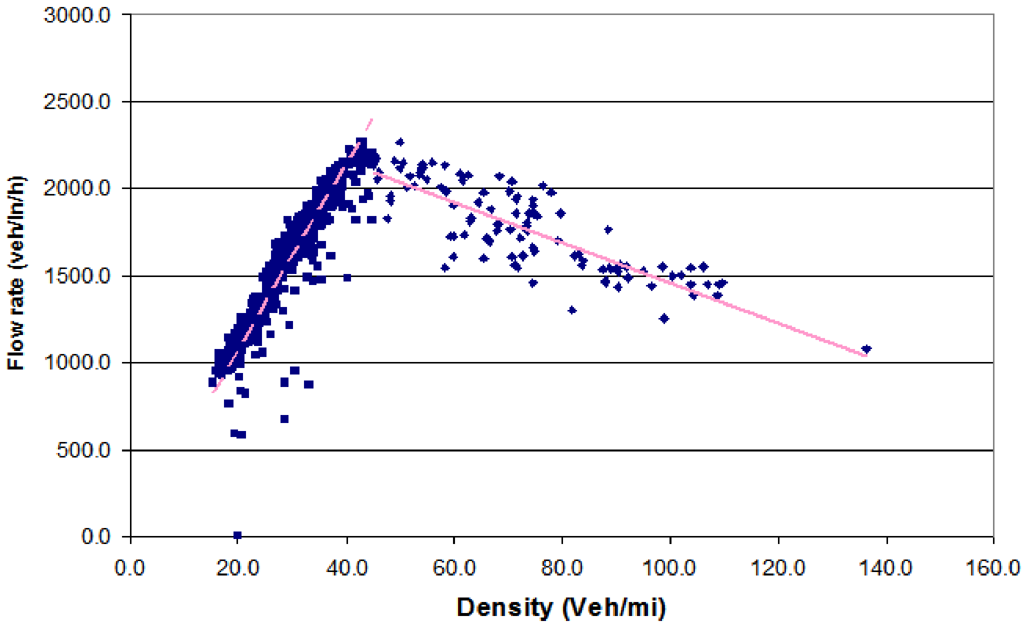


Figure 42. Example of flow rate vs. density

Since the flow-density relationship shown in the figure appears to be linear, the forward-wave speed and free-flow speed should be equal at an estimated value of 53.5 mi/h. The maximum flow rate under these conditions is 2,260 veh/ln/h at 43 veh/mi/lane. The projection of the trend line in the oversaturated area to the horizontal axis suggests a jam density of 227 veh/mi/lane. This suggests a spacing of 5,280/227 = 23.3 ft between the front bumpers of successive vehicles at jam density. The corresponding figure from the HCM, as shown previously in Figure 8 is 190 veh/mi/lane, representing a spacing of 27.8 feet. Since the highest recorded density on this figure was 140 veh/mi/lane, the linear projection of the trend line to the horizontal axis might not be appropriate.

The slope of the trend line in this area represents the backward wave speed, which is estimated at approximately 11 mph. This is slightly lower than the 16.5 mph estimate from the HCM. In comparing these values, it must be pointed out that the HCM figures represent broad national averages obtained from many locations, whereas the data presented here represent a single location. The HCM does not present any data on the variability of this value.

VMT or VHTT can be easily obtained from the STEWARD performance measures report. From these reports, the VMT and VHTT are 6,272 veh-mi and 108 veh-hr, respectively. The corresponding space-mean speed is 58.1mi/hr.

The difference between the detector data and STEWARD measures comes mainly from the definition of the link. STEWARD defines the link from the upstream station to the current station and the measures are calculated from the average volume and speed of these two stations. Figure 43 shows the speed-density relationship at this location. The density is calculated from the occupancy using a constant effective vehicle length of 17.3 feet and the speed is the measured time-mean speed.

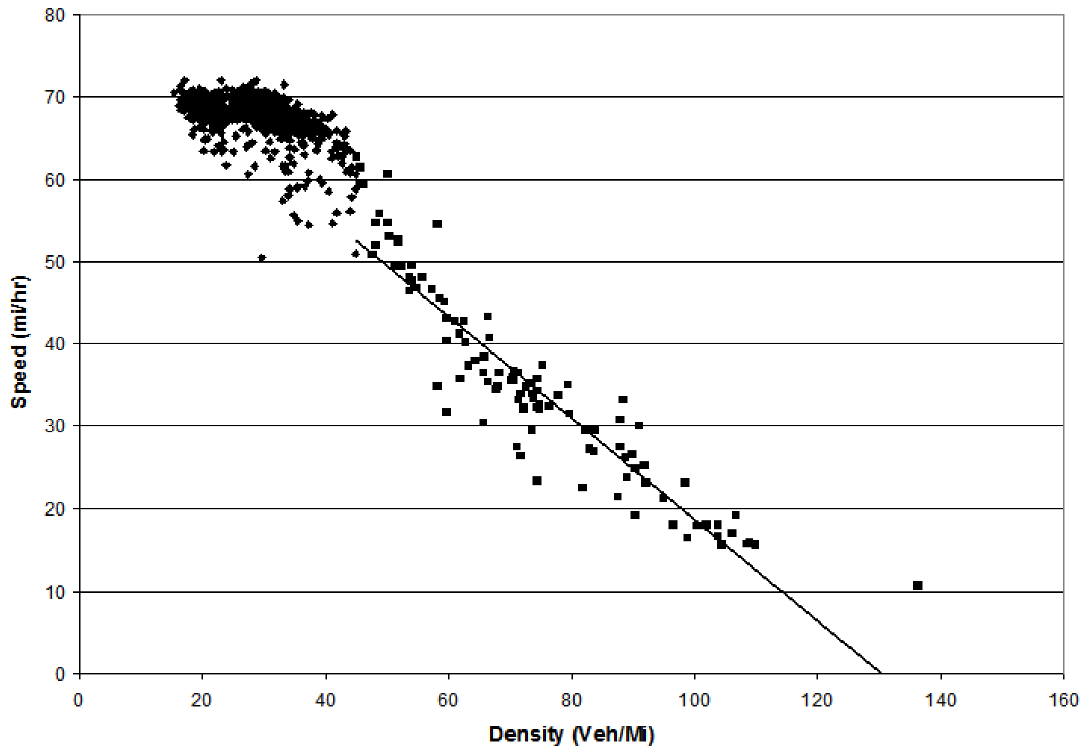


Figure 43. Example of Speed vs. density

This figure shows that, in the low-density area (uncongested region), the speed approaches the free-flow speed. As the density increases, the relationship fits better with Greenberg's logarithmic model (R^2=0.9076) than the Greenshields linear relationship (R^2=0.864). The jam density is calculated as 150.3 veh/lane/mi. The speed-density equation thus becomes

$$U = -47.109 \times \ln(\frac{K}{150.3})$$

The jam density computed here suggests a spacing of approximately 35 feet between the front bumpers of successive vehicles. This is somewhat higher than the values previously mentioned. Jam density is a somewhat abstract concept because it never actually occurs. It is simply a mathematical property of the speed-flow-density model.

From the preceding discussion on speed, flow rate and density, it is clear that density may be computed either from the flow rate and speed or from the occupancy and EVL. If the detectors are producing credible data, then these two methods should produce comparable results. The results of the two methods for the same example are shown in Figure 44. The density computed from the speed and flow rate is plotted against the occupancy values obtained directly from the detectors. Note that a tight linear relationship is demonstrated in this figure, suggesting that the detectors are producing data with a credible speed-flow-density relationship.

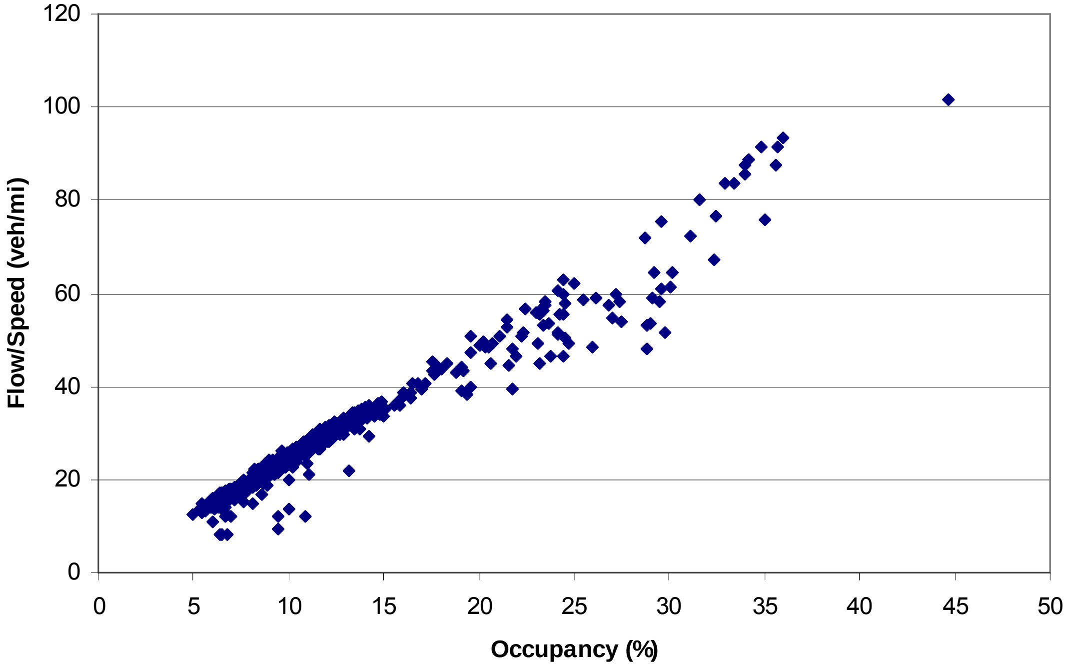


Figure 44. Comparison of density computation methods

9.2 Crash and Incident Analysis Applications

Crashes and other incidents cause perturbations that show up in the archived data. The occurrence of incidents should be evident in each of the basic archive data items, including the flow rate, speed and occupancy. Flow rates and speeds are likely to decrease during the period of an incident, and occupancy is likely to increase because of the higher traffic densities. This example will investigate the effect of a selected crash on the basic archive data and the measures that are derived from these data. The FDOT CARS system described in Section 7 was the source of the crash data

9.2.1 Overall Crash Characteristics

The sample data retrieved from CARS covered all of 2008. A total of 196 crashes on I-95 between Station Milepost: 338.0 and 348.7 were included. The locations with respect to the roadway were:

- North bound: 74
- Middle/median: 47
- South bound: 75

Figure 45 shows the number of crashes by milepost. “NB,” “SB,” and “MID” in the legend indicate the side of the road as northbound, southbound and median. The frequency of crashes was higher at specific locations, such as southbound side of state milepost 350 where I-95 merges with Acosta expressway and the northbound side of state milepost 351 where I-95 merges with I-10.

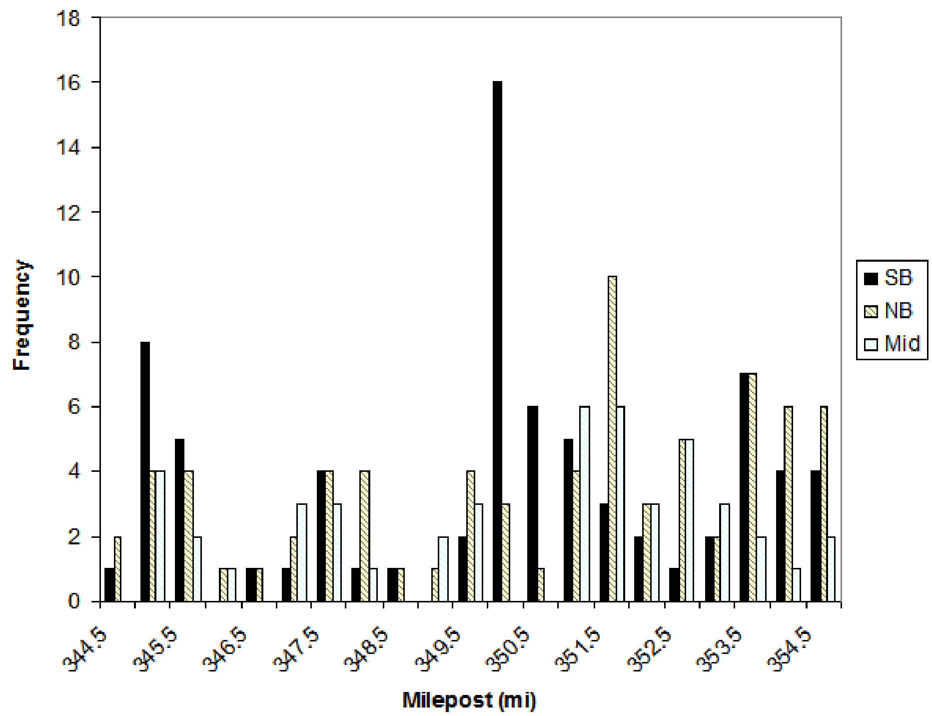


Figure 45. Number of crashes by milepost

9.2.2 Sample Crash Description

Incident data from the CARS system can be associated with STEWARD archive data to support a more detailed investigation of the effect on the operation of the facility. The selected crash had the following characteristics.

- CARS crash number: 769,954,660
- Location: Milepost 343.947
- Date: 10-3-08 12:56 p.m.
- Lane: 3
- Weather condition: Clear
- Total number of vehicles: 2
- Total number of injuries: 2

This incident occurred on a Friday at the south end of the Jacksonville downtown area. It took place on the southbound roadway. Figure 46 shows the location of incident on a satellite photo.

Hourly flow rates, speed variation, occupancy have been suggested as the most common precursors for incident analysis. In this example, hourly flow rates, occupancy, speed and speed variance will be verified for this incident. Also, the overall delay caused by this incident will be calculated by a number of methods.

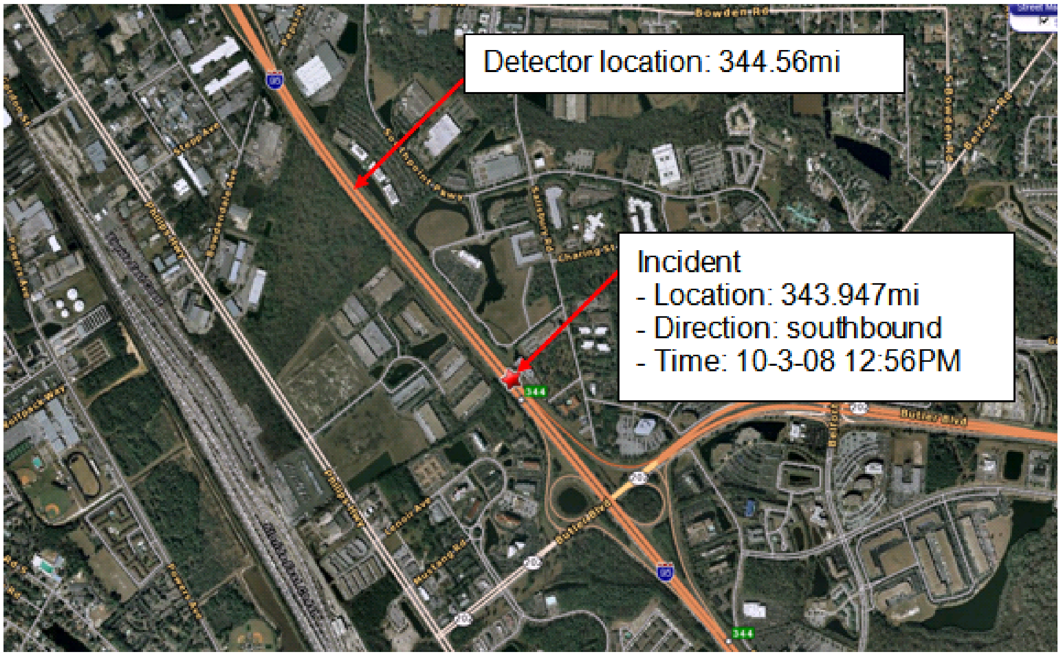


Figure 46. Incident location on a satellite photo

9.2.3 Hourly Flow Rates

Figure 47 shows the hourly flow rate per lane at milepost 344.56 mi on 10-3-08.

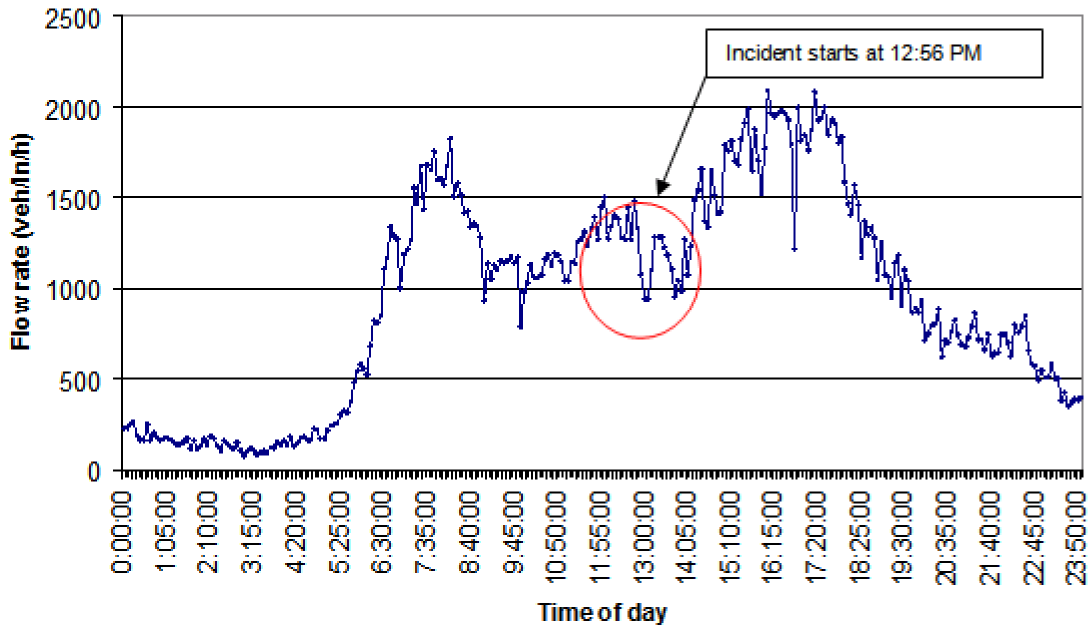


Figure 47. Hourly flow rates during the incident

CARS reported that the incident occurred at 12:56 p.m. The effect is evident on the flow rate graph, which shows a decrease near 12:40 p.m. The red circle in the figure shows the two flow rate drops during the incident, which occurred between 1:00 p.m. and 1:45 p.m., and the values were below the 1,000 veh/ln/h.

To establish the normal operation, Figure 48 shows the hourly flow rate per lane at the same location on three Fridays without incidents during the same time frame. This figure presents the average flow rates for three days (10-10-08, 10-17-08, and 10-24-08). It shows a similar trend of the flow rates except the flow does not drop at the time of the incident.

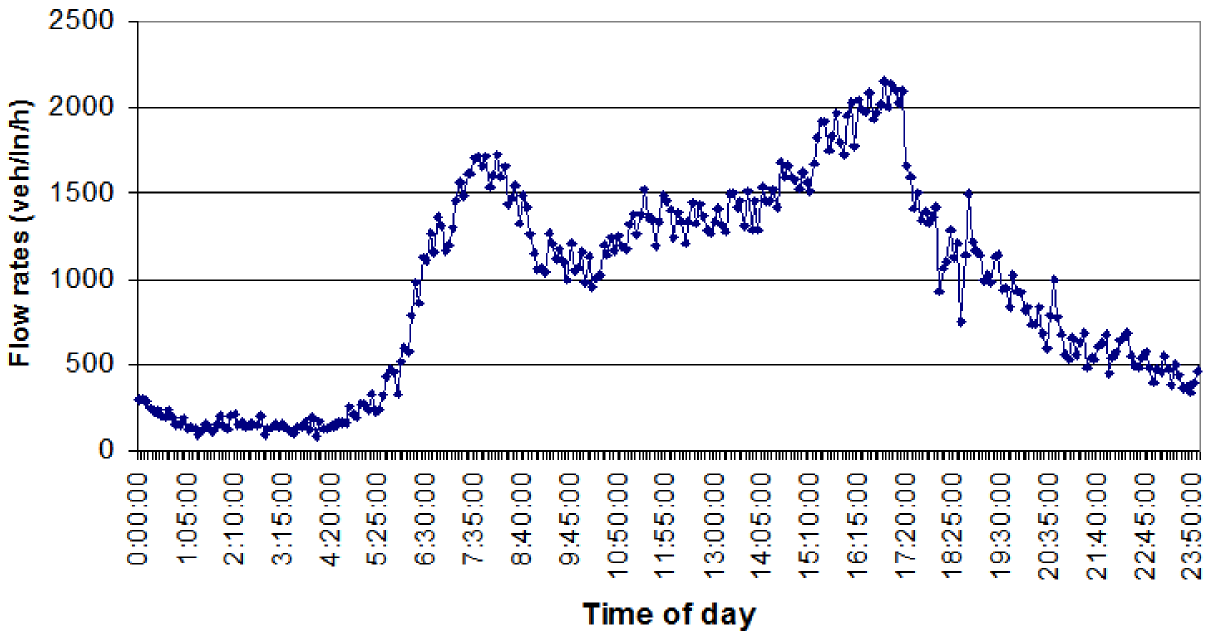


Figure 48. Hourly flow rates during incident-free operation

Figure 49 shows five-minute volume counts for the incident and non-incident cases. The differences are plotted in Figure 50 as the cumulative volume difference between the non-incident case and the incident case from 12:00. From this figure, it is observed that the incident starts at about 12:55 and the capacity decreases. Also, the queue starts to build up until the cumulative difference between the incident volume count and non-incident volume count reaches 1,058 vehicles at 14:10. After that, the queue starts to discharge until 15:15 and the cumulative differences become stabilized at approximately 900veh/lane/hr.

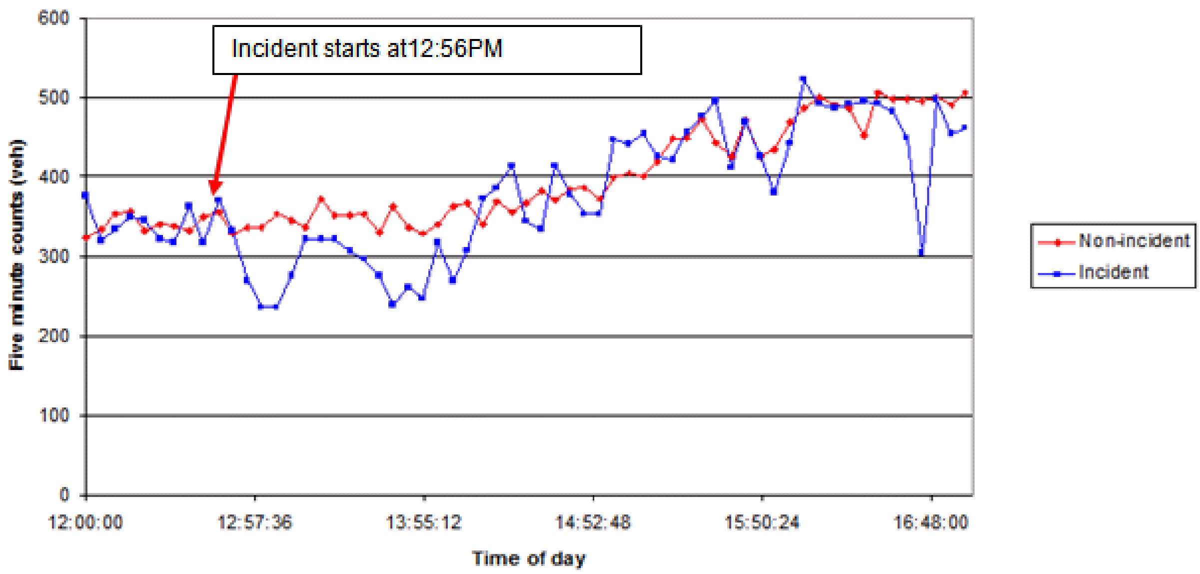


Figure 49: Comparison of five-minute counts for the incident and non-incident case

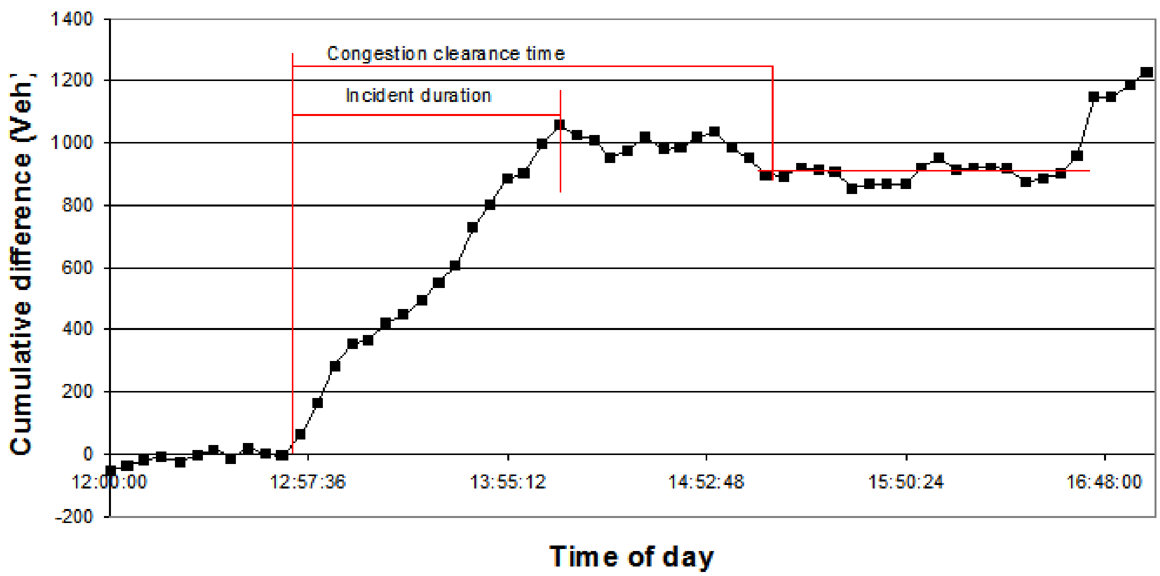


Figure 50. Cumulative differences between incident and non-incident volume counts

From this information, the delay can be estimated using queuing analysis.

Mean arrival rate (λ) = 9,610veh / 145min
= 3,977veh/hr

Reduced service rate (μ_R) = 4,498veh / 80min
= 3,374 veh/hr

Service rate (μ) = (9,610veh – 4,498veh) / 65min
= 4,719 veh/hr

Time duration of queue (t_Q) = $\dfrac{t_R \times (\mu - \mu_R)}{(\mu - \lambda)}$

= 80 x (4,719-3,374) / (4,719-3,977)
= 145min

Total delay = $\dfrac{t_R \times t_Q \times (\lambda - \mu_R)}{2}$

= 80 min x 145 min x (3,977 – 3,374) / 2
= 971.5 veh/hr

Average vehicle delay = (971.5 veh-hr / 9610 veh)
= 6.1 veh-min/veh

9.2.4 Occupancy

Figure 51 shows an abrupt occupancy increase at the same location, milepost 344.56 mi. Between 12:45 p.m. to 2:30 p.m., the occupancy increases over 45 percent.

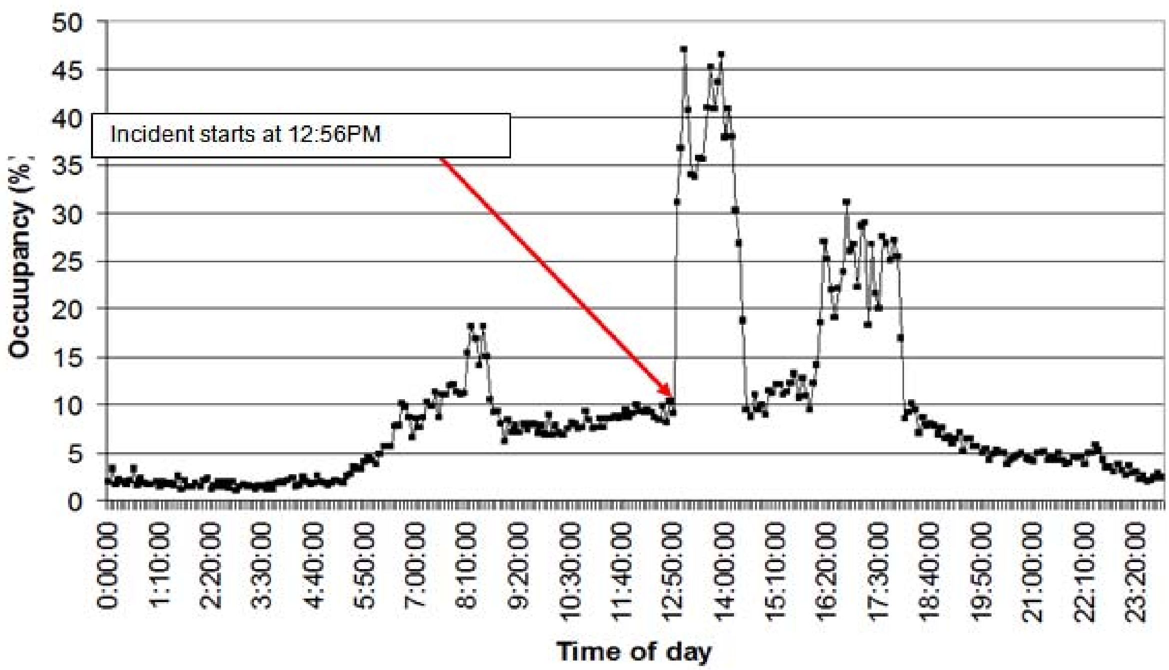


Figure 51. Effect of the incident on occupancy

Figure 52 presents a contour graph to show the changes in the occupancy over the time and space. The x-axis shows the milepost and the y-axis shows time of day. The incident started near the red circle (milepost 344 at 12:40) and the occupancy increases are evident on this figure.

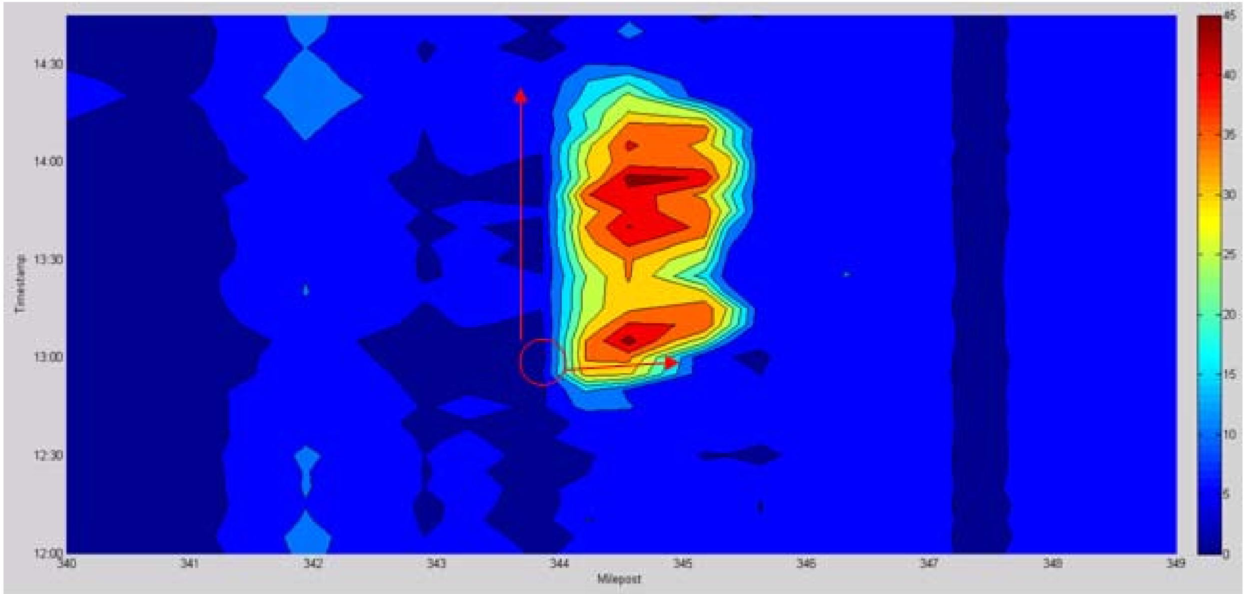


Figure 52. Occupancy contour graph during the incident

9.2.5 Estimation of Delay from Speed

Figure 53 shows the speed changes at the same location. Between 12:45 to 14:30, there were speed drops up to 3 mph that lasted until 14:00.

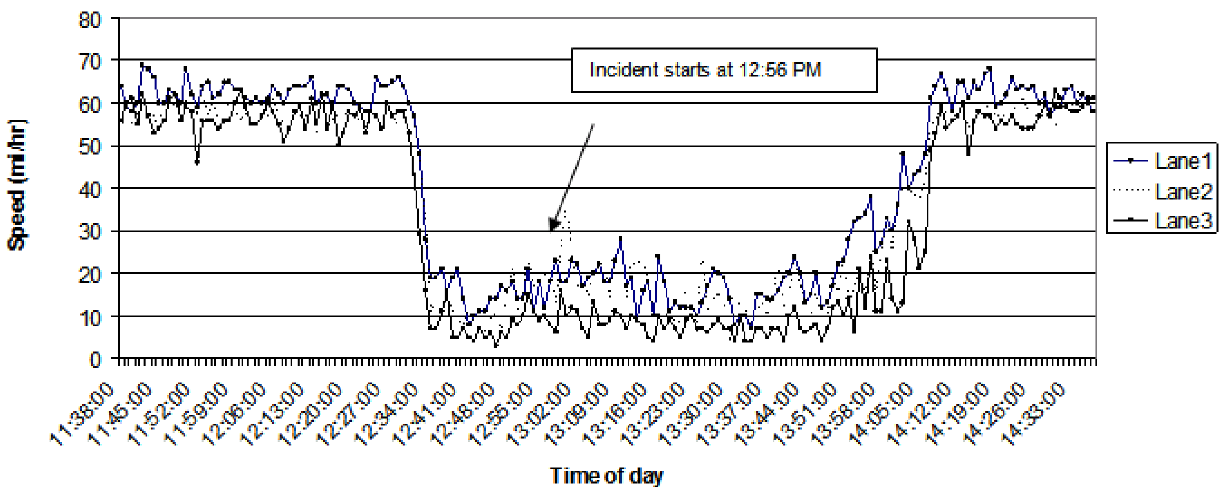


Figure 53. Effect of the incident on speed

As demonstrated previously, density can be estimated from the occupancy data. Using the density, flow and speed data, total delay and average individual delay could be estimated with shock wave analysis. Figure 54 shows the speed variation over the time and space and shock waves can be observed. A backward-forming shock wave is shown as W_{AB}, and a forward recovery shock wave is shown as W_{BD}. This incident shows a temporary capacity increase around 13:00, which creates a temporary backward shock wave (W_{BD1}) and forward recovery shock wave (W_{AB1}) in the middle of the incident. By definition, the area of a time-space domain of congestion multiplied by the density of the traffic flows under congestion is the total vehicle-hours of travel in congestion.

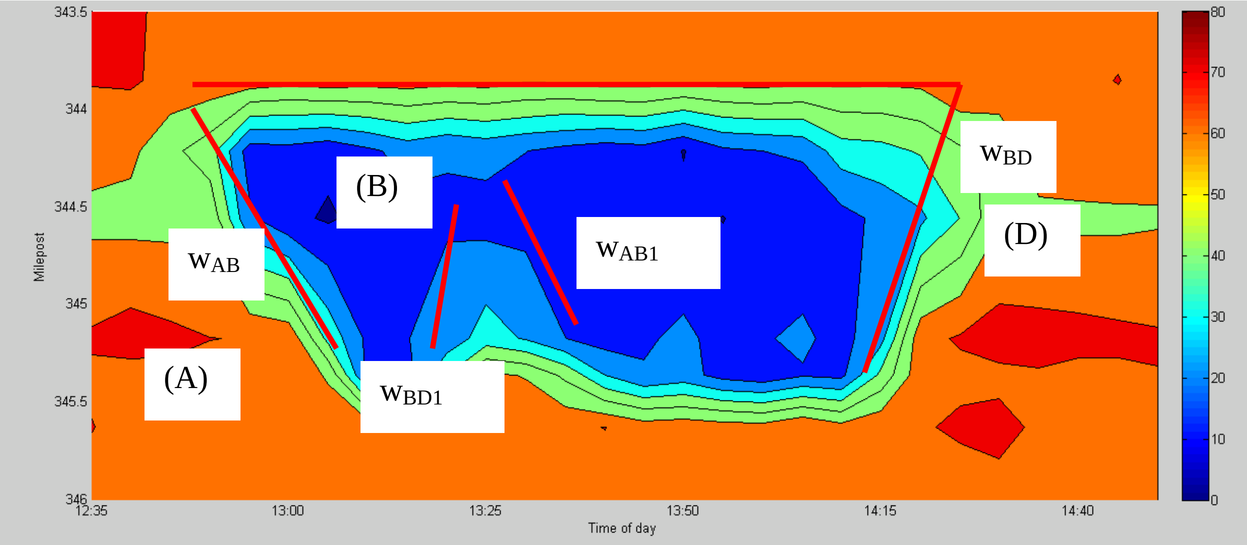


Figure 54. Speed contour graph during the incident

Total delay can be calculated from the difference between the travel time without incident and the travel time with incident. For the calculation of the travel time without the incident, averages of the three-day density for each station are used.

Total travel time = (Density x (distance x time))
= 642 veh-hr

Total delay = (travel time$_{\text{non-incident}}$ -Total travel time$_{\text{incident}}$)
= 510 veh-hr
Average vehicle delay = (delay/flow)
= 9.03veh-min/veh

9.2.6 Estimation of Delay from the Travel Time Reliability Report

The total delay can be also calculated from the Travel Time Reliability Report. Congestion delay is referenced for purposes of that report to a travel time index of 1.5. The travel time index is defined as the ratio of the actual travel time to the travel time at the free-flow speed. The speed limit will be used to represent the free-flow speed. The unit of measurement is accumulated minutes of delay.

From the definition, congestion delay is calculated as follows:

Total congested delay = distance * volume * (1/speed – 1.5 / (speed limit))
(When speed is less than (speed limit)/1.5)
= 159.8 (veh-hr)

But the total delay consists of the delays from the incident (non-recurring) and everyday congestions (recurring). Non-recurring delay can be estimated from the average delay from the same location without incidents. The same section does not have an incident from 10/10/08 to 1/24/08 on every Friday. The daily average congestion delay is 10.4 (veh-hr) and it can be assumed as the recurring delay.

The delay from the incident (non-recurring delay) would be calculated from the difference of the total delay and the recurring delay:

Non-recurring delay = total congested delay - recurring delay
= 159.8 - 10.4
= 149.4 (veh-hr)

Average vehicle delay = (971.5 veh-hr/ 9,610 veh)
= 1.6 veh-min/veh

9.2.7 Comparison of Delay Estimation Methods

Three different methods were applied to estimate the total delay and the average vehicle delay. Each method has a sound theoretical basis but substantial differences in their results were observed, largely due to differences in assumptions and definitions. It is clear that the travel time reliability method estimates the smallest delay of the three methods but it bases its computation on a different definition of delay because it ignores any delay that occurs when the speeds are less than the free-flow speed, but greater than two-thirds of the free-flow speed. It also covers a time period that extends beyond the incident on both sides because of the need to start and end the analysis at the beginning of an hour for purposes of this report. The speed/shock wave method estimates the highest delay because it focuses exclusively on the time period in which delays were observed.

The purpose of this exercise was to identify the various approaches to computing incident delay from the archived data. A substantial effort beyond the scope of this project would be required to provide useful guidance on their relative merits.

9.2.8 Lane Volume Balance Ratio

Figure 55 shows the LVBR during the incident in the form of a contour plot. The x-axis shows the milepost and the y-axis shows the time. Several peak values of LVBR are found during the incident near milepost 344 mi. At some points, the LVBR approaches 8.0, indicating a severe unbalance in the lane utilization.

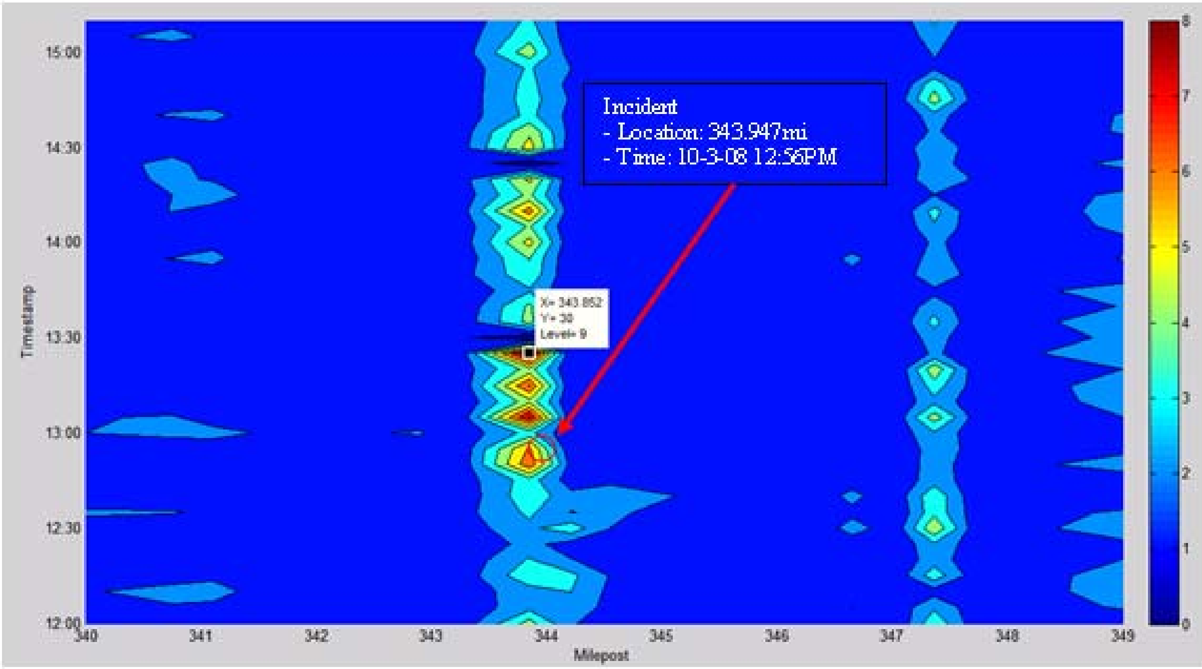


Figure 55. Lane volume balance ratio during the incident

These larger LVBR values would show that this incident didn't block the entire roadway and the vehicles were moving through some of the lanes. This condition could be verified from Figure 56.

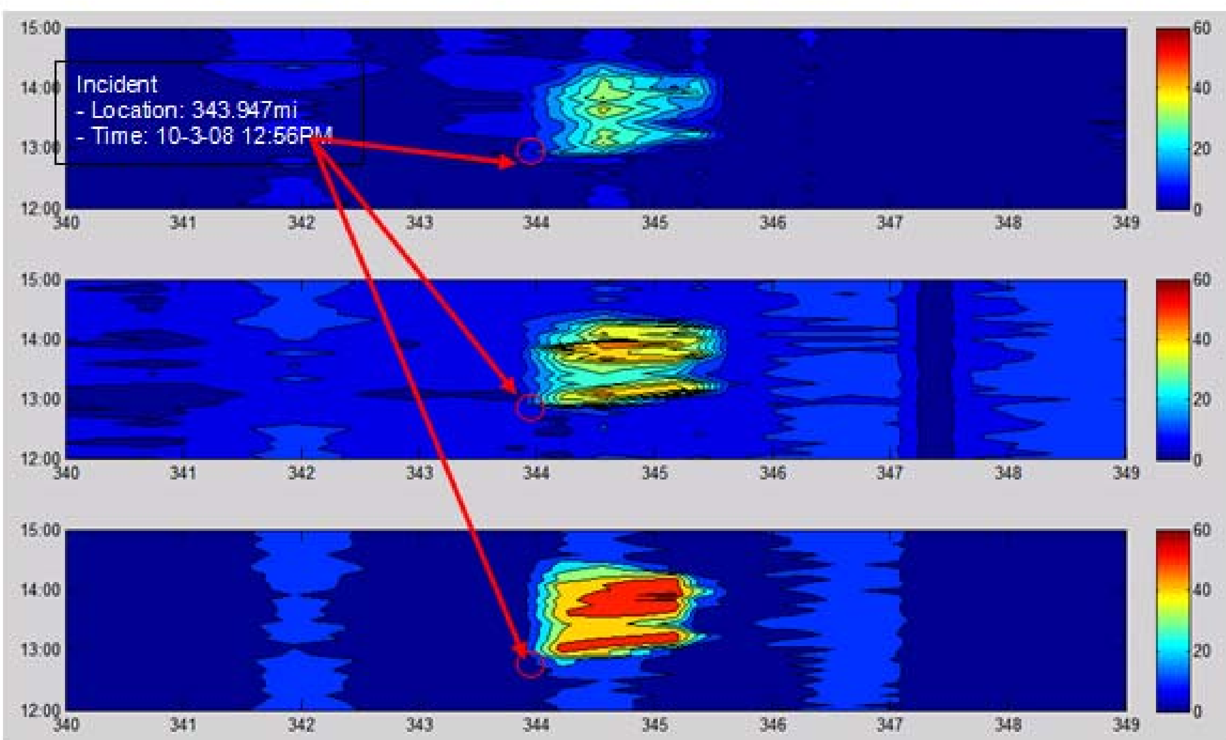


Figure 56. Occupancy contours by lane during the incident

Three contour graphs represent the occupancy in each of the three lanes. The x-axis shows the milepost and the y-axis shows the time. The vehicles are moving from right to left (decreasing milepost). As the figure shows, the incident occurs in Lane 3 (rightmost lane) and therefore, the occupancy values from Lane 1 and Lane 2 are less affected than Lane 3.

9.2.9 Speed Variance

The speed variance is also a potential indicator of the incident. Figure 57 shows the speed coefficient of variation (CV) at the incident location. It shows peak values around 8:30, 12:50, 16:00 and 18:00, but there is only one incident record in CARS in that area on the day in question. Inspection of the Florida Highway Patrol incident log indicated four incidents on that day, but none of them could be expected to affect this location.

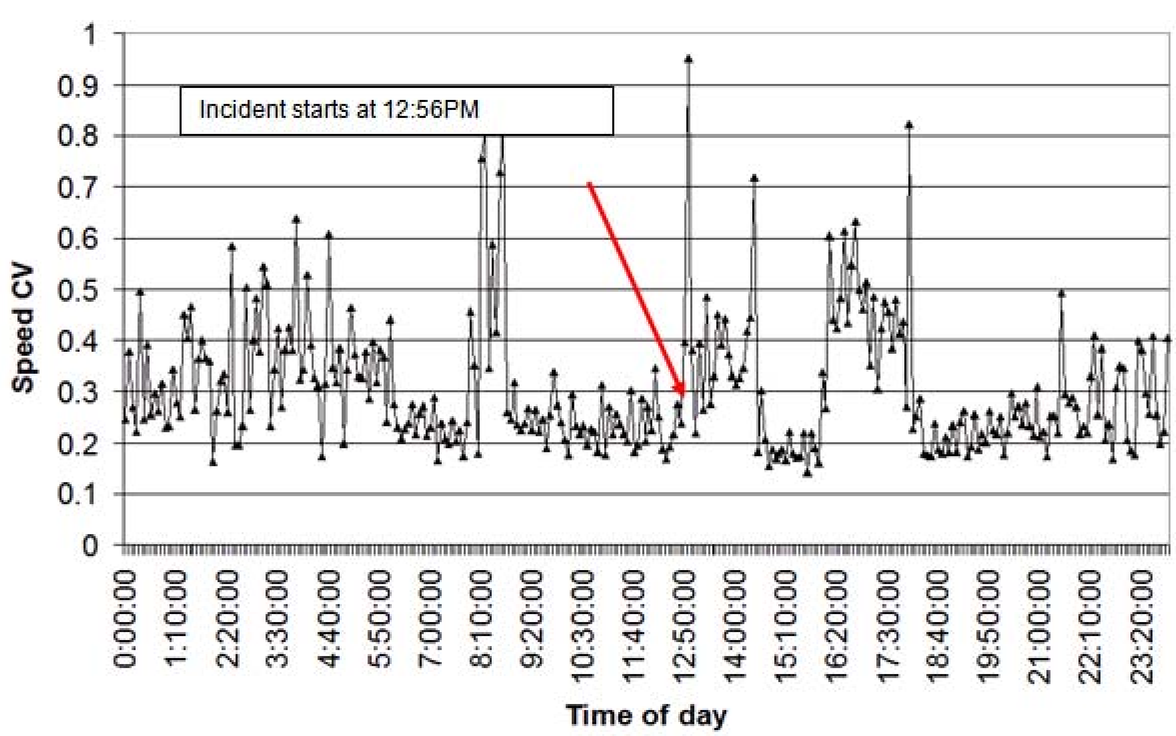


Figure 57. Speed variation by time of day

Speed variance is a measure that is unique to STEWARD and is not used for operational analysis. It was created mainly to support future research into congestion modeling and incident analysis. It certainly provides an indication of some type of perturbation in the traffic stream. No specific conclusions on its value can be drawn from this single example. The speed CV definitely peaked during the incident, but similar peaks were observed at other times of the day that were incident-free. It is interesting to note that some of the other peaks occurred during the a.m. and p.m. peak periods. There is a chance that peaks in the speed CV are simply a natural phenomenon that is associated with the onset and resolution of congestion. It appears, however, that a substantial investigation well beyond the scope of this project would be required to support definitive conclusions.

9.2.10 Kinetic Energy

The concept of kinetic energy as the product of speed and flow rate was introduced in Chapter 4. Like speed variance, this measure is not used by SunGuide for operational analysis, but it was incorporated into STEWARD to support future research. Because it reflects changes in both flow rate and speed, it offers greater sensitivity to traffic stream perturbations.

The sensitivity of kinetic energy is illustrated in Figure 58, which shows the variation in volume and kinetic energy throughout the period of the incident. It is clear from this figure that the kinetic energy dropped much more dramatically than the volume, reaching a level near zero during the incident. Other less significant variations were observed outside of the incident time range.

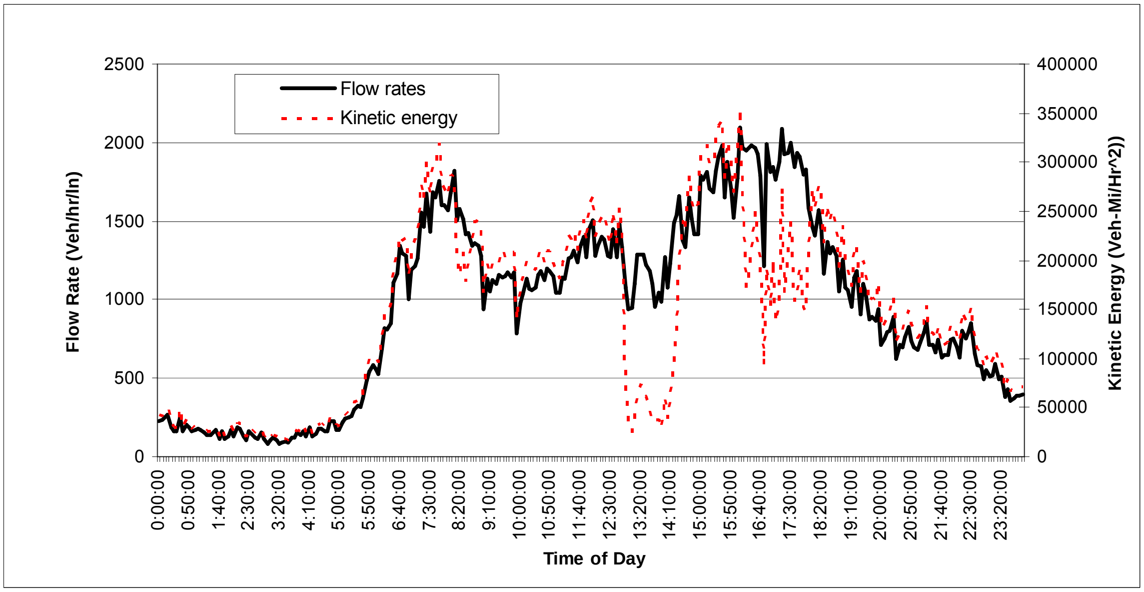


Figure 58. Flow rate and kinetic energy at the incident by time of day

No quantitative conclusions can be drawn from this one example, however, it can be said that kinetic energy has a potential application for screening the archived data over long periods of time to isolate periods of perturbation for further analysis.

9.3 Analysis of Managed Lanes

The concept of managed lanes is gaining popularity in congested urban areas. Managed lanes fall into two categories:

- High Occupancy Vehicle (HOV) lanes in which the use is restricted to vehicles with a specified minimum occupancy
- High Occupancy Toll (HOT) lanes in which a toll fee is charged to all vehicles using the lane

HOV lanes have been in use for several years. HOT lanes are a more recent concept, largely because they require an ITS infrastructure to support their use. Both types of managed lanes are in use in SunGuide systems in Florida. The archive data available from STEWARD offers an excellent potential for studying the effectiveness of managed lanes. This section explores the measures that can be computed for this purpose and presents an example of a managed lane analysis.

The procedures for evaluating the performance of managed lanes were incorporated into a utility program called "High Occupancy Toll Traffic Evaluation Report" (HOTTER), which can analyze a specified section of roadway to provide performance information for the managed lanes and the general use lanes. The operation of this utility program is described in the STEWARD Final Report Appendix 4d.

The archived data for analysis are downloaded from the STEWARD Web site in the manner described previously in this report. The required user specified parameters include:

- The lane numbers for the managed lanes
- The type of lane (HOT or HOV)
- The passenger occupancy for HOV and general use lanes
- The cost to the motorist per vehicle mile for HOT lane use
- The spatial and temporal limits of the analysis

9.3.1 Managed Lane Performance Measures

The basic performance measures obtained directly from the data include:

- Managed Lane Volume, V_m
- Managed Lane Speed, S_m
- General Lane Volume, V_g
- General Lane Speed, S_g
- Vehicle Speed Difference, $D_s = S_m - S_g$
- Vehicle Speed Ratio, $R_s = S_m / S_g$

With information on the facility length the following performance measures can be derived:

- Vehicle miles traveled in the managed lanes, $VMT_m = L * V_m$
- Vehicle miles traveled in the general lanes, $VMT_g = L * V_g$
- Total vehicle miles traveled, $VMT = VMT_m + VMT_g$
- Travel time per vehicle in the managed lanes, $TTV_m = L / S_m$
- Travel time per vehicle in the general lanes, $TTV_g = L * S_g$
- Vehicle hours spent in the managed lanes, $VH_m = TTV_m * V_m$
- Vehicle hours spent in the general lanes, $VH_g = TTV_g * V_g$
- Total vehicle hours spent, $VH = VH_m + VH_g$
- Average vehicle speed for all lanes, $VS = (VMT_m + VMT_g) / (VH_m + VH_g)$

The above measures may be most usefully applied to before and after situations to determine changes in facility productivity resulting from the managed lane operation. The operational effectiveness of a managed lane may also be assessed in an absolute sense (i.e., without a before-and-after study) by comparing the average vehicle speeds and travel times in the managed lanes and the general lanes. The following measures may be obtained:

- Travel time difference, D_{tt} = TTV_g - TTV_m
- Speed difference, D_s = $S_m – S_p$
- Speed Ratio, R_s = S_m / S_p

Negative values in the differences or ratios less than 1.0 would indicate that the operation in the HOV lane was worse than the general lanes.

Additional performance measures that could be computed for HOT lanes, based on the price per vehicle mile, PVM, include:

- Cost per vehicle-hour saved, CVH = PVM * L /D_{tt}
- Revenue, R = PVM * VMT_m

Additional performance measures that could be computed for HOV lanes, based on the passenger occupancy in the managed lanes and the general lanes, PPVm and PPVg include:

- Passenger miles traveled in the managed lanes, PMT_m =VMT_m * PPV_m
- Passenger miles traveled in the general lanes, PMT_g =VMT_g * PPV_g
- Total passenger miles traveled, PMT = PMT_m + PMT_g
- Travel time per vehicle in the managed lanes, TTV_m = L/ S_m
- Travel time per vehicle in the general lanes, TTV_g = L * S_g
- Passenger hours spent in the managed lanes, PH_m = VH_m * PPV_m
- Passenger hours spent in the general lanes, PH_g = VH_g * PPV_g
- Total passenger hours spent, PH = PH_m + PH_g

The average passenger speed, PS, for the facility may be computed as a passenger occupancy-weighted average of the vehicle speeds in the managed lanes and the general lanes. An increase in speed for high occupancy vehicles, coupled with generally higher vehicle occupancy should increase the average passenger speed to a level greater than the average vehicle speed, VS. The relationship between vehicle speeds in the HOV lanes and the general lanes provides an indication of the advantage given to the HOV lanes at the expense of the general lanes. It does not necessarily reflect the overall value of the HOV lane to the transportation system. For example, an HOV lane that accommodates little or no traffic would provide a great advantage to its occupants but would be of limited value to the transportation system.

The relationship between the average passenger speed and the average vehicle speed on the facility offers a better measure of the value of the HOV lane operation because it also reflects the degree of utilization of the HOV lane, in terms of both the traffic volumes and the passenger occupancy levels. For purposes of this discussion, measures based on this relationship will be defined as HOV performance measures. The following measures may be computed:

- HOV performance difference, PS-VS, expressed in mph
- HOV performance ratio, PS/VS

Both measures reflect the degree to which the average passenger is moving faster than the average vehicle. If there is no difference in the two speeds, then it is difficult to argue that the HOV lane provides any value to the transportation system.

9.3.2 Example Results

The performance measures have been incorporated into an experimental version of HOTTER to demonstrate their potential. There is very limited experience with their application at this point. A data set from a section of I-95, which now includes and HOV lane in District 4 was selected for demonstration. The common performance measures are shown below:

- Vehicle-miles traveled in the managed lanes: 25,145
- Vehicle-hours spent in the managed lanes: 372
- Average speed in the managed lanes (mph): 67.5
- Vehicle-miles traveled in the general lanes: 143,069
- Vehicle-hours spent in the general lanes: 2,374
- Average speed in the general lanes (mph): 60.3
- Vehicle-miles traveled in all lanes: 168,214
- Vehicle-hours spent in all lanes: 2,746
- Average vehicle speed in all lanes (mph): 61.2
- Travel time per vehicle in the managed lanes (min): 22.4
- Travel time per vehicle in the general lanes (min): 25.1
- Travel time difference (min); 2.7
- Vehicle speed difference (mph): 7.28
- Vehicle speed ratio: 1.12

In the absence of empirical data, it was assumed that the HOV lane had an average occupancy of 2.1 PPV and that the general lanes had an average occupancy of 1.2 PPV. The HOV operational analysis results were as follows:

- Passenger miles traveled in the managed lanes: 52,803
- Passenger miles traveled in the general lanes: 171,682
- Total passenger miles traveled: 224,486
- Passenger hours spent in the managed lanes: 781
- Passenger hours spent in the general lanes: 2,848
- Total passenger hours spent: 3,630
- Average passenger speed for the facility (mph): 64.89
- HOV performance difference (mph): 3.65
- HOV performance ratio: 1.06

There are currently no HOT lane facilities providing data to STEWARD. Therefore, to demonstrate the HOT lane analysis capabilities of HOTTER, it was hypothetically assumed that the HOV lane was instead a HOT lane with a pricing of $1.00 per trip. The results indicated a cost of $22.19 per vehicle-hour of travel time saved in comparison with the general lanes.

HOT lanes would normally be expected to offer a substantially greater travel time difference to attract participation by the motorist. Since this example is hypothetical, the only conclusion that can be drawn is that the speed difference associated with the HOV operation would not be worth $1.00 to many drivers. The main purpose for including the example was to illustrate the potential to evaluate a real HOT lane from the STEWARD data at some point in the future.

These examples demonstrate the ability to produce potentially useful results; however, more experience with this application in addition to stakeholder feedback will be required before meaningful application guidelines can be developed.

9.4 Travel Time Reliability

Travel time reliability and the measures by which it can be assessed have been mentioned throughout this report. The need for reporting of travel-time-related measures was introduced in Chapter 4. The STEWARD report that presents these measures was described in detail in Chapter 7. A number of research projects using STEWARD data for dealing with travel time reliability were summarized in Chapter 8. The use of the travel time report for assessing incident delay was discussed earlier in this chapter.

Because of the importance of travel time reliability assessment, an example will be presented here using data from STEWARD. A southbound section of Interstate 95 in Jacksonville between the entrance from Emerson Street and the entrance from WB Butler Blvd. will be used to demonstrate the travel time reporting features. The data sample covered the period 4 p.m. to 6 p.m. for all weekdays in 2008. A total of 253 days are represented in this example. Not all stations reported valid data for the entire period. The number of valid days per station ranged from 59 to 249. The average number of days of valid data per station was 211. This relatively short section (3.58.mi) was chosen for demonstration to simplify the discussion.

The travel time reliability table produced by STEWARD for this example is presented in Table 29.

Table 29. Results for the Travel Time Reliability Example

Segment	Av Speed	Travel Time Av	Travel Time 95%	TT Index	% On Time	Buffer Index	Ontime Delay	Congestion Delay
Units	mph	(Min/Veh)					(Min/Veh)	
Entrance from Emerson								
North of Spring Glen Rd	58.07	0.44	0.68	1.19	69.76	0.56	0.05	0.02
South of Spring Glen Rd	57.34	0.4	0.74	1.28	65.4	0.86	0.07	0.04
South of University Blvd	56.46	0.39	0.65	1.26	64.03	0.69	0.05	0.03
Exit to University Blvd EB	60.81	0.37	0.54	1.12	75.5	0.49	0.03	0.01
Exit to University Blvd WB	53.68	0.34	0.63	1.39	55	0.85	0.07	0.04
Between University and Bowden	50.63	0.26	0.48	1.49	52.03	0.86	0.07	0.05
Entrance from Bowden	44.54	0.58	1.12	1.93	39.45	0.92	0.23	0.17
South of Bowden	41.35	0.47	0.75	1.76	30.65	0.58	0.16	0.11
North of Butler Blvd	54.86	0.39	0.5	1.22	46.18	0.27	0.03	0
Exit to Butler Blvd	66.55	0.33	0.36	0.98	98.67	0.09	0	0
Entrance from Butler WB	73.47	0.15	0.16	0.89	99.44	0.09	0	0
Entrance from Butler EB	66.49	0.38	0.41	0.99	96.91	0.1	0	0
Totals:	N/A	4.5	7.02	1.29	66.1	0.53	0.76	0.47

Most of the segments in this section were relatively congestion free, but some congestion and delays may be observed in the segments near the center of the section. The following measures are presented in the table for each segment:

- *Average Speed:* The segment speeds ranged from 41.36 to 73.47 mph
- *Average Travel Time:* The segment travel times ranged from 0.15 to 0.44 min/veh. The average travel time for the section was 3.58 minutes.
- *95 Percentile Travel Time:* This value was based on the travel time distribution. The values ranged from 0.16 to 1.12 min/veh. The 95-percentile travel time for the section was 5.36 minutes.
- *Travel Time Index:* This value represents the ratio of the average travel time to the travel time at the speed limit. The maximum value for any segment was 1.93. Note that some values fell slightly below 1.0, indicating that the actual speeds exceeded the speed limit by a small amount. The overall travel time index for the section was 1.29, indicating a moderate level of congestion.
- *Percent on Time:* This value represents the percent of vehicles that were able to make their trip within 10 mph of the speed limit. This value varied by segment from 30.65 percent to nearly 100 percent with the higher values at the south end of the section. The average value was 66.1 percent, indicating that about one-third of the vehicles were not able to complete their trips within 10 mph of the speed limit.
- *Buffer Index:* The buffer index is defined in Table 7-8 as (95 percent Travel Time – Average Travel time) / Average Travel time. It is intended to convey the amount of extra time a person would have to allow to be 95 percent sure of being able to make the trip within the allotted time. The buffer index for the section is 0.53. Therefore, based on an average travel time of 4.5 minutes, a person would have to allow 4.5 * (1+0.53), or 6.88 minutes for this portion of the trip.
- *On Time Delay:* This value represents the extra time spent in the section over and above the time that would be spent at 10 mph below the speed limit. The total for the section is 0.76 minutes.
- *Congestion Delay:* This value represents the extra time spent in the section over and above the time that would be spent at a travel time index of 1.5. The total for the section is 0.47 minutes. The congestion delay is lower than the on time delay because it is referenced to a lower speed. Congestion delay can generally be taken as an indication of a capacity deficiency, whereas on time delay is considered to be more related to driver satisfaction.

10 CONCLUSIONS AND RECOMMENDATIONS

Based on the results of this study, the following conclusions and recommendations are offered:

10.1 Conclusions

The goals of the project described in this report were 1) to design a data archiving system capable of producing a set of useful reports and 2) to demonstrate the value of the system to researchers and practitioners. Both of these goals have been met.

This project has created an important resource for a wide variety of traffic data users in Florida, including both practitioners and researchers. The Web site developed as a part of the project provides the capability to download several reports summarized over a range of temporal and spatial requirements. The data can serve a variety of purposes:

- Identification of detector malfunctions
- Calibration guidance for detectors
- Quality assessment tests on data
- Development of daily performance measures
- Fulfillment of periodic reporting requirements
- Evaluation of special projects, such as managed lanes
- Provision of data for research and special studies

There are several projects and activities that have already benefited from the available data. As noted in the report, University of Florida and Florida International University researchers have already made good use of the data. The Web site has shown continued high level of activity. It is anticipated that activity levels will increase as more data become available and awareness of the STEWARD capabilities increases.

While the system implementation schedules in the districts created some delays in the provision of archive data, cooperation at the district level was excellent, and was a strong factor in the success of the project. A fully functional scheme is now in place for automated transmittal and processing of archive data. It appears that feedback from the project team to the districts and to the SunGuide contractor was helpful in resolving some technical issues with SunGuide.

The diagnostic reports furnished to the districts should be valuable in the maintenance of their detector systems. These reports indicate that, in general, the detectors are functioning well. The quality assurance procedures indicate that the completeness and validity of the data is on a par with, and sometimes exceeds, the corresponding measures in other systems throughout this country.

The traffic volume data produced by the SunGuide archive should be useful to the district and statewide traffic counting programs. The capability to examine data from all detector stations and to create traffic count files in both the District and Central Office formats should facilitate the extraction of counts. Preliminary experience indicates that the accuracy of the count data varies among stations for reasons enumerated in the report. It appears, however that, with careful selection of stations, the FDOT traffic counting programs will benefit from the availability of the data.

The documentation included in the appendices to this report covers the installation, operation and maintenance aspects of the data management systems and the Web site in detail. STEWARD is now fully functional and ready to pass from the research and development phase to the operational phase.

A number of specific observations can be made from the insight developed during the course of the project:

- It has been demonstrated that the archive data characteristics are consistent with the principles of traffic flow theory. Relationships between the macroscopic descriptors of traffic flow demonstrated good agreement with those found in the literature, and with the empirical data presented in the Highway Capacity Manual.

- The speed, flow rate and occupancy values produced by the RTMS detectors are not measured independently but are derived based on proprietary algorithms. A comparison of the relationship between the density estimated from the flow rate and speed values and the density estimated from the vehicle occupancy values suggests that the measures are at least internally consistent.

- Several measures such as lane volume balance ratio, speed variation, kinetic energy and effective vehicle length were incorporated into the reports. These measures are not widely used for operational purposes but they were demonstrated through the use of examples to offer some potential for future applications.

- With the exception of known problem areas, usually resulting from construction, the traffic sensor subsystems in all districts appear to be functioning reliably and producing credible data.

- One specific detector problem area involves the failure of the detectors to report data during periods of extremely low volume, typically from, 1 a.m. to 5 a.m.. The cause of this problem is unknown.

- The quantity of data that must be transmitted daily from each TMC can be accommodated by the ETL procedures that have been developed for this purpose.

- The current facilities for data processing are adequate to accommodate the prototype system operation, but additional speed, storage and bandwidth will be required if the system utilization expands significantly beyond its present level.

- The quality assurance procedures described in the literature can be improved by incorporating additional QA tests that consider the relationships between the data from all of the lanes at a detector station as well as the consistency of the data between adjacent detector stations. The lane volume balance ratio at a given station is a good example of a characteristic that is not generally considered in current procedures. The consistency of the effective vehicle length between adjacent stations provides additional useful information.

- Detectors that produce unreasonable traffic volumes generally require adjustment and calibration. Threshold levels are required for determining when volumes are too high. Investigation of the distribution of maximum flow rates suggests that a threshold level of 2,900 veh/ln/h, which is approximately 20 percent greater than the typical capacity suggested by the HCM is appropriate for screening detectors. Flow rates in excess of this threshold occur in less than 1 percent of the observations.

- The traffic counts produced by the detectors can be extracted in a practical manner to augment the FDOT traffic counting programs. Some care needs to be exercised in choosing the appropriate detector stations and days for extraction. The desktop processing utilities have proven to be very helpful for this purpose.

- Effective vehicle length and lane volume balance ratios offer useful information at medium to high volumes. These measures can be misleading under very light traffic. They should only be applied during the 7 a.m. to 7 p.m. timeframe.

- Delay resulting from incidents may be estimated by a variety of methods. Those that focus directly on density at the point of the incident are likely to produce more credible results.

- Speed and kinetic energy fluctuations were shown to be associated with the sample incident that was studied; however no quantitative conclusions could be formulated. A substantially larger study of many incidents would be required to support definitive observations.

10.2 Recommendations

Recommendations from this study fall into two categories, including recommendations about future system operation and future research. The following recommendations are officered on the system operation:

1. STEWARD should continue to operate as long as resources permit. The current usage, as evidenced by Web site activity, justifies continued operation. It is reasonable to expect that the usage will increase. The University of Florida should continue to apply resources from the Center for Multimodal Solutions for Congestion Mitigation to the extent that they are available.

2. Funding should be sought to establish a permanent home for STEWARD, probably as a part of the FDOT Intelligent Transportation Systems establishment.

3. The STEWARD Web site should be maintained and a library page should be created for reports from projects that utilized STEWARD data.

4. Data from other SunGuide archive systems should be brought on board as they become operational. The step-by-step facility configuration process was documented in detail with that in mind.

5. The links between STEWARD and the SunSim Project being carried out by Florida International University should be strengthened.

6. The workshop material developed under this project should continue to be used to promote the use of STEWARD. The material was developed in accordance with the project scope for live instructor delivery. This material could be expanded for interactive web delivery. In that format, it would reach a much wider audience of intended users.

7. Districts should consider adding detectors to lanes upstream of exit ramps where they are currently omitted. Such detectors would greatly improve the accuracy of traffic counts extracted from the data. The additional detectors should be configured to provide archive data but not to be used in travel time estimation.

8. Districts should also consider placing detectors on entrance and exit ramps to form a closed system for input/output balance evaluation. The ability to perform input/output analysis very important to identifying inaccuracies in traffic volumes from detectors.

9. The districts could best improve the counting accuracy of their detectors by focusing their efforts initially on stations selected in consultation with the Statistics Office. This strategy would serve the dual purpose of providing the statistics office with useful data and creating a set of benchmark stations that could be used as a reference to evaluate other nearby stations.

STEWARD will continue to be a resource for research projects that need freeway operational data. The University of Florida and other universities within Florida should continue to develop research proposals to further the body of knowledge in freeway operations and congestion management. The potential use of archived data for this purpose was covered in this report. The following additional possibilities are suggested:

1. Development of the speed-flow density relationships: These relationships are the basis of the computational methodology of Chapter 10 of the 2010 Highway Capacity Manual. More detailed knowledge of the relationships could support an improved methodology.

2. Investigation of the basic continuity relationship described in Section 3 of this report, with the idea of developing shock wave analysis procedures from the archived data.

3. Continued research into managed lane operations using the HOTTER desktop utility program described in this report.

4. Continued research into improving the validity of traffic count data for FDOT traffic counting programs

5. Continued research into travel time variability reporting with a view to using actual data from STEWARD instead of a surrogate modeling process.

6. Investigation of the validity of data from various types of detectors (radar, loop video, etc.). More insight is needed into the cost/performance tradeoffs for these devices.

7. Investigation of turbulence as a predictor of incidents. This would require a significant study effort involving a substantial amount of archive data at aggregation levels below the five-minute level available from the STEWARD Web site. It would also require a substantial amount of incident data. The objectives of a project of this nature would be to 1) Develop means of quantifying turbulence using traffic flow principles, 2) Develop a statistical model that describes the relationships between the turbulence measures and the incidents and 3) Verify the model with a new set of traffic data.

The research team appreciates the support and cooperation of the funding agencies, SunGuide operators and STEWARD users in the development and operation of this system.

References

[1] Bertini, R. L. and J. Makle, “Beyond Archiving: Developing and Attracting Users of an Archived Data User Service,” TRB 87th Annual Meeting, Washington, D.C., 2008.

[2] 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 Final Report,” University of Florida Transportation Research Center Report FR-51449, 2008.

[3] Hranac, R., “Transforming Archived ITS data into information,” TRB 88th Annual Meeting, Washington, D.C, 2009.

[4]. PATH, "PeMS System Overview," Retrieved from http://pems.eecs.berkeley.edu/, 2005.

[5] Chen, C., “Freeway Performance Measurement System (PeMS),” California PATH research report, UCB-ITS-PRR-2003-22, 2003.

[6] Bertini, R., "Portland Transportation Archive Listing (PORTAL)," Retrieved from http://portal.its.pdx.edu/ , 2009.

[7] Tufte, K. A., S. Ahn, R. L. Bertini, B. Auffray and J. Rucker, “Toward the systematic Improvement of Data Quality in the Portland, Oregon Regional Transportation Archive Listing (PORTAL),” TRB 86th Annual Meeting, Washington, D.C., 2007.

[8] Bertini, R. L., S. Hansen, S. Matthews, A. Rodriguez and A. Delcambre, “PORTAL: Implementing a New Generation Archived Data User Service in Portland, Oregon,” *12th World Congress on ITS*, San Francisco 2005.

[9] Computer Science Corporation, “CHART II System Architecture” Maryland DOT, 2005.

[10] Maryland DOT, "CHART," Retrieved from http://www.chart.state.md.us/, 2009.

[11] Pack, M. L., "Regional Integrated Transportation Information System," Retrieved from http://www.cattlab.umd.edu/index.php?page=research&a=00023, 2009

[12] Wongsuphasawat, K., M., L. Pack, D. Filippova, M. VanDaniker, and A. Olea, “Visual Analytics for Transportation Incident Datasets,” TRB 88th Annual Meeting Washington, D.C., Transportation Research Board, 2009.

[13] Dellenback, S. and T. Duncan, “SunGuide: Software Users Manual,” Southwest Research Institute, 2008.

[14] Duncan, T. and J. Halbert, “SunGuide Database Design Document 3.1.0,” Southwest Research Institute, 2008.

[15] Turner, S., R. Margiotta, and T. Lomax, “Monitoring Urban Freeways in 2003: Current Conditions and Trends from Archived Operations Data,” *Publication FHWA-HOP-05-018*, Federal Highway Administration, 2004.

[16] May, A. D. *Traffic Flow Theory*, Prentice-Hall, 1990.

[17] Transportation Research Board. *Highway Capacity Manual*, *2010 Edition*, (Publication in progress)

TECHNICAL APPENDICES BOUND IN A SEPARATE VOLUME

Appendix 1: STEWARD System Description

Appendix 2: STEWARD Operation

Appendix 3: STEWARD Web Interface

Appendix 4: STEWARD Software Utility Documentation: