Statewide Transportation Engineering Warehouse for Archived Regional Data (STEWARD)

Florida Department of Transportation
University of Florida Transportation Research Center

Phase III Final Report

December 15, 2009

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

Disclaimer

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

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	5. Report Date
Development of a Central Data Warehouse for Statewide ITS and Transportation Data in Florida Phase III: Final Report	December 15, 2009
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Kenneth G Courage and Seokjoo Lee	FR-72734
9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
University of Florida Transportation Research Center PO Box 116580 Gainesville, FL 32611-6580	11. Contract or Grant No. BD545, RPWO # 93
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
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	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	18. Distribution Statement
ITS data archive, Central data warehouse	No Restrictions

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	381	

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

Most of the traffic management centers (TMCs) in Florida are operated by SunGuide, a system of hardware and software that was developed specifically for the Florida Department of Transportation (FDOT). 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:

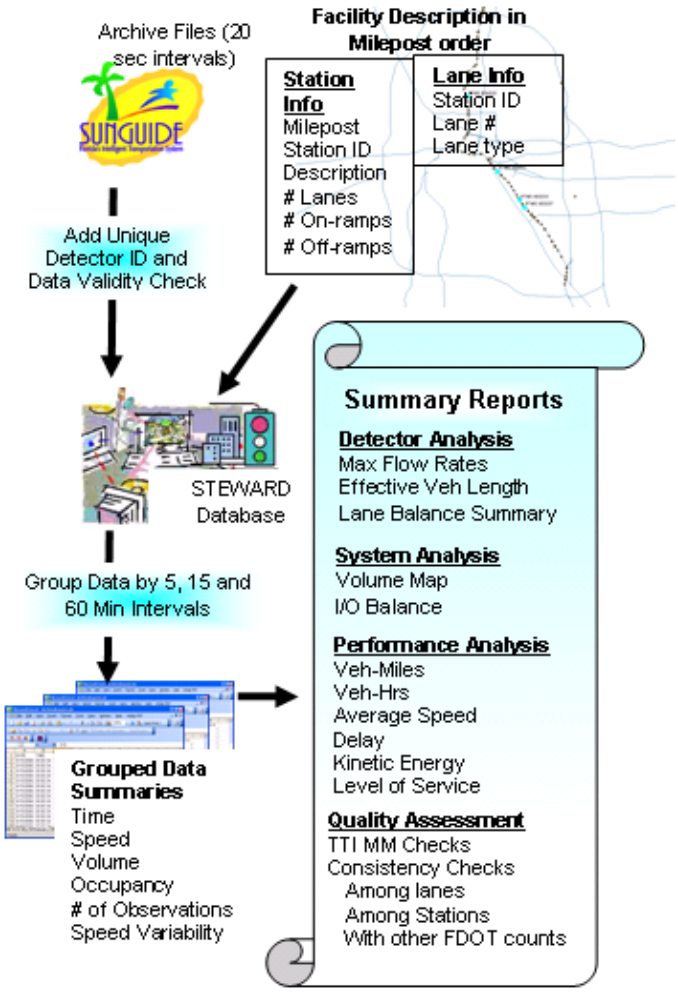
1. Present a series of workshops for potential providers and users of archived data.
2. Expand the scope of the STEWARD database to include additional SunGuide detectors.
3. Operate the system to provide data and reports to agencies and researchers.
4. Analyze the data to explore congestion modeling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

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 ½ 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:

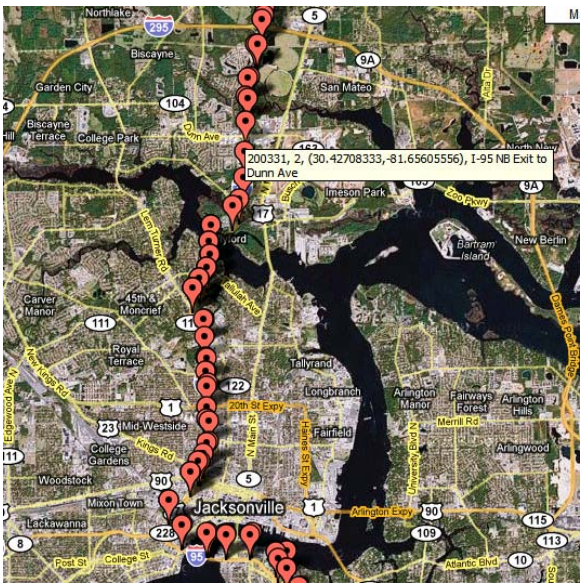
1. The software, including the database and data management components
2. The website manager, which provides database query, mapping and report generation components
3. Several desktop utility programs that process data from and create special analysis reports
4. The final report, which summarizes the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

The STEWARD website provides the interface between the archived data and the users of those data. The following items are available on the website 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 website 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

1 INTRODUCTION AND SUMMARY 1
1.1 Statement of the Problem 1
1.2 Project Objectives 1
1.3 Summary of Project Tasks 1
1.3.1 Literature Review 1
1.3.2 Analysis of the Traffic Archive Data 2
1.3.3 Data Management System Development 2
1.3.4 Statewide Deployment 2
1.3.5 Quality Assurance 2
1.3.6 Workshop Presentations 2
1.4 Organization of the Report 2
1.4.1 Content of the Report Body 2
1.4.2 Supplemental Material 3
1.5 Key Definitions for Geographical Elements 4
2 LITERATURE REVIEW 6
2.1 Traffic Data Warehouses in Other States 6
2.1.1 California – PeMS System 6
2.1.2 Oregon – PORTAL System 8
2.1.3 Maryland - CHART System 10
2.1.4 Summary of Traffic Data Archive Systems 11
2.2 The SunGuide Traffic Management System in Florida 11
2.2.1 Operational Features 11
2.2.2 Archive Functions 13
2.3 Quality Control of Traffic Data 14
3 APPLICABLE FREEWAY TRAFFIC FLOW PRINCIPLES 16
3.1 Speed Flow Rate and Density Relationships 16
3.1.1 The Fundamental Diagram 16
3.1.2 Other Speed-Flow-Density Models 18
3.1.3 Highway Capacity Manual Treatment of Speed, Flow and Density 19
3.2 Level of Service (LOS) 21
3.3 Platoon Propagation 22
3.4 Maximum Flow Rates 22
3.5 Effective Vehicle Lengths 22
3.6 Lane Volume Balance Ratio 22
3.7 Input/Output (I/O) Volume Balance 22
4 CENTRAL DATA WAREHOUSE REQUIREMENTS 23
4.1 Functional Requirements 23
4.1.1 Raw SunGuide Archive Data 23
4.1.2 TMC Configuration Requirements 24
4.1.3 ETL Requirements 27
4.2 Data Transfer Automation Requirements 27
4.3 Reporting Requirements 28
5 CENTRAL DATA WAREHOUSE DEVELOPMENT 30
5.1 STEWARD System Description 30

5.1.1 System Overview 30
5.1.2 Database Design and Architecture 31
5.1.3 Data Flow 31
5.2 STEWARD Operation 31
5.2.1 Oracle Database Program Installation 32
5.2.2 STEWARD Deployment 32
5.2.3 STEWARD Web Installation 33
5.2.4 STEWARD Management 33
5.3 Internet Access Features 33
5.3.1 Overview of the STEWARD Web Interface 33
5.3.2 STEWARD Web Architecture 34
5.3.3 Maps 35
5.3.4 Report Levels 35
5.3.5 Report Types 37
5.4 ETL Operations 37
5.5 Current Status of the System 38
6 DEVELOPMENT OF QUALITY ASSURANCE PROCEDURES 41
6.1 Level 1 Completeness Test 41
6.1.1 Availability of the District Data 42
6.1.2 Missing Detector Data 42
6.1.3 All Zero or Stuck Detectors 43
6.1.4 Duplicate or Negative-Scan Data 44
6.1.5 Level 1 Test Summary 44
6.2 Level 2 Data Validity Test 44
6.2.1 Maximum Volume Test 45
6.2.2 Maximum Occupancy Test 45
6.2.3 Minimum and Maximum Speed Tests 45
6.2.4 Multivariate Consistency Test 46
6.2.5 Truncated Occupancy Values Test 46
6.2.6 Maximum Estimated Density 46
6.2.7 Summary of Level 2 Tests 47
6.3 Level 3 Station Data Validation 47
6.3.1 Maximum Flow Rates 48
6.3.2 Speed Flow Relationships 50
6.3.3 Effective Vehicle Lengths 52
6.3.4 Lane Volume Balance Ratio 56
6.3.5 Hourly and Daily Volume Variation 61
6.4 Level 4: System Level Tests 63
6.4.1 Continuity of EVL 63
6.4.2 Continuity of Volume 63
7 OPERATIONAL APPLICATIONS FOR ARCHIVED DATA 65
7.1 Summary of Available Reports 65
7.1.1 Diagnostic Procedures 65
7.1.2 Station Level Reports 66
7.1.3 Section Level Reports 68
7.1.4 Facility Level Reports 76

7.2 Traffic Volume Data for Traffic Counting Programs 79
7.2.1 Converting ITS Data to FDOT Counts 79
7.2.2 Reporting of Traffic Volume Trends 80
7.3 Integration with Statewide Crash Data Records 82
7.4 Integration with the Roadway Characteristics Inventory 83
7.5 General Support for Periodic Reporting Requirements 83
7.6 Diagnostic Support for TMC Detector Operation and Maintenance 83
7.7 Other Applications for the STEWARD Reports 83
7.7.1 Work-Zone Crash Analysis 83
7.7.2 Support for Identification of Recurring Congestion 84
7.7.3 Travel Time Reliability Reporting 84
8 RESEARCH APPLICATIONS FOR ARCHIVED DATA 85
8.1 Analysis of Breakdown at a Freeway Ramp 85
8.2 Simulation Support for SunGuide 85
8.3 Other TRC Research Applications 86
8.3.1 Modeling the Location of Crashes within Work Zones 86
8.3.2 A Case Study in Spatial Misclassification of Work Zone Crashes 87
8.3.3 Analyzing the Effectiveness of Enhanced Penalty Zones 87
8.3.4 Capacity of Florida Freeways, FDOT Project BDK-75-977-08 87
8.3.5 Travel Time Reliability 87
8.3.6 Freeway Work Zone Capacity 87
8.3.7 NCHRP 3-87, Proactive Ramp Management 87
8.3.8 Doctoral Dissertation Project 88
8.4 Assessment of Turbulence as a Predictor of Incidents 88
8.5 Other STEWARD Users 88
9 DATA ANALYSIS EXAMPLES 89
9.1 Speed-Flow-Density Relationships 89
9.1.1 Speed 89
9.1.2 Density 90
9.1.3 Flow Rate 91
9.1.4 Examples of Relationships 91
9.2 Crash and Incident Analysis Applications 94
9.2.1 Overall Crash Characteristics 94
9.2.2 Sample Crash Description 95
9.2.3 Hourly Flow Rates 96
9.2.4 Occupancy 99
9.2.5 Estimation of Delay from Speed 100
9.2.6 Estimation of Delay from the Travel Time Reliability Report 101
9.2.7 Comparison of Delay Estimation Methods 102
9.2.8 Lane Volume Balance Ratio 102
9.2.9 Speed Variance 104
9.2.10 Kinetic Energy 105
9.3 Analysis of Managed Lanes 105
9.3.1 Managed Lane Performance Measures 106
9.3.2 Example Results 108
9.4 Travel Time Reliability 109

10 CONCLUSIONS AND RECOMMENDATIONS 111
10.1 Conclusions 111
10.2 Recommendations 113
TECHNICAL APPENDICES .....
Appendix 1: STEWARD System Description 118
Appendix 2: STEWARD Operation .....
Appendix 3: STEWARD Web Interface .....
Appendix 4: STEWARD Software Utility Documentation: .....
WORKSHOP PRESENTATION MATERIAL .....

List of Figures

Figure 1: Demonstration of the Fervor application 11
Figure 2: SunGuide system architecture 13
Figure 3: Fundamental diagram relating speed, flow rate and density 17
Figure 4: Determination of parameters from the fundamental diagram 17
Figure 5: Current concept of the speed-flow-density relationship 18
Figure 6: Typical speed-flow relationship presented in the HCM 19
Figure 7: Speed-flow relationships for basic freeway segments in the HCM 20
Figure 8: HCM flow-density relationship for basic freeway segments 21
Figure 9: Example of the raw data from the SunGuide TSS archive 23
Figure 10: Summary of required functions and data flow 24
Figure 11: Automated data flow diagram for the SunGuide archive data 28
Figure 12: STEWARD system configuration 31
Figure 13: Example screen capture from the installation process 32
Figure 14: STEWARD web architecture 34
Figure 15: STEWARD overview page 35
Figure 16: Example of an interactive satellite photo map 36
Figure 17: GIS map example 36
Figure 18: ETL Utility data flow 39
Figure 19: Sample daily report from the ETL process 39
Figure 20: Flow rate histogram from four selected stations 48
Figure 21: Cumulative percentage over 2400 veh/ln/h from I-95 stations 50
Figure 22: Speed vs Flow rates at station 210371, Lane 1 51
Figure 23: Speed *vs*. Flow rates at station 210371, Lane 2 51
Figure 24 Speed vs Flow rates at station 210371, Lane 3 52
Figure 25: Effect of flow rate on EVL at a selected station 53
Figure 26: EVL histogram from four selected stations 54
Figure 27: Lane-by-lane effective vehicle length distribution 55
Figure 28: Flow-occupancy relationship for low effective vehicle lengths 56
Figure 29: Flow-occupancy relationship for normal effective vehicle lengths 56
Figure 30: Effect of flow rate on lane volume balance ratio at a selected station 57
Figure 31: Lane volume balance ratio histogram from four selected stations 58
Figure 32: Location of Station 210032 60
Figure 33: Location of Station 200201 61
Figure 34: Average flow rates by time of day for a selected station 62
Figure 35: Total 24 hr weekday volume for one year at a selected station 62
Figure 36: EVLs at stations on a section of I-95 63
Figure 37: Hourly traffic volumes at stations on a section of I-95 64
Figure 38: Example SunGuide TSS and Statistics Office count comparison 79
Figure 39: Overview of the ITSCounts data flow 80
Figure 40: Sample page from a facility count analysis report 81
Figure 41: Example of flow rate *vs.* speed 91
Figure 42: Example of flow rate *vs.* density 92
Figure 43: Example of Speed *vs.* density 93

Figure 44: Comparison of density computation methods 94
Figure 45: Number of crashes by milepost 95
Figure 46: Incident location on a satellite photo 96
Figure 47: Hourly flow rates during the incident 96
Figure 48: Hourly flow rates during incident-free operation 97
Figure 49: Comparison of 5 minute counts for the incident and non-incident case 98
Figure 50: Cumulative differences between incident and non-incident volume counts 98
Figure 51: Effect of the incident on occupancy 99
Figure 52: Occupancy contour graph during the incident 100
Figure 53: Effect of the incident on speed 100
Figure 54: Speed contour graph during the incident 101
Figure 55: Lane volume balance ratio during the incident 103
Figure 56: Occupancy contours by lane during the incident 103
Figure 57: Speed variation by time of day 104
Figure 58: Flow rate and kinetic energy at the incident by time of day 105

List of Tables

Table 1: Evaluation Criteria for the PeMS Loop Detector Data 8
Table 2: PeMS Traffic Data Quality Statistics 8
Table 3: Evaluation Criteria for the PORTAL Detector Data 9
Table 4: Comparison of Traffic Data Warehouses in other States 12
Table 5: Basic Rules of the QC Criteria 14
Table 6: Quality Control Criteria 15
Table 7: Level of Service Thresholds in the 2010 HCM 21
Table 8: Status of STEWARD Facilities and Stations 38
Table 9: TSS Data Availability in STEWARD 38
Table 10: STEWARD Web Statistics for June , July and August 2009 40
Table 11: Maximum Flow Rate Frequencies 49
Table 12: Histogram of the Effective Vehicle Lengths 55
Table 13: Histogram of Lane Volume Balance Ratio 58
Table 14: Frequency of LVBR Greater than 10.0 59
Table 15: Column Description for the *All data fields* Report 66
Table 16: Column Descriptions for the *Station Traffic Counts R*eport 67
Table 17: Column Description for the *Maximum Flow R*eport 68
Table 18: Column Description for the *Effective Vehicle Length R*eport 69
Table 19: Column Descriptions for the *Performance Measures R*eport 70
Table 20: Column Descriptions for the *Performance Measures* Report Baseline 71
Table 21: Sample Performance Measures Report 72
Table 22: Column Description for the *Travel Time Reliability* Report 73
Table 23: Column Descriptions for the *Travel Time Reliability* Report Baseline 74
Table 24: Travel Time Reliability Frequency Table Column Descriptions 74
Table 25: Sample Performance Measures Report 75
Table 26: Column Descriptions for the *All Data Fields* Report 76
Table 27: Column Descriptions for the *Volume Map and I/O Balance* Report 77
Table 28: Column Descriptions for the *Facility Traffic Counts* Report 78
Table 29: Results for the Travel Time Reliability Example 109

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

Most of the TMCs in Florida are operated by SunGuide, a system of hardware and software that was developed specifically for the Florida Department of Transportation (FDOT). 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 ½ 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, November 5, 2009 (15 participants)
- District 2: 513 Weil Hall (UF Campus), Gainesville, December 7, 2009 (Attendance: ___)
- District 4: SunGuide Center, Ft Lauderdale, December 2, 2009 (Attendance: ___)
- District 5: District Office, Deland December 8, 2009 (Attendance: ___)
- District 6 SunGuide Conference Room 2**,** Miami, December 9, 2009 (Attendance: ___)
- District 7: Tampa Bay SunGuide Center**,** Tampa, December 1, 2009 (Attendance: ___)

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 which 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 5AM to 10PM (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% 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	Intermittent	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, OR 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 > 100MPH
4	Speed < 5MPH
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.

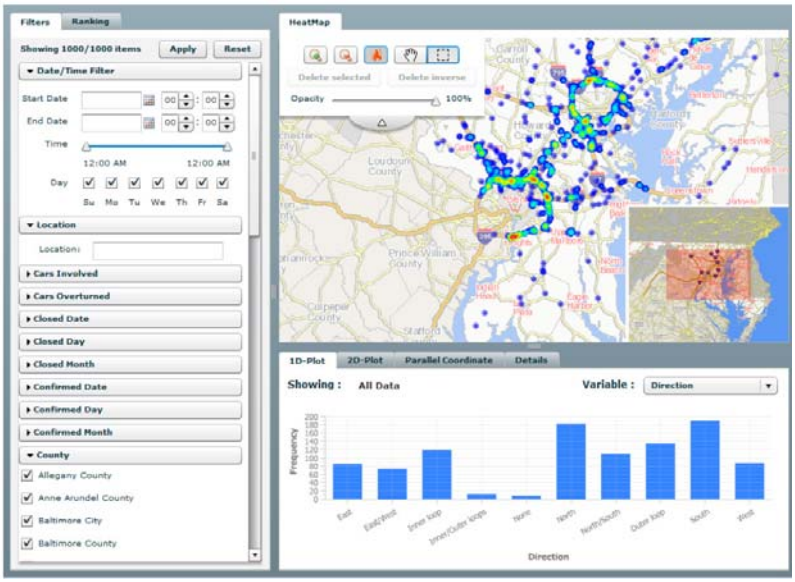


Figure 1: Demonstration of the Fervor application

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 eleven 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).

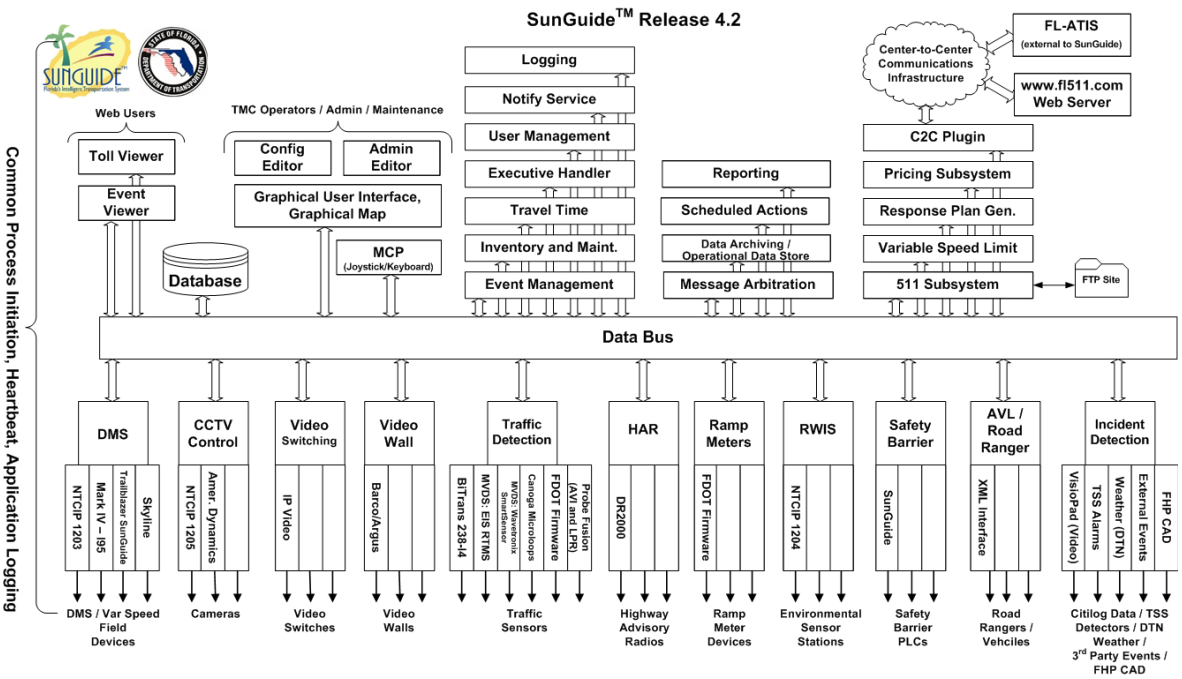


Figure 2: SunGuide system architecture

The SunGuide database archives all of the data from these eleven 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 > 3000 (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 8 consecutive identical volume-occupancy-speed values. That is, the volume AND occupancy AND speed values have more than 8 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

The original model of these relationships was developed by Greenshields 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.

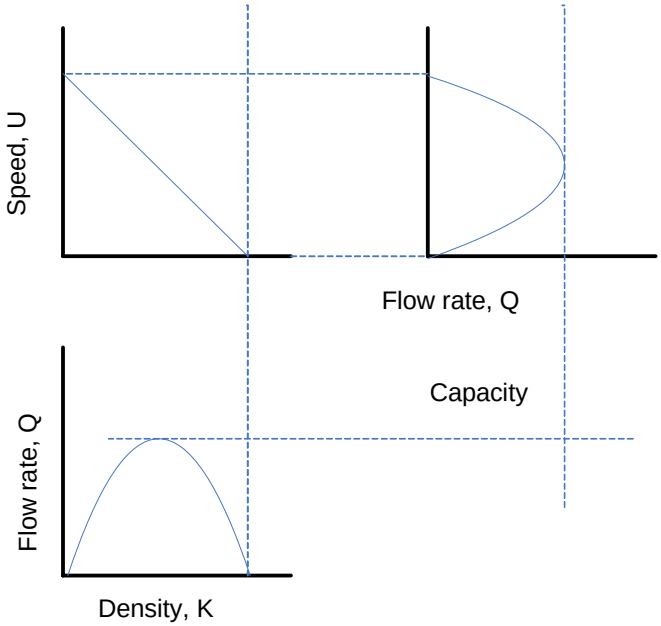


Figure 3: Fundamental diagram relating speed, flow rate and density

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

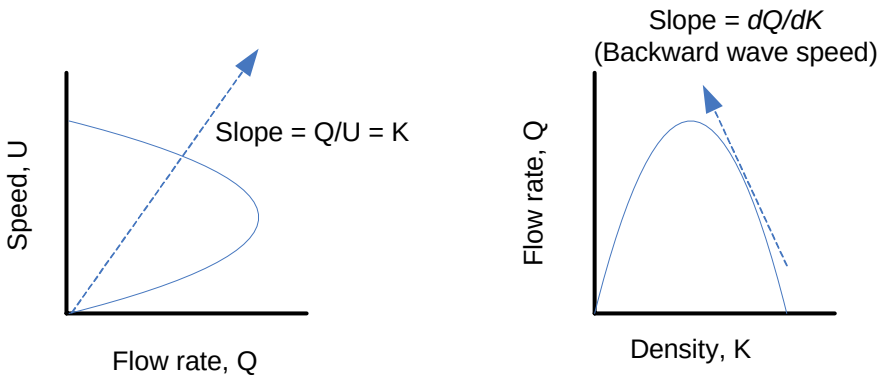


Figure 4: Determination of parameters from the fundamental diagram

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 $(Q_2-Q_1)/(K_2-K_1)$.

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.

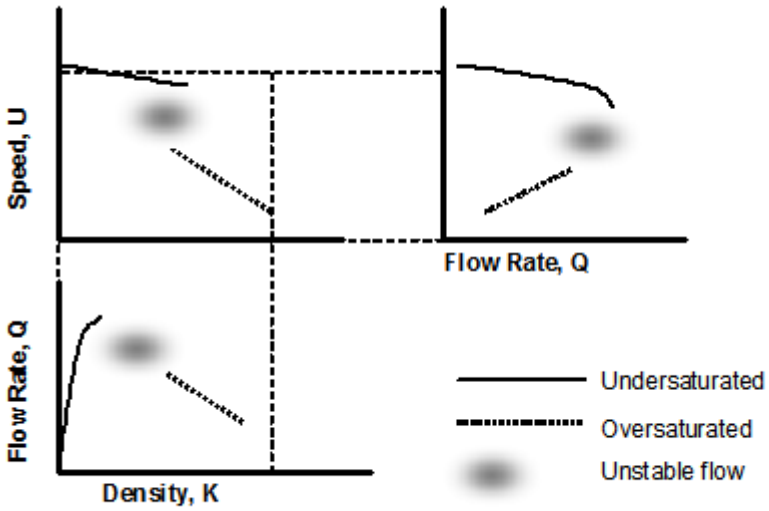


Figure 5: Current concept of the speed-flow-density relationship

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.

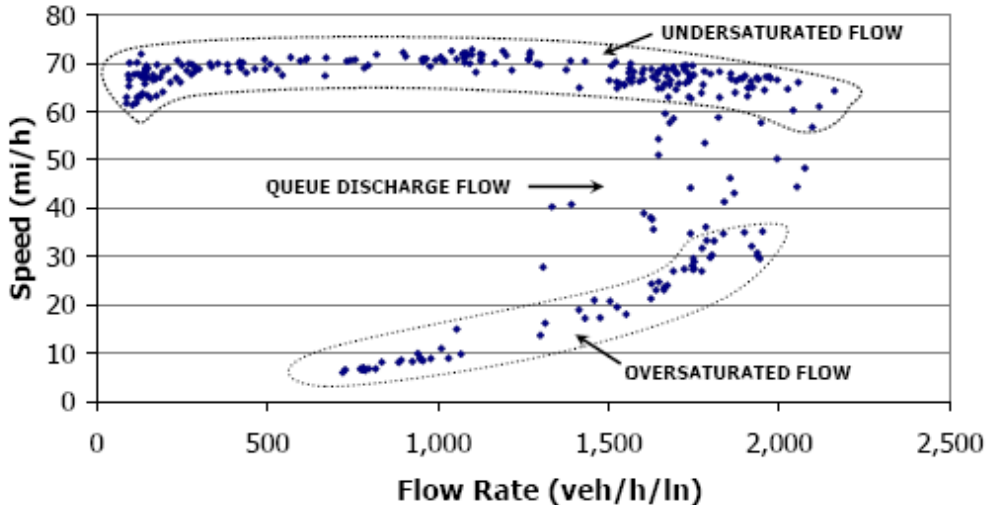


Figure 6: Typical speed-flow relationship presented in the HCM

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 ½ of the free flow speed. As indicated in Figure 7, the 2010 HCM places the capacity at speeds between 50 and 54 mph.

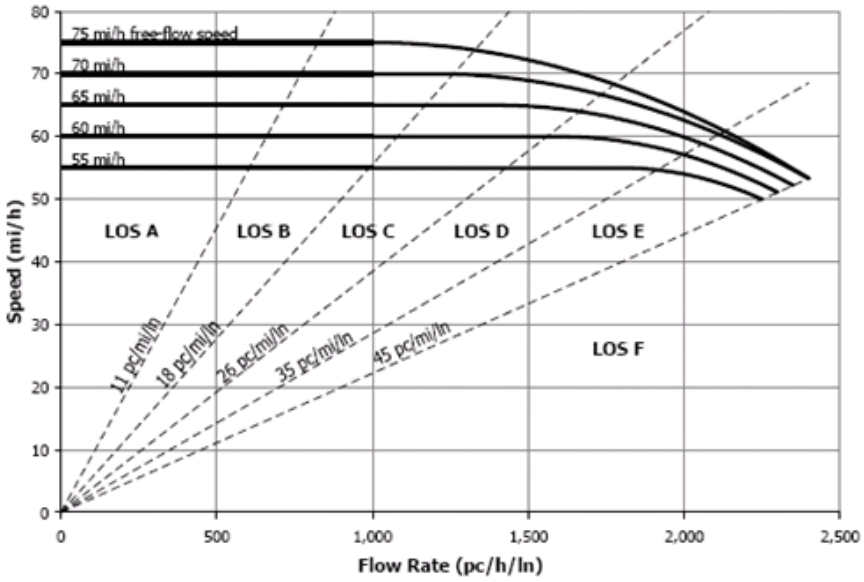


Figure 7: Speed-flow relationships for basic freeway segments in the HCM

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 ft, the space gap between vehicles (front bumper to rear bumper) would be slightly less than 12 ft at jam density.

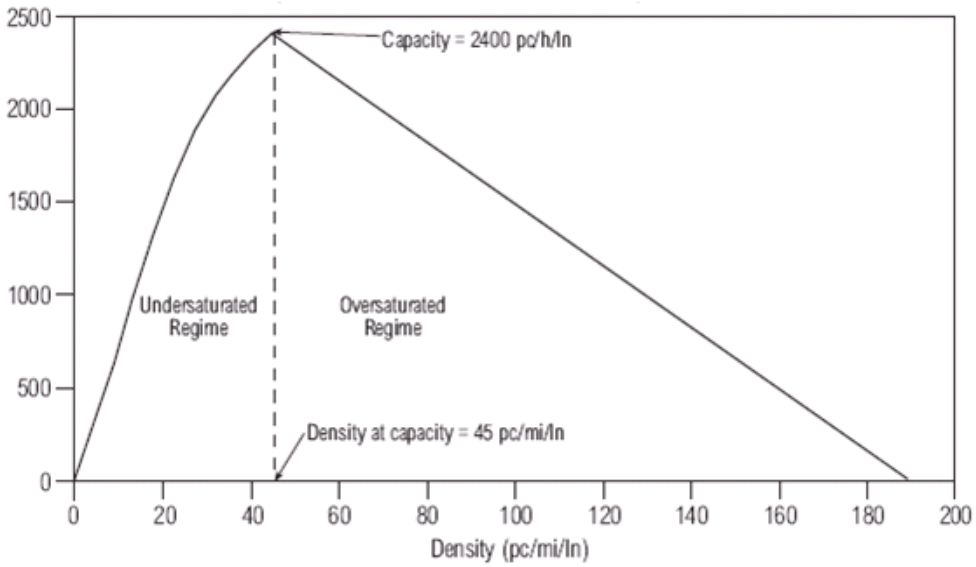


Figure 8: HCM flow-density relationship for basic freeway segments

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-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 2400 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.

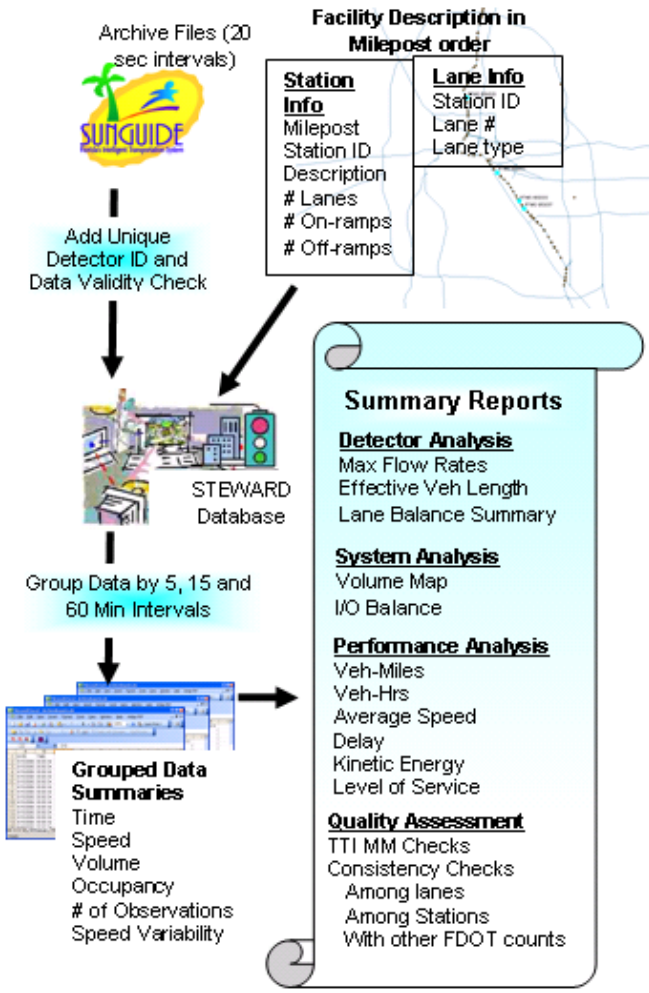


Figure 10: Summary of required functions and data flow

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:* his 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 id's 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. The facility assignment must be carried out by district personnel.
- 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.

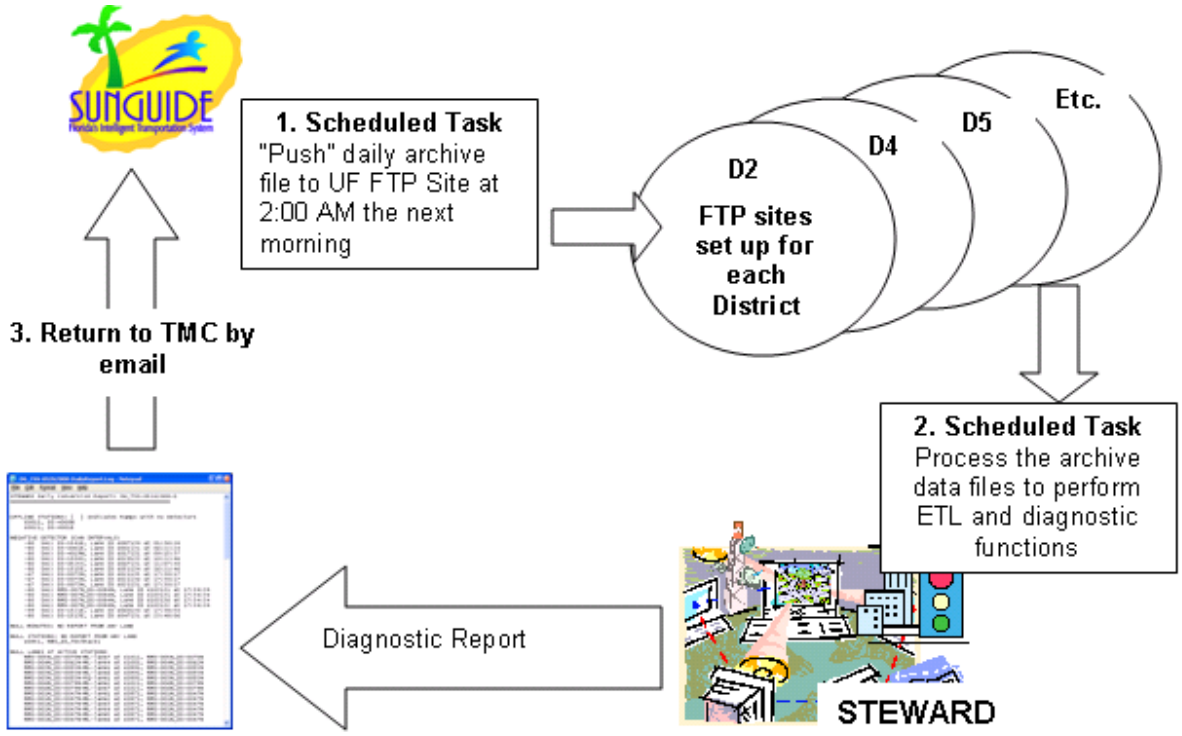


Figure 11: Automated data flow diagram for the SunGuide archive data

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 ten 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.

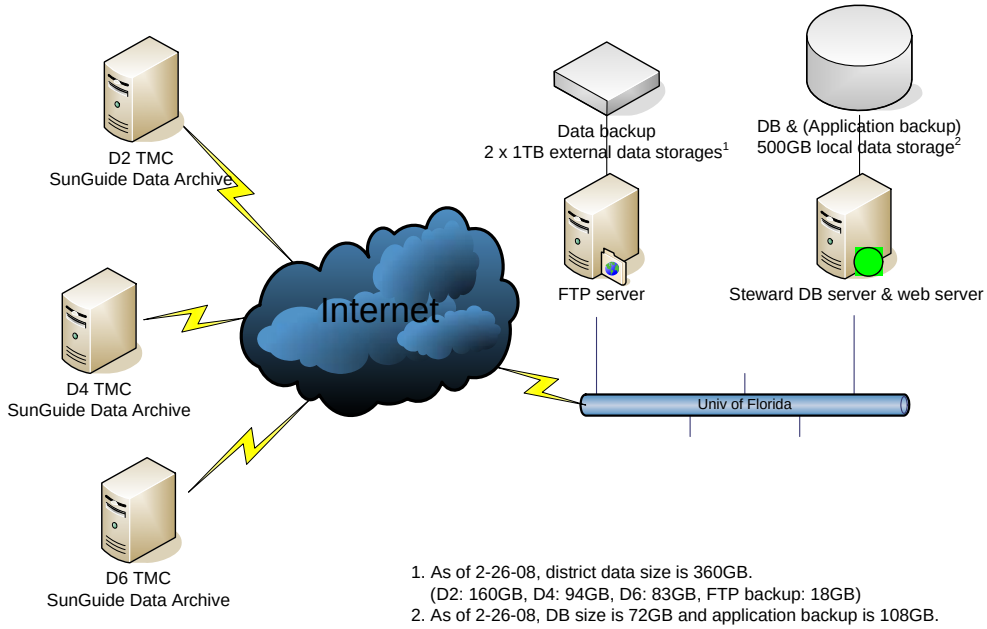


Figure 12: STEWARD system configuration

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 data base 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.

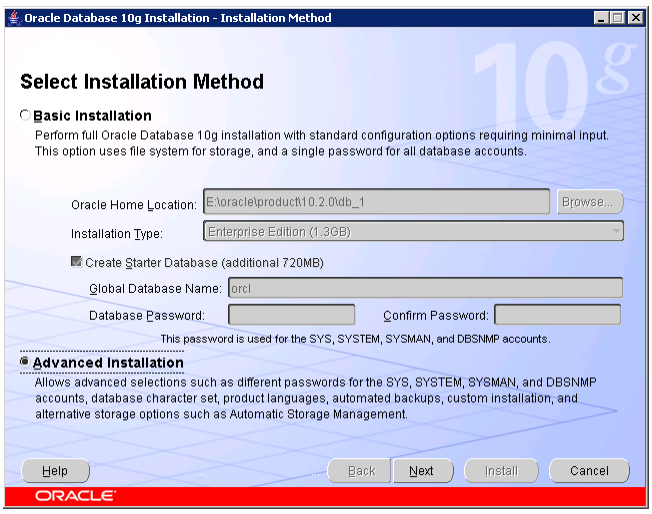


Figure 13: Example screen capture from the installation process

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.

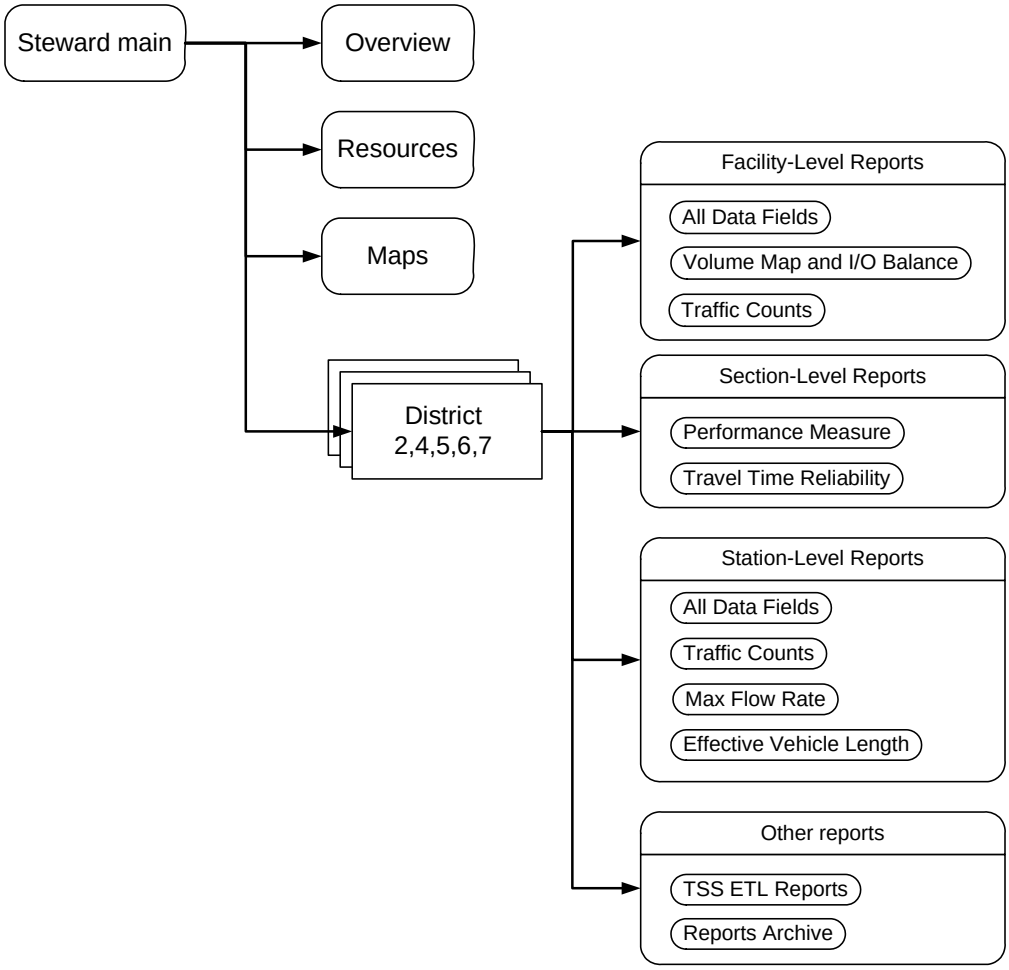


Figure 14: STEWARD web architecture

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.

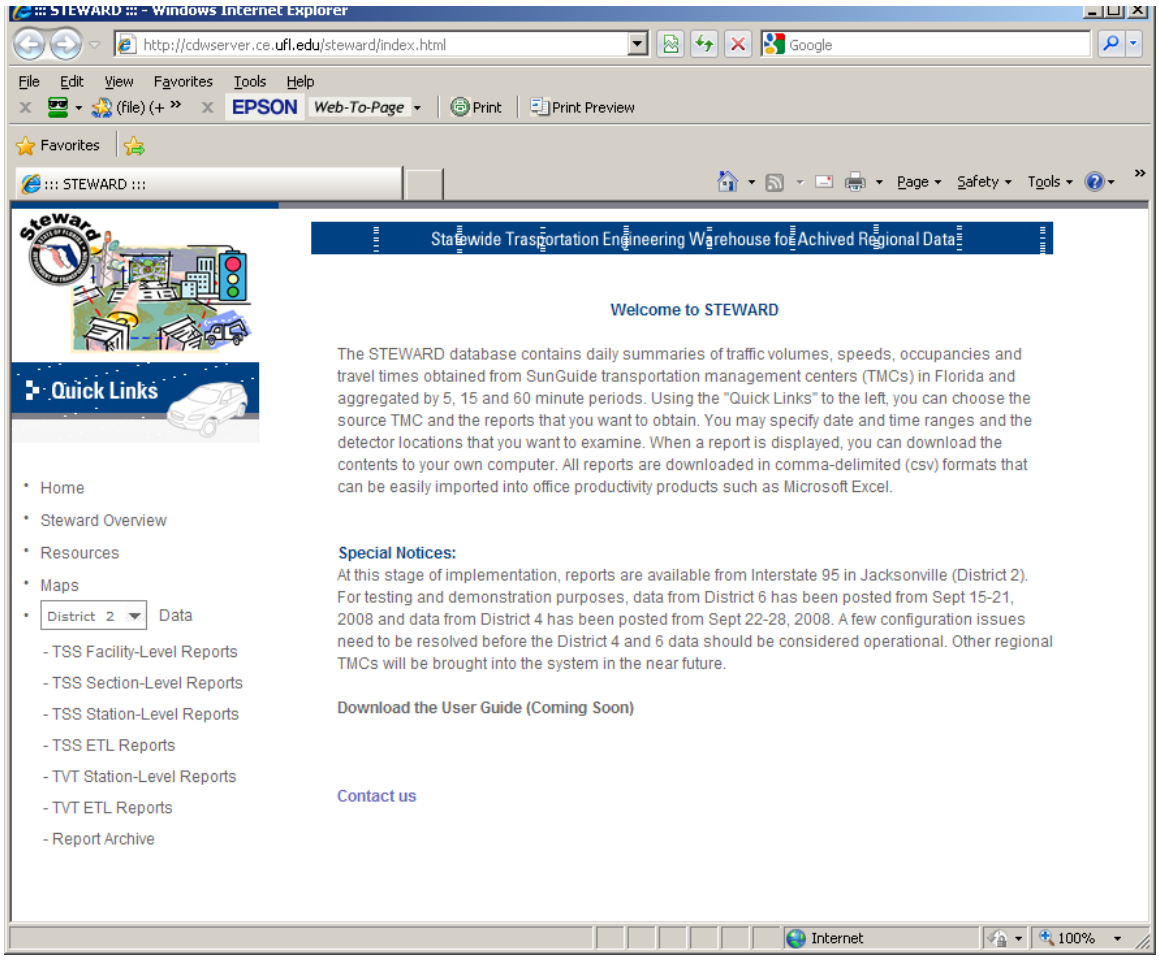


Figure 15: STEWARD overview page

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 (5 minutes, 15 minutes or 1 hour)

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

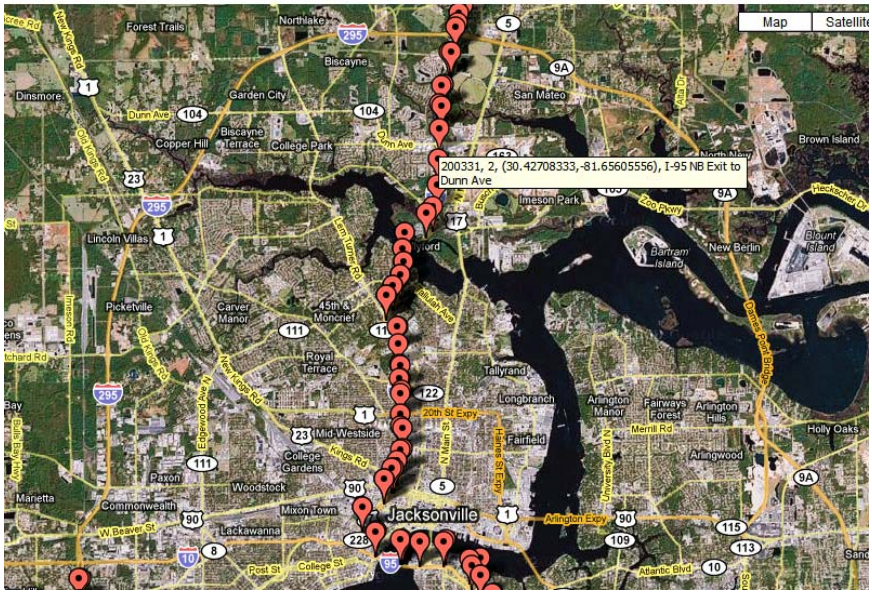


Figure 16: Example of an interactive satellite photo map

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 5, 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-8 lanes at one location. Approximately 4200 lanes are covered by STEWARD stations.

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 9-30-2009.

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, FL.

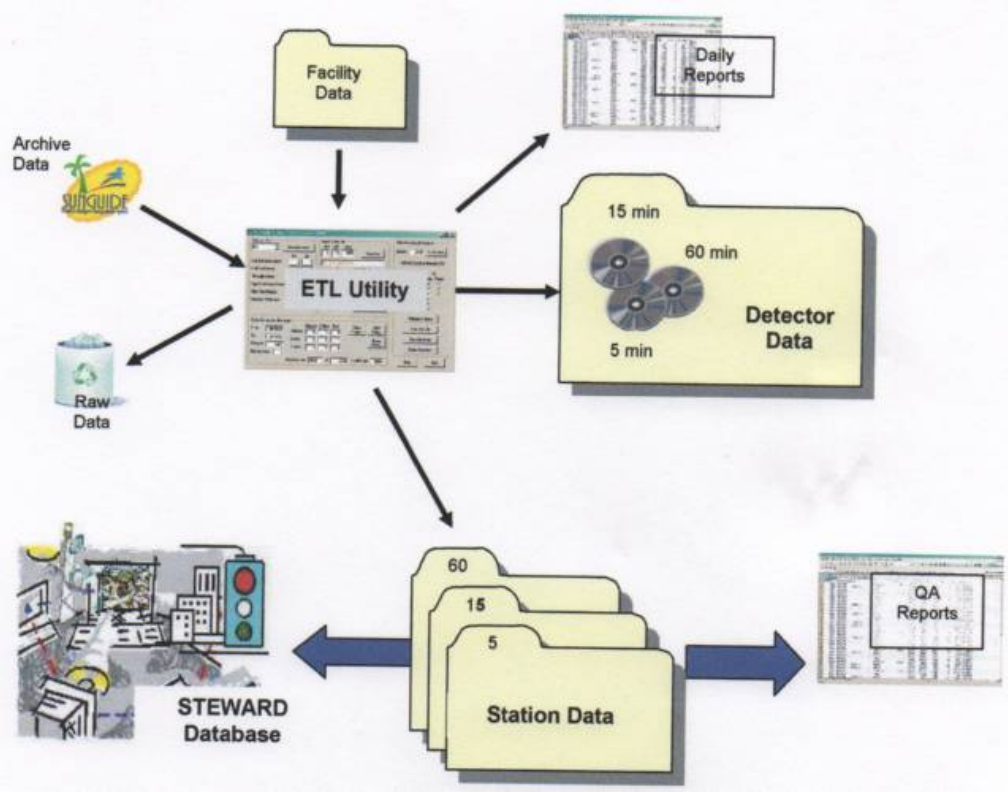


Figure 18: ETL Utility data flow

```
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 US 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 3: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 fourteen 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 12/2/08 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:

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

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

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

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%) 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 20sec, so 15 data observations are tested during each 5 min 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 5min (1)	Number of stuck data (V/S/O) for 5min (2)	Total expected number of traffic for 5min (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% of the daily traffic data records were reported with zero or negative scan intervals. It appears that these problems have been resolved by the SunGuide contractor. 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 20second data
- 25 vehicles for 30 sec data
- 250 vehicles for 5min data
- 3000 vehicles for 1hr 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% of the observations failed this test.

6.2.2 Maximum Occupancy Test

Maximum allowable occupancy for each lane data should be less than

- 95% for 20 or 30 second data
- 80% for 1 to 5min 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% 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 1 to 5min 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%. The only exception is the minimum speed test in District 6 which showed a failure rate of 3.32%. 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 20second 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 one percent and be truncated to a zero value. This test was applied to 20second 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 Oct. 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: Oct., 2008, weekdays (23 days)
- Time: Morning peak (7:00 AM~10:00 AM) rush-hour traffic

The following four stations were selected from twenty three 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 5min, 15min or 1hr traffic volumes. Excessive maximum flow rates could be the result of detector over counting. In the Level 2 tests the threshold for maximum flow rate from an individual lane was set at 3000 veh/ln/h. For purposes of this test, a station level threshold value of 2400 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 15min traffic volume data. From the cumulative percentages, 99.3% of flow rates are less than 2400 veh/ln/h, which was selected as the threshold criterion.

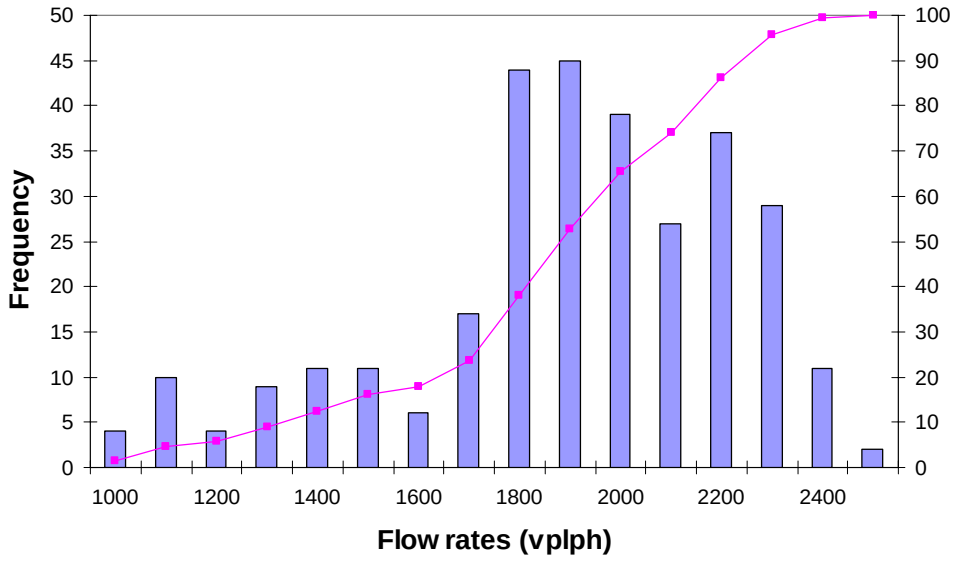


Figure 20: Flow rate histogram from four selected stations

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%
1000	94	12.04%
1100	118	14.39%
1200	128	16.93%
1300	129	19.50%
1400	229	24.05%
1500	247	28.97%
1600	267	34.28%
1700	346	41.17%
1800	378	48.69%
1900	414	56.92%
2000	521	67.29%
2100	515	77.54%
2200	378	85.06%
2300	287	90.77%
2400	225	95.24%
2500	98	97.19%
2600	41	98.01%
2700	26	98.53%
2800	20	98.93%
2900	21	99.34%
3000	16	99.66%
3100	8	99.82%
3200	3	99.88%
3300	4	99.96%
3400	2	100.00%

Table 11: Maximum Flow Rate Frequencies

The flow rate observations in excess of 2400 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 2400 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 **Error! Reference source not found.**, the observations that exceeded 2400 veh/ln/h were all below 3400 veh/ln/h. The cumulative distribution reached the 99% level at a flow rate of approximately 2900 veh/ln/h. Raising the threshold to 2900, which is approximately 20% higher than the HCM's value of 2400, will increase the usefulness of the maximum flow rate test without an unreasonable departure from the HCM results. Therefore a value of 2900 veh/ln/h will be applied.

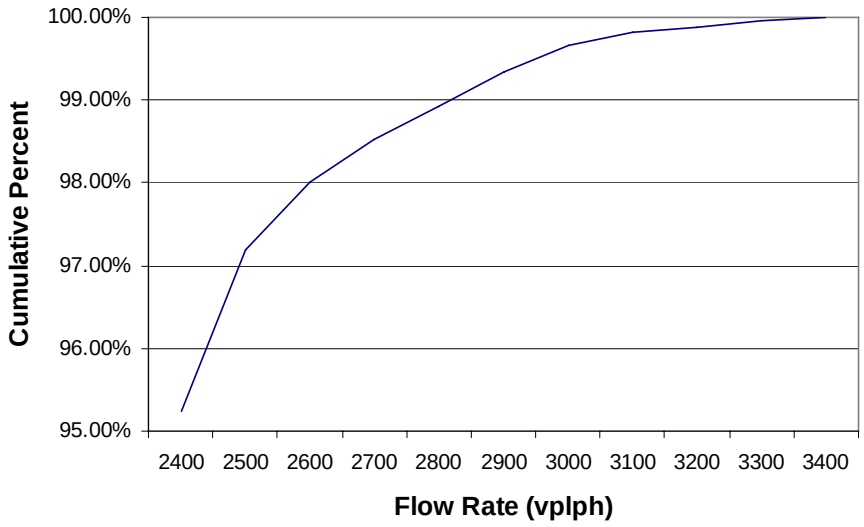


Figure 21: Cumulative percentage over 2400 veh/ln/h from I-95 stations

From Table 11, the maximum flow rates exceed 2400vplph 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 3500 vphpl. Lanes 2 and 3 show smaller peaks in the range of 2300 to 2600 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.

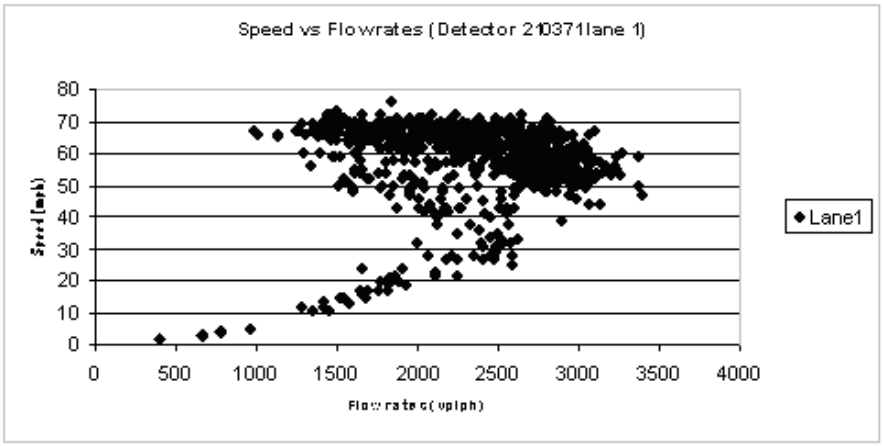


Figure 22: Speed vs Flow rates at station 210371, Lane 1

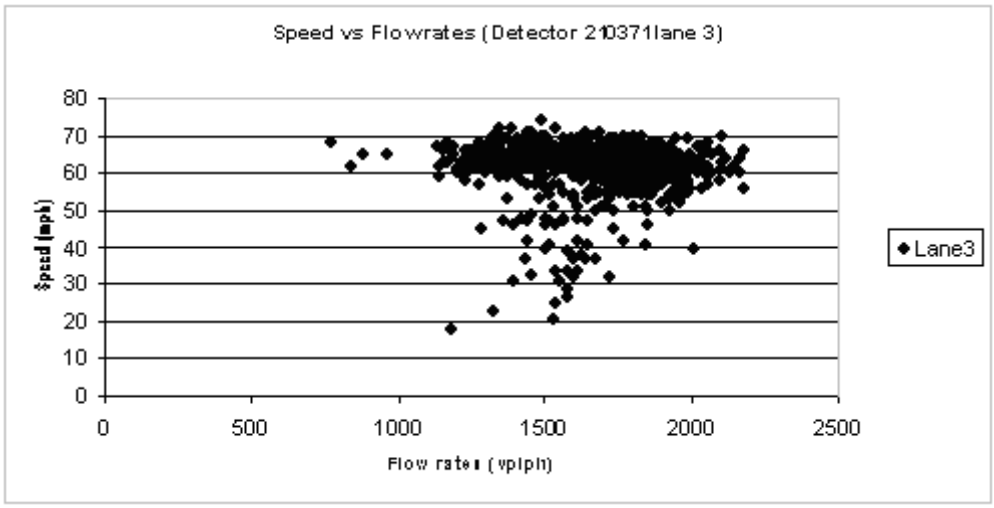


Figure 23: Speed *vs.* Flow rates at station 210371, Lane 2

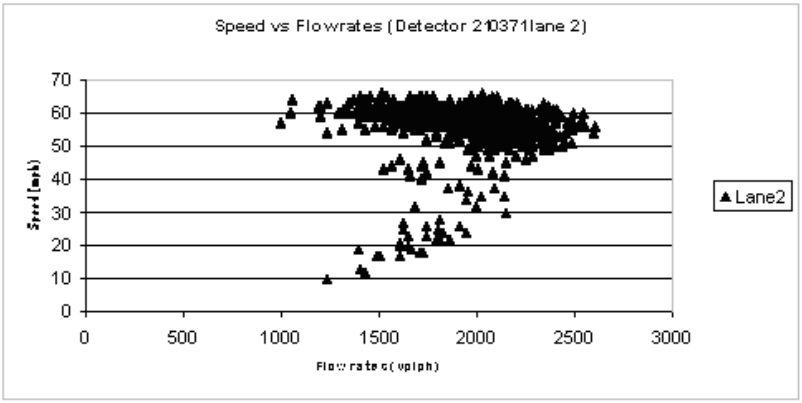


Figure 24 Speed vs Flow rates at station 210371, Lane 3

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:

$$Q=K*U$$

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

$$U \approx V$$

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

(Effective vehicle length) = V * 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%) 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 21ft, which is the sum of average passenger vehicle length (15ft) and average detector length (6ft). 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.

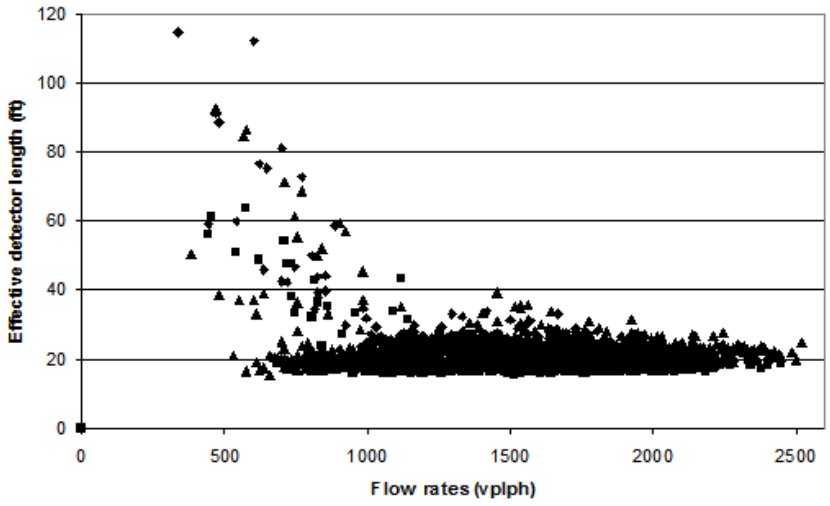


Figure 25: Effect of flow rate on EVL at a selected station

The flow rate histograms for the four selected stations were created as shown in Figure 26 from 5 minute data. From the cumulative percentages, 96.7% of the EVLs are less than 30ft, which was selected as the threshold criterion for purposes of this test.

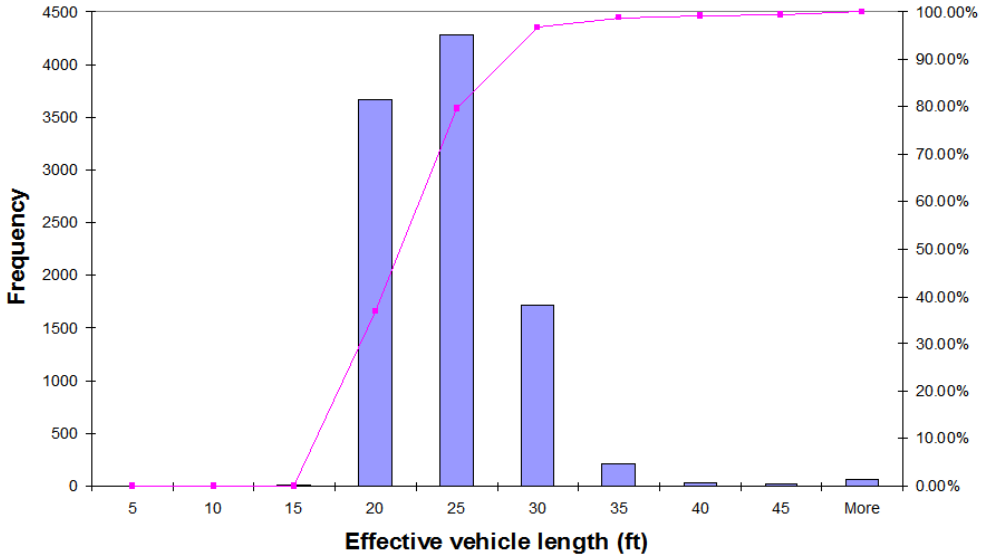


Figure 26: EVL histogram from four selected stations

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 ft and Lane 1 is 16.0 ft. 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% and 5% of the intervals showed lengths less than 15 ft in Lanes 1 and 2, respectively. Approximately 3% and 8% of the intervals showed lengths greater than 30 ft in Lanes 1 and 2, respectively.

Figure 28 and Figure 29 show the occupancy-flow relationships for low effective vehicle length cases (< 10 ft) and normal cases (>10 ft). 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

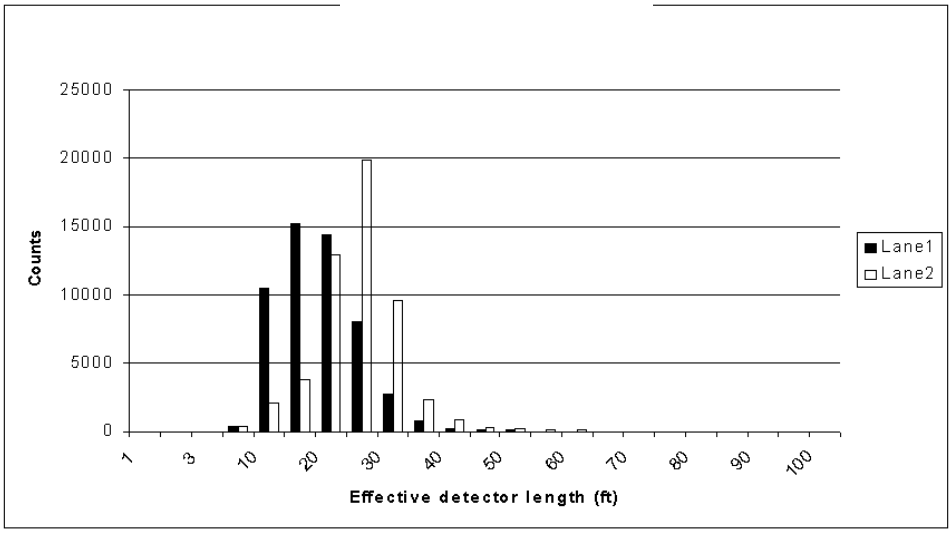


Figure 27: Lane-by-lane effective vehicle length distribution

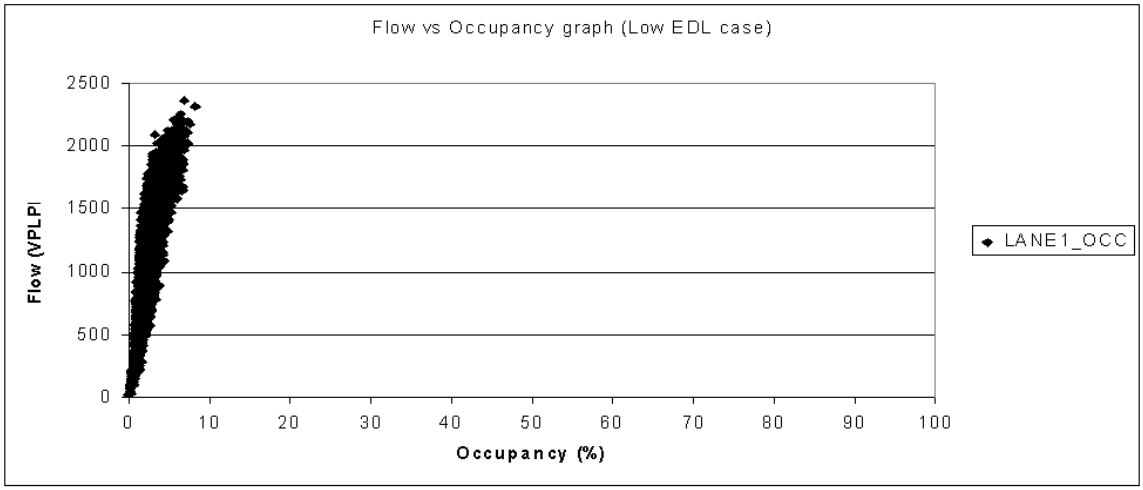


Figure 28: Flow-occupancy relationship for low effective vehicle lengths

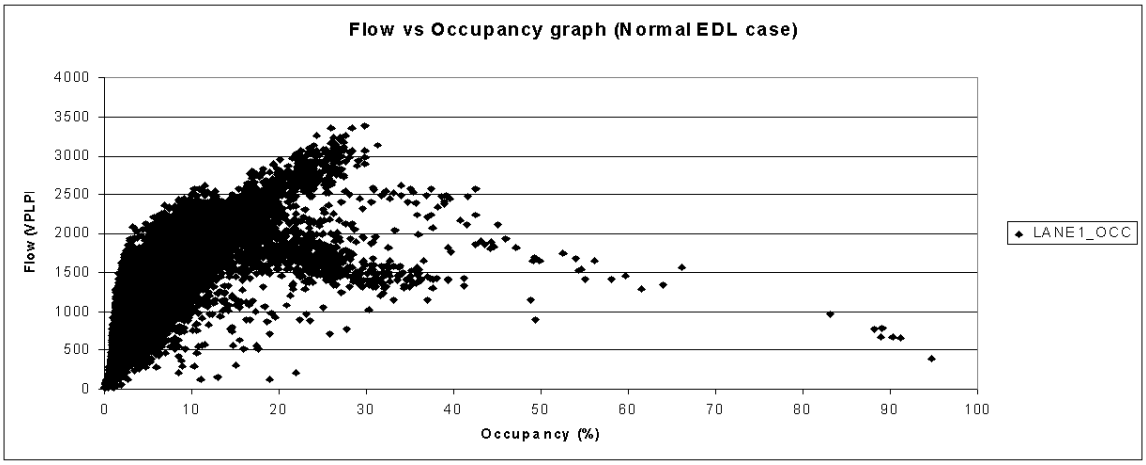


Figure 29: Flow-occupancy relationship for normal effective vehicle lengths

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 5min, 15min and 1hr 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% of lane LVBR values are less than 2. During the estimation, the observations with the large occupancy values (>18%) 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 Oct. 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% of lane volume balance observations are less than 2 and approximately 5% of the observations are greater than 10.

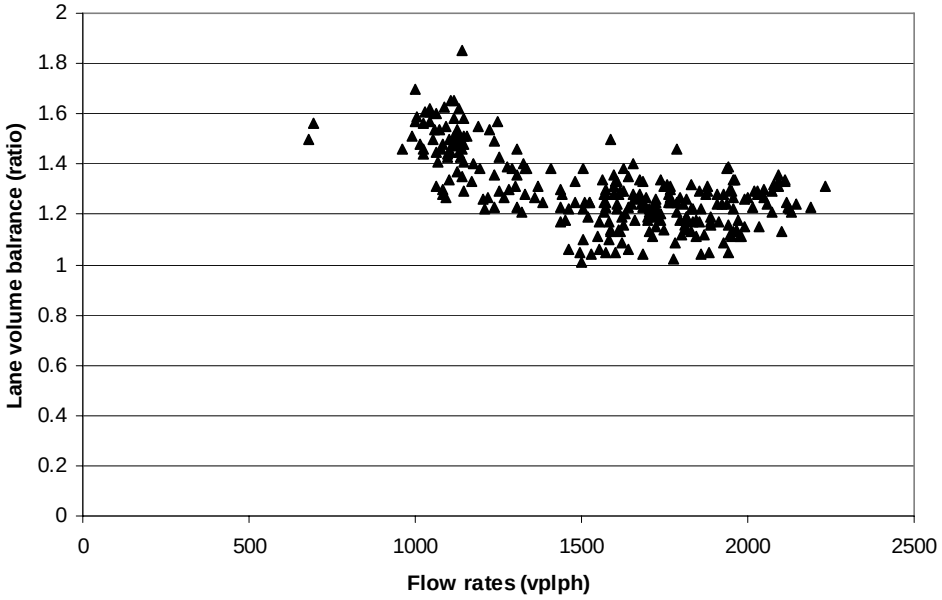


Figure 30: Effect of flow rate on lane volume balance ratio at a selected station

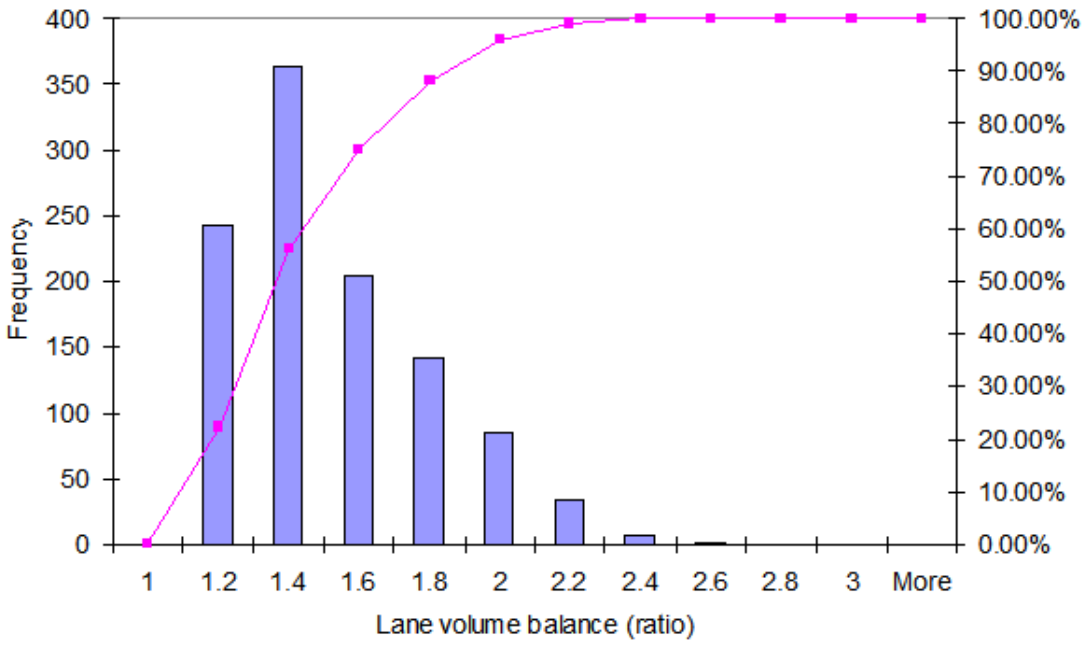


Figure 31: Lane volume balance ratio histogram from four selected stations

Table 13: Histogram of Lane Volume Balance Ratio

Lane Volume Balance Ratio	Frequency	Cumulative %
1	3348	6.17%
2	32226	65.57%
3	8917	82.01%
4	3086	87.70%
5	1557	90.57%
10	2583	95.33%
15	1061	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 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% 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.

Table 14: Frequency of LVBR Greater than 10.0

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

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 33shows, this station is located at the south of Norwood Ave 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% confidence bounds based on the computed standard deviation. The highest flow rates are observed during the morning peak around 7:30 AM.

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.

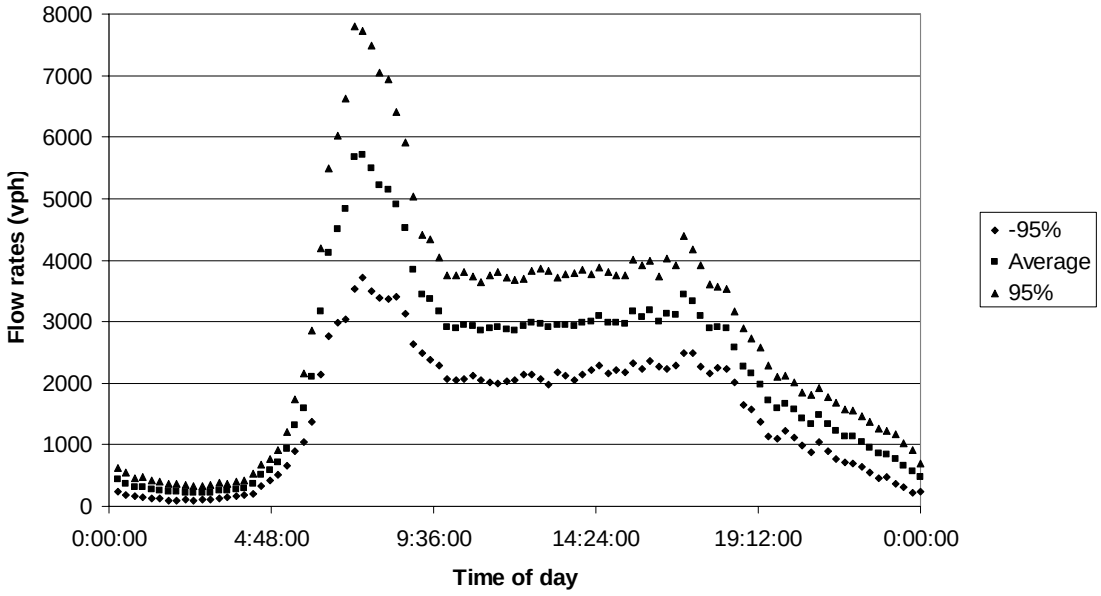


Figure 34: Average flow rates by time of day for a selected station

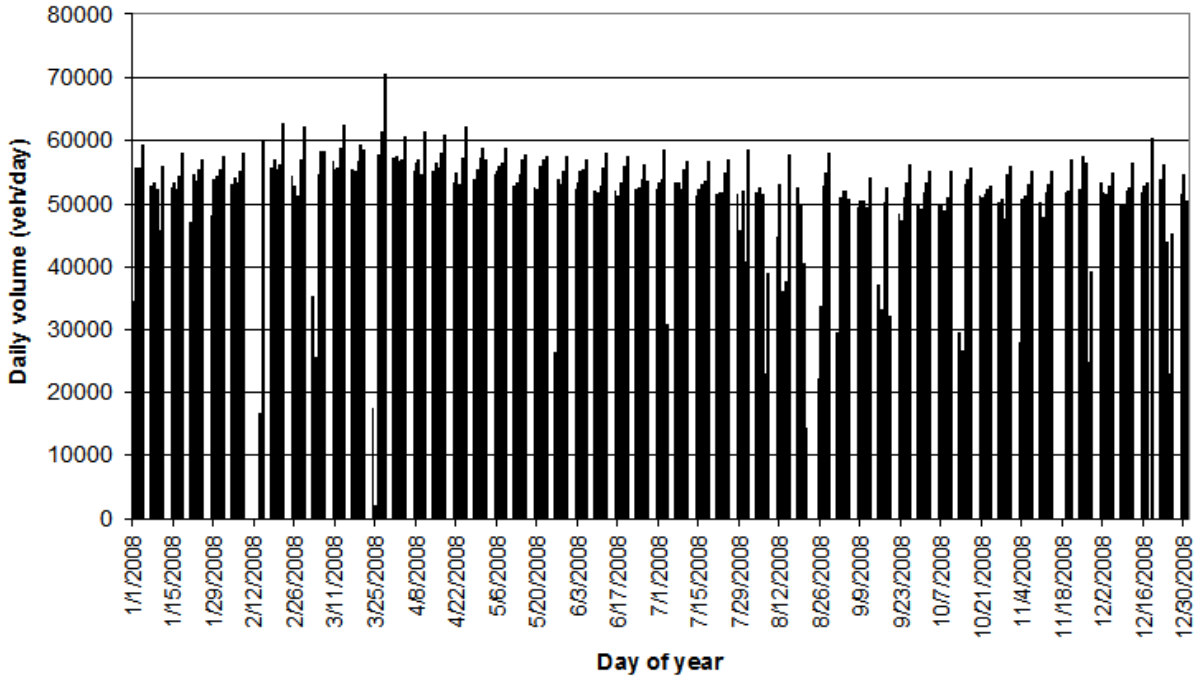


Figure 35: Total 24 hr weekday volume for one year at a selected station

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: District2 I-95 northbound, south of I-10
- Date: Oct. 6th, 2008, Monday
- Time: Morning peak (7:00 AM~10:00 AM) 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 ft.

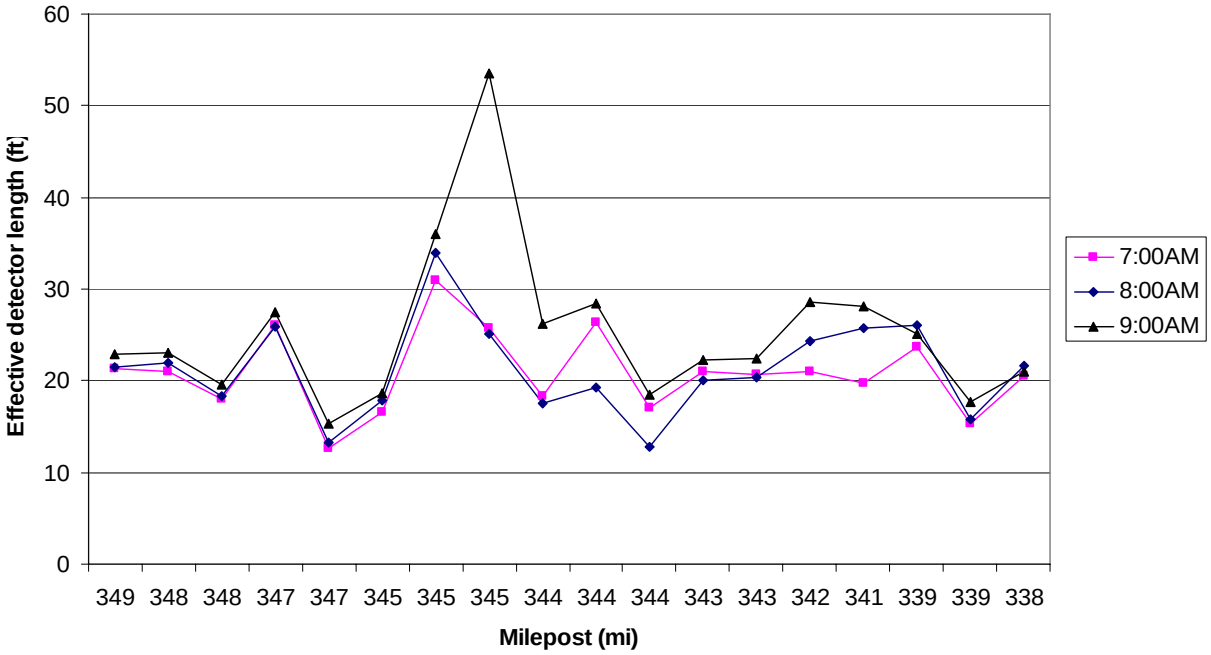


Figure 36: EVLs at stations on a section of I-95

The EVLs were fairly consistent over the three hour period at most stations. Only one station showed values that were consistently above 30 ft. One station exhibited a peak value above 50 ft, 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 1 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 9 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.

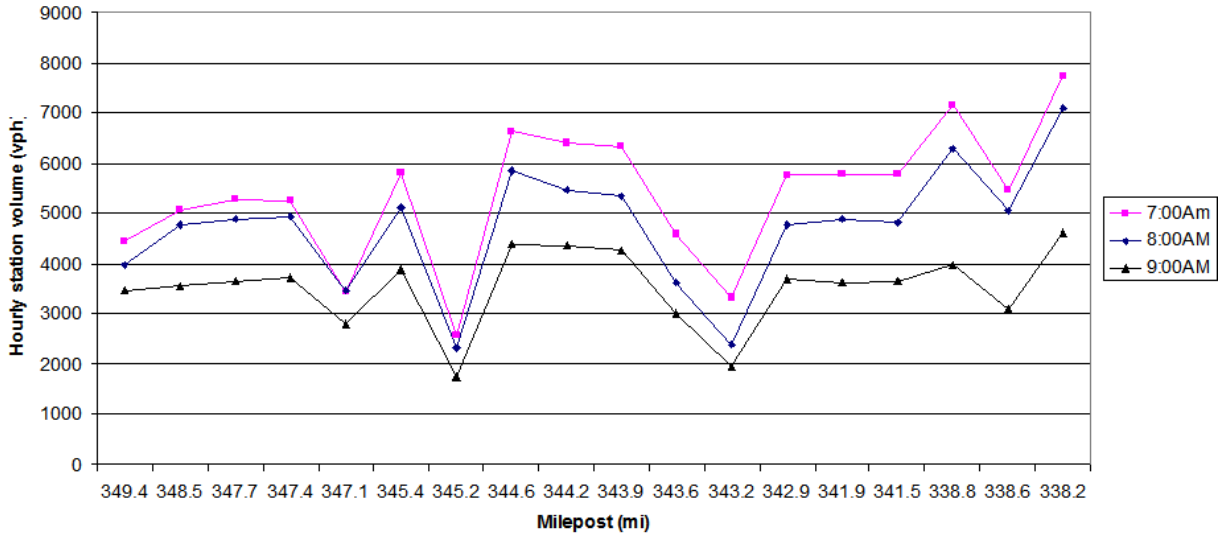


Figure 37: Hourly traffic volumes at stations on a section of I-95

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% 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 1440 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	N/A	Date
TIME		
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(2200 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 2/3 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 1000 ft north and 700 ft 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.

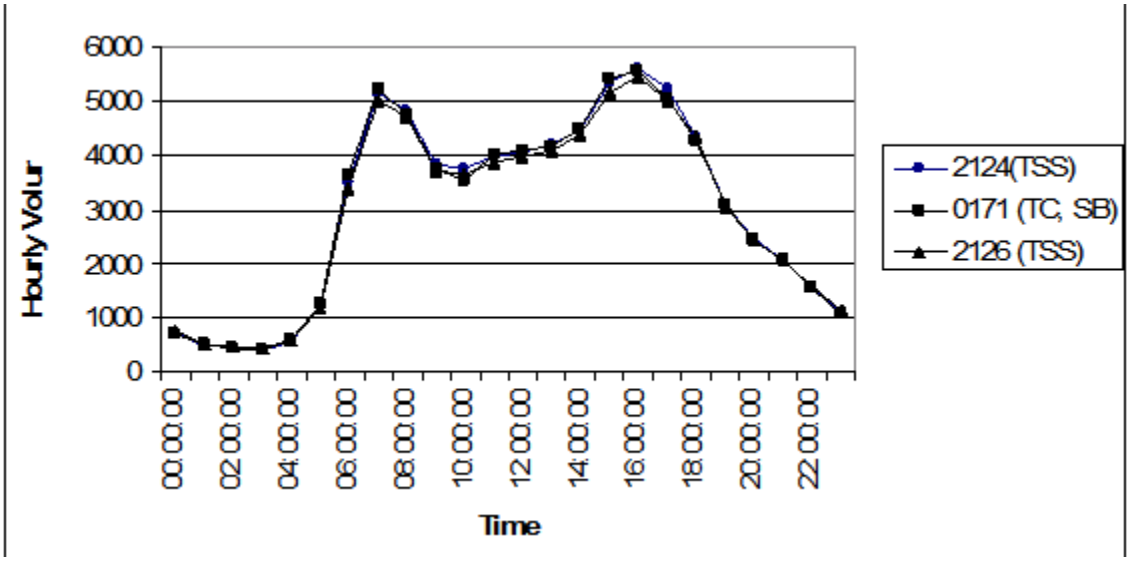


Figure 38: Example SunGuide TSS and Statistics Office count comparison

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.

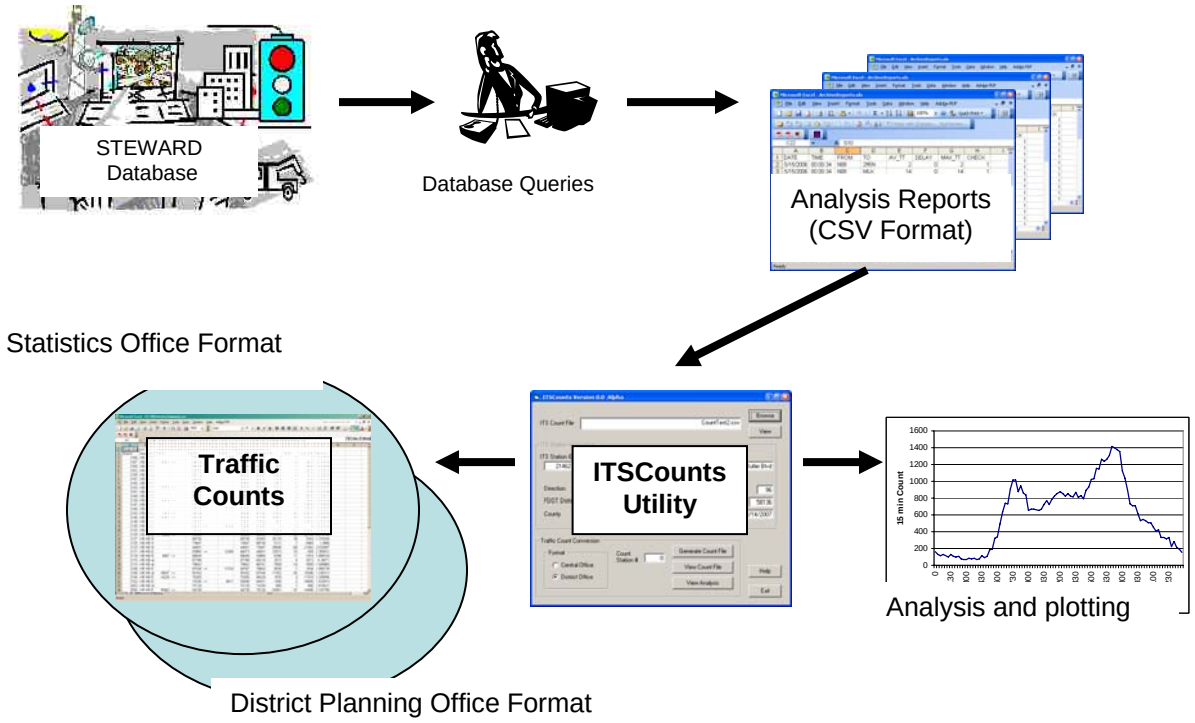


Figure 39: Overview of the ITSCounts data flow

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 website.

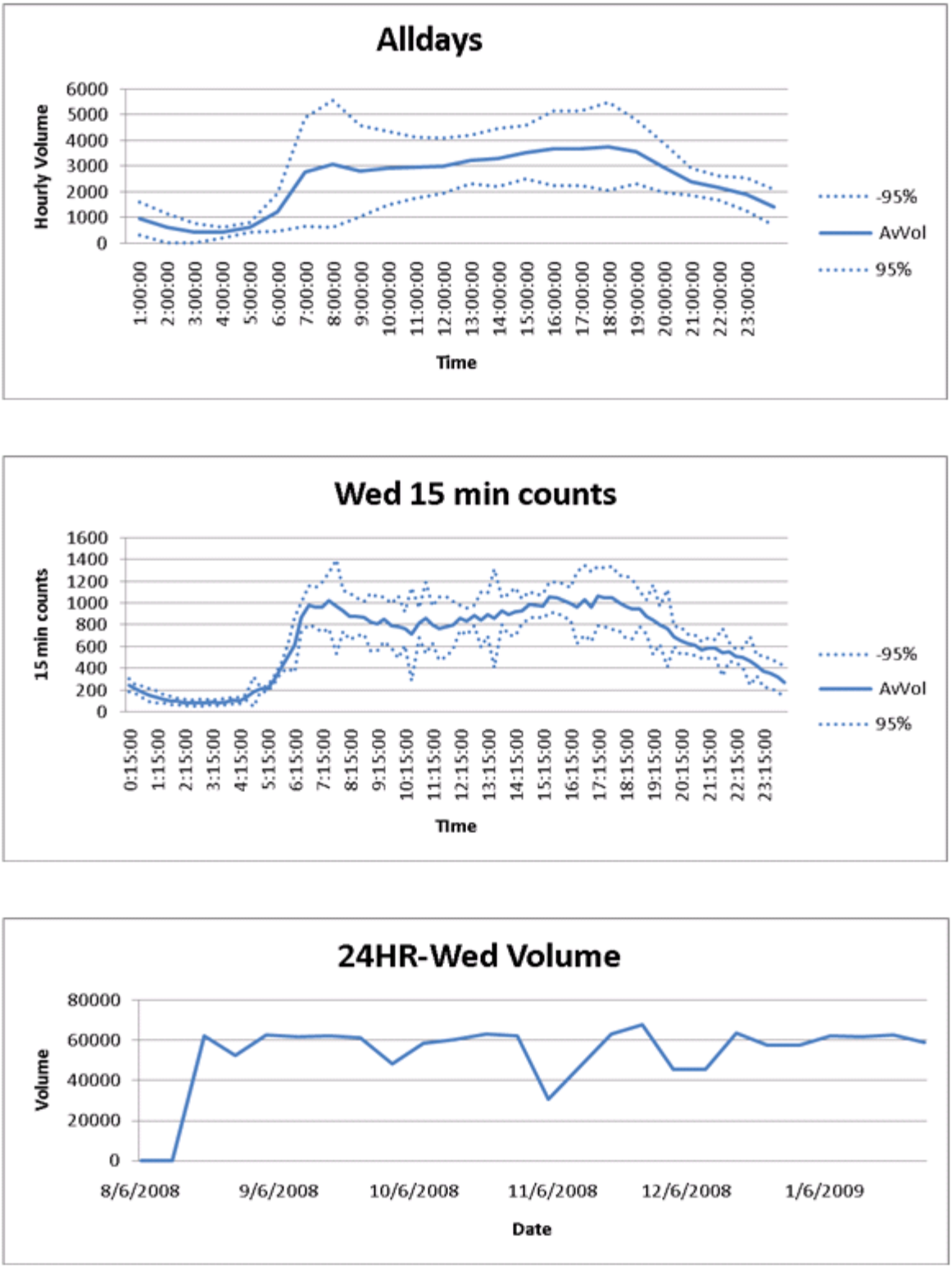


Figure 40: Sample page from a facility count analysis report

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, etc. 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 study is being performed by the University of Florida, in cooperation with the FDOT Jacksonville Traffic Management Center and the Florida Highway Patrol. 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, PA 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 website 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 1 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 and
- 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 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:

- **I**dentification 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 (5280 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 Rd in Jacksonville. The freeway at this station carried northbound traffic in 3 lanes. The analysis period included all of October 2008 during the morning peak (7:00 AM~10:00 AM). 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 2236 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 2236 / 65.6 = 34.1 veh/mi/lane.

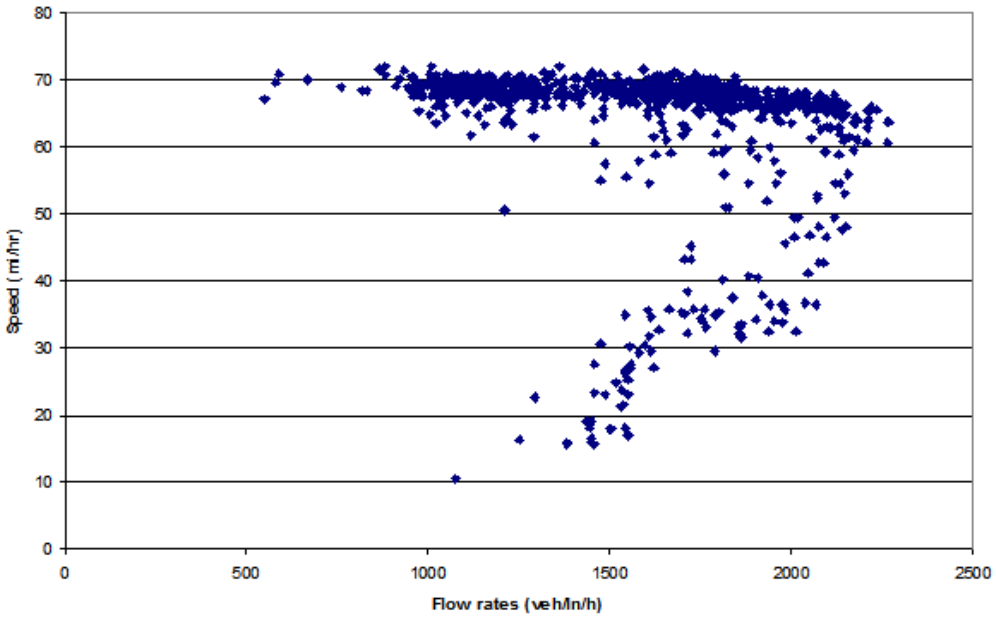


Figure 41: Example of flow rate *vs.* speed

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 17.3ft. 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.

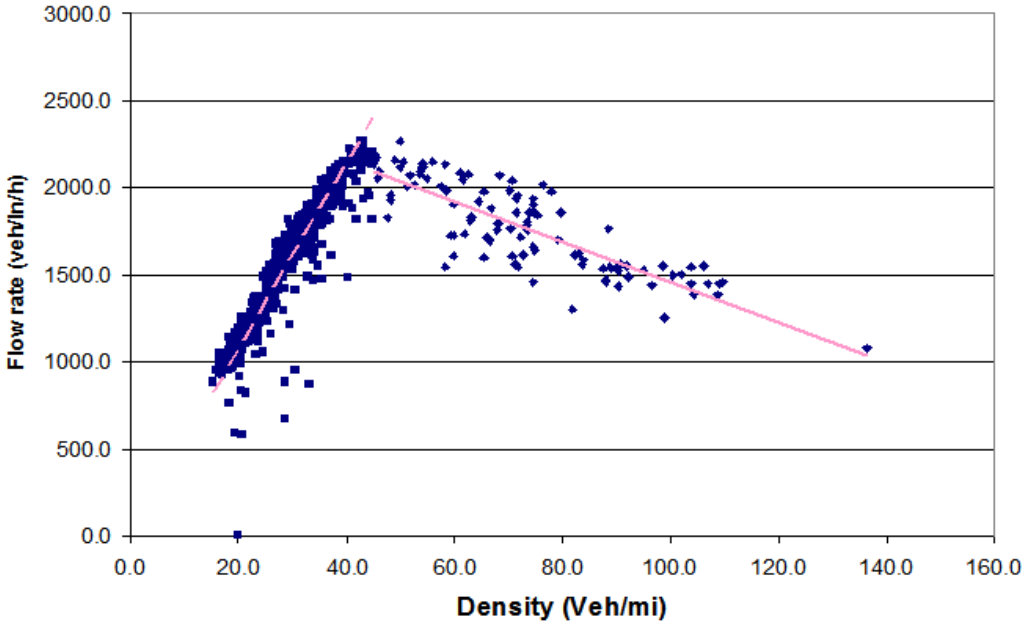


Figure 42: Example of flow rate *vs.* density

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 2260 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 5280/227 = 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 17.3ft and the speed is the measured time-mean speed.

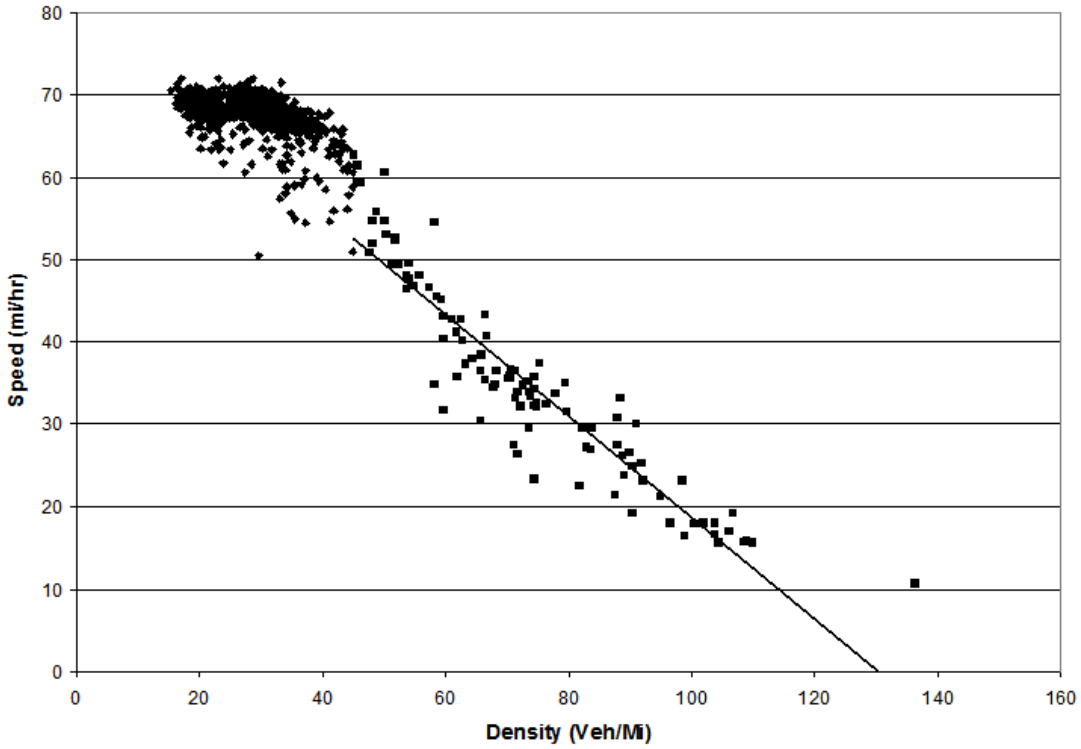


Figure 43: Example of Speed *vs.* density

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 ft 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.

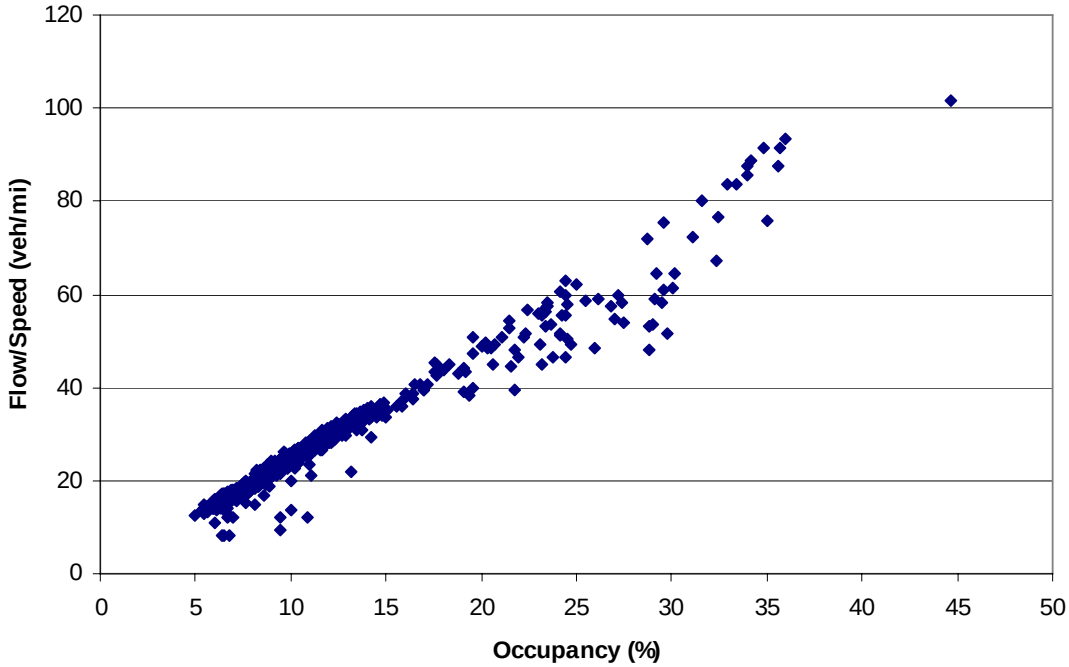


Figure 44: Comparison of density computation methods

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.

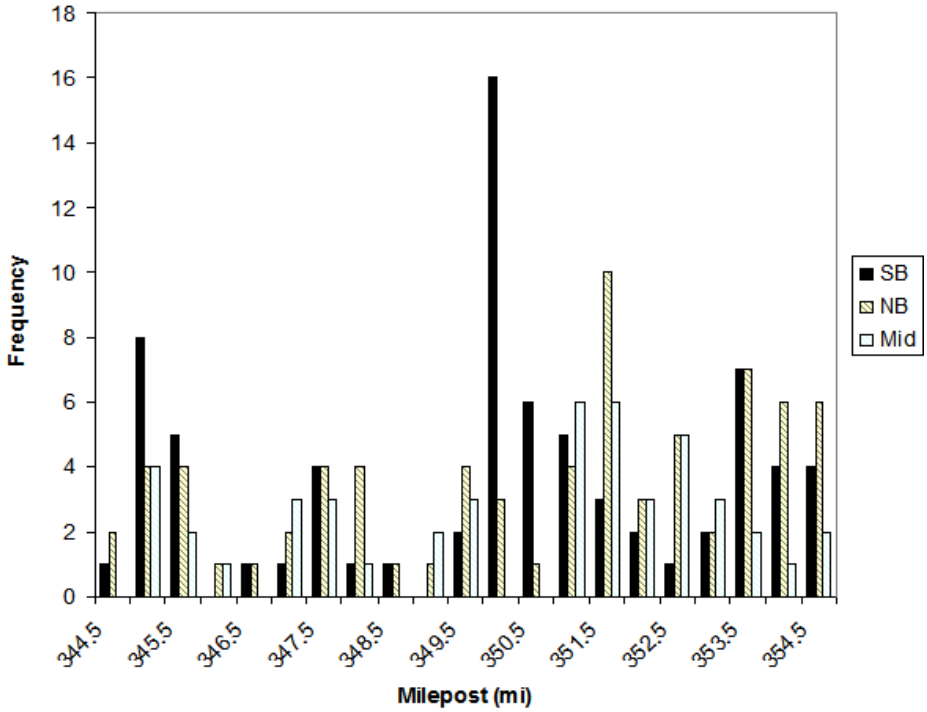


Figure 45: Number of crashes by milepost

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: 769954660
- Location: Milepost 343.947
- Date: 10-3-08 12:56PM
- 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.

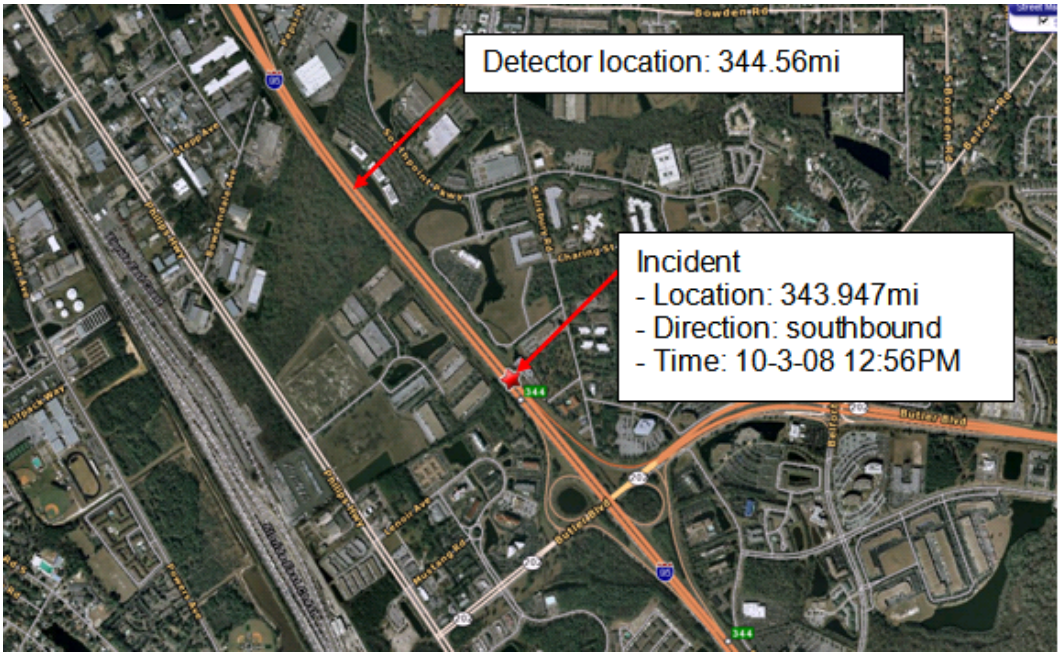


Figure 46: Incident location on a satellite photo

9.2.3 Hourly Flow Rates

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

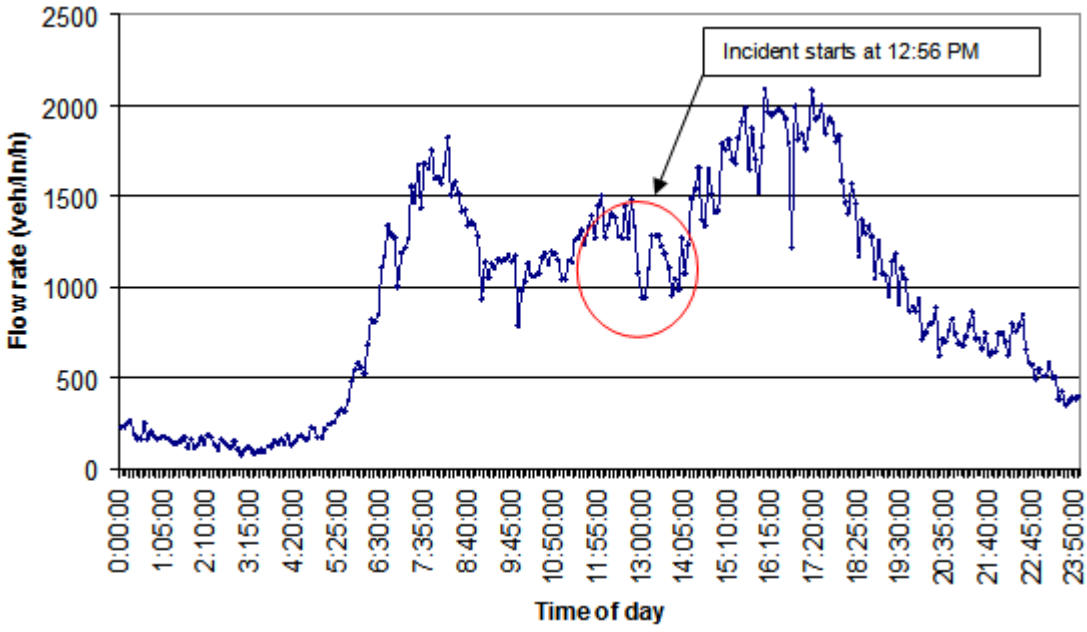


Figure 47: Hourly flow rates during the incident

CARS reported that the incident occurred at 12:56 PM. The effect is evident on the flow rate graph, which shows a decrease near 12:40 PM. The red circle in the figure shows the two flow rate drops during the incident, which occurred between 1:00 PM and 1:45 PM and the values were below the 1000 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.

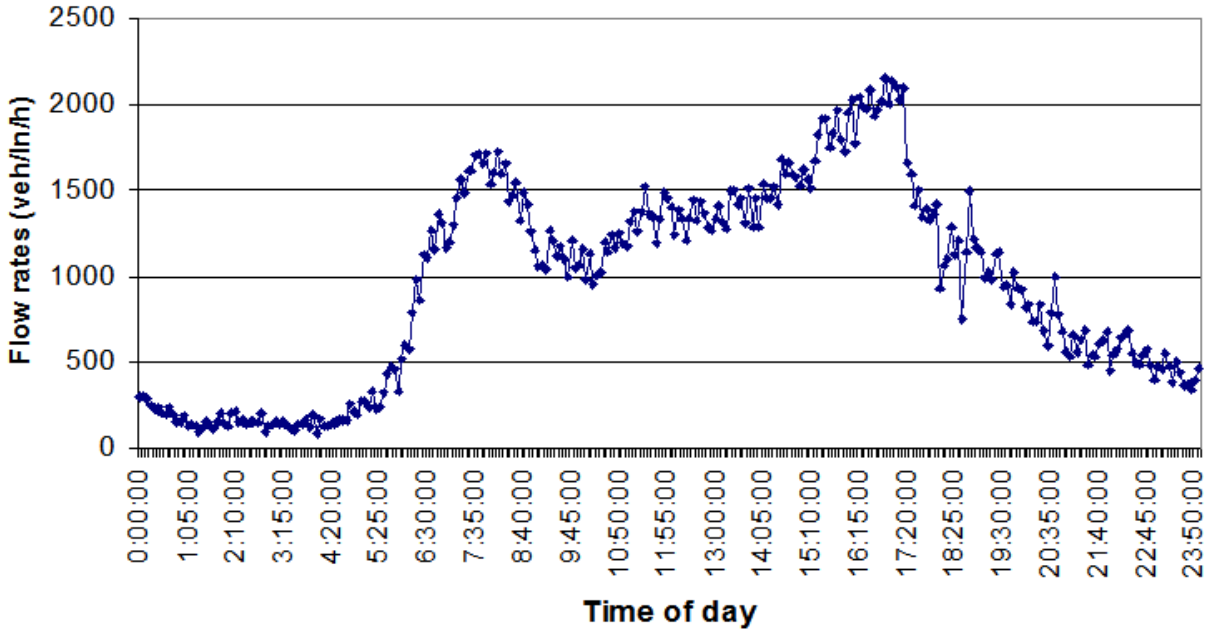


Figure 48: Hourly flow rates during incident-free operation

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.

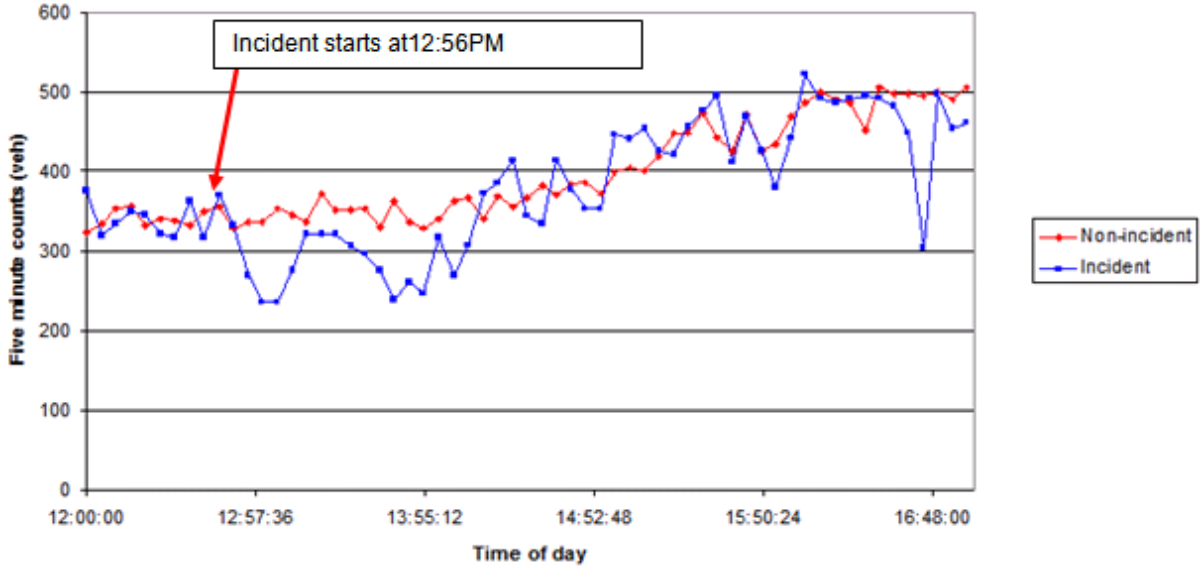


Figure 49: Comparison of 5 minute counts for the incident and non-incident case

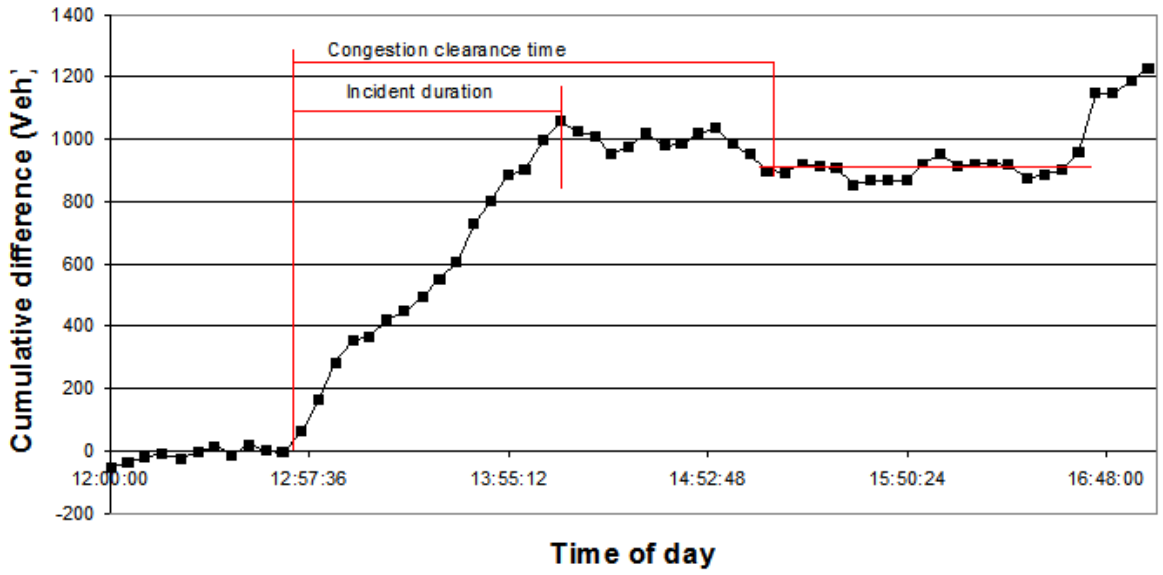


Figure 50: Cumulative differences between incident and non-incident volume counts

From this information, the delay can be estimated using queuing analysis.
Mean arrival rate (λ) = 9610veh / 145min
= 3977veh/hr

Reduced service rate (μ_R) = 4498veh/ 80min
= 3374 veh/hr

Service rate (μ) = (9610veh - 4498veh)/ 65min
= 4719 veh/hr

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

= 80 x (4719-3374) / (4719-3977)
= 145min

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

= 80min x 145min x (3977 – 3374) /2
= 971.5 veh-hr

Average vehicle delay = (971.5 veh-hr/ 9610veh)
= 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 PM to 2:30 PM, the occupancy increases over 45%.

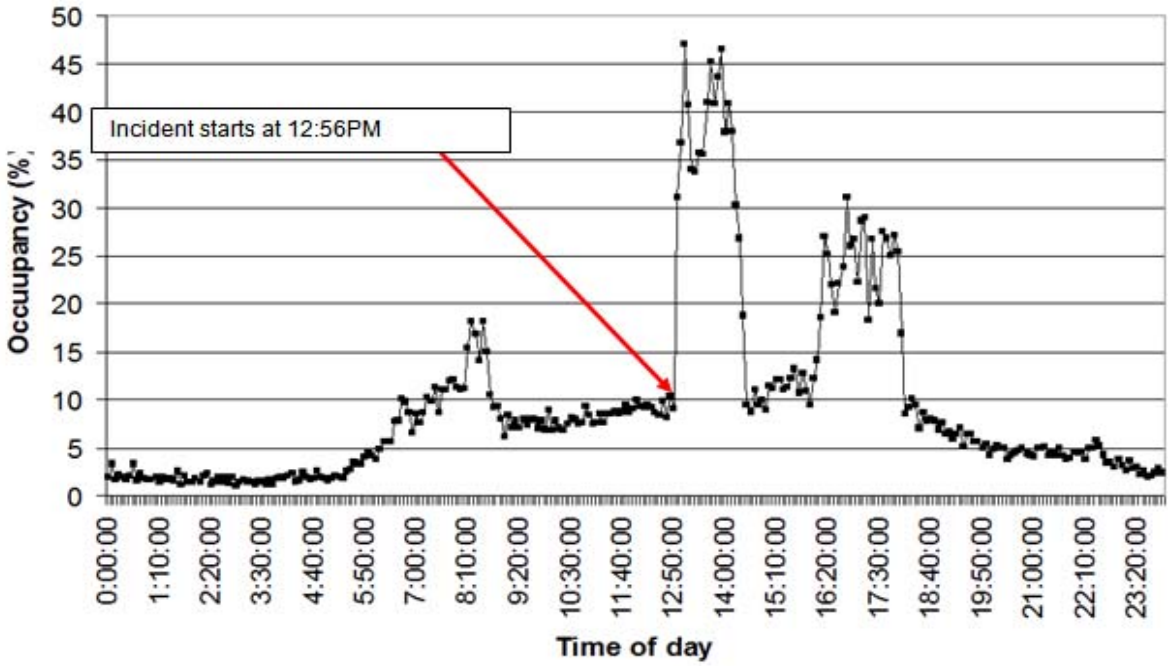


Figure 51: Effect of the incident on occupancy

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.

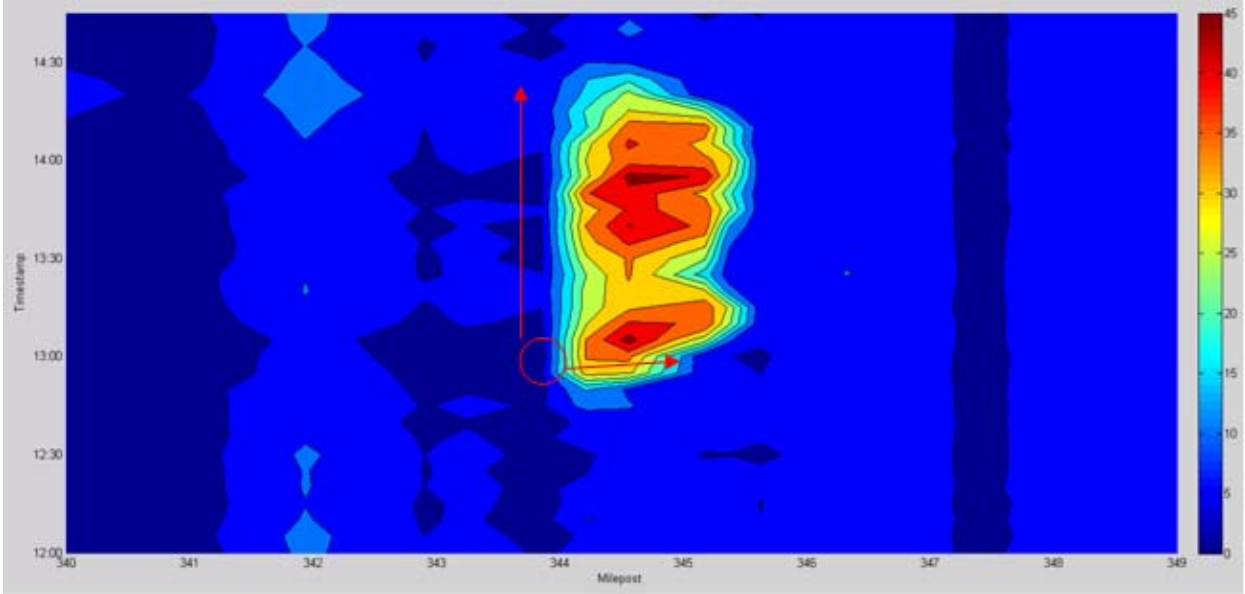


Figure 52: Occupancy contour graph during the incident

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.

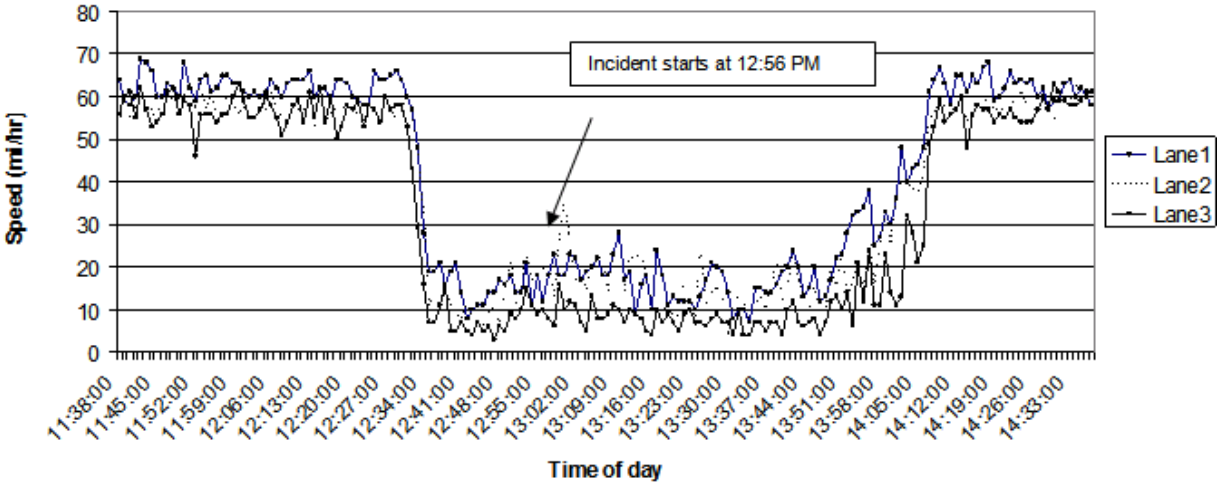


Figure 53: Effect of the incident on speed

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.

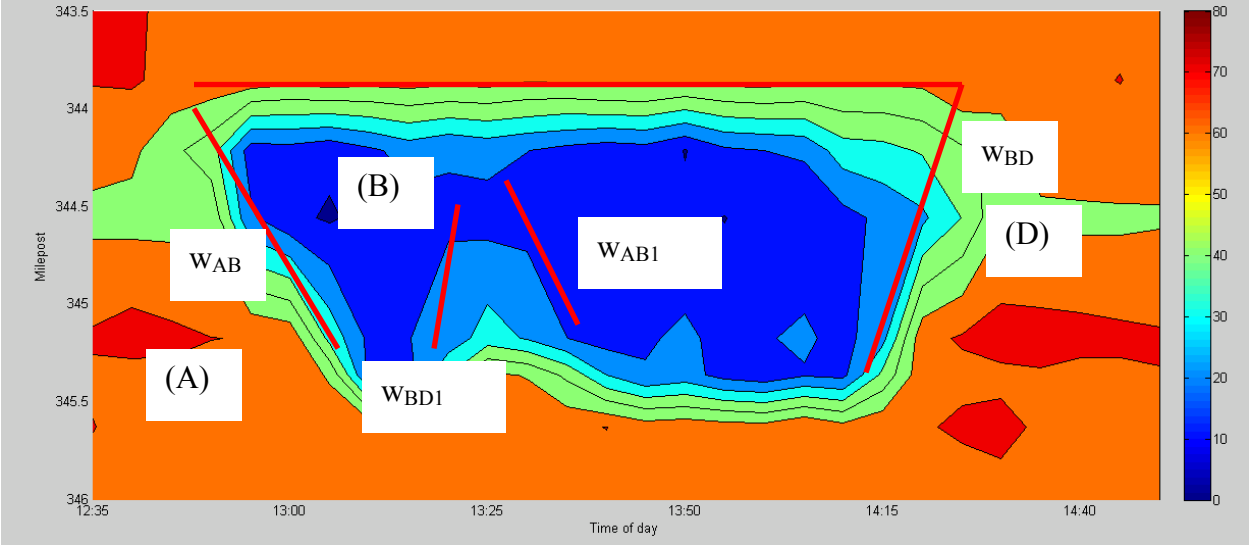


Figure 54: Speed contour graph during the incident

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$_{non\text{-}incident}$ -Total travel time$_{incident}$)
= 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/ 9610 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 2/3 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.

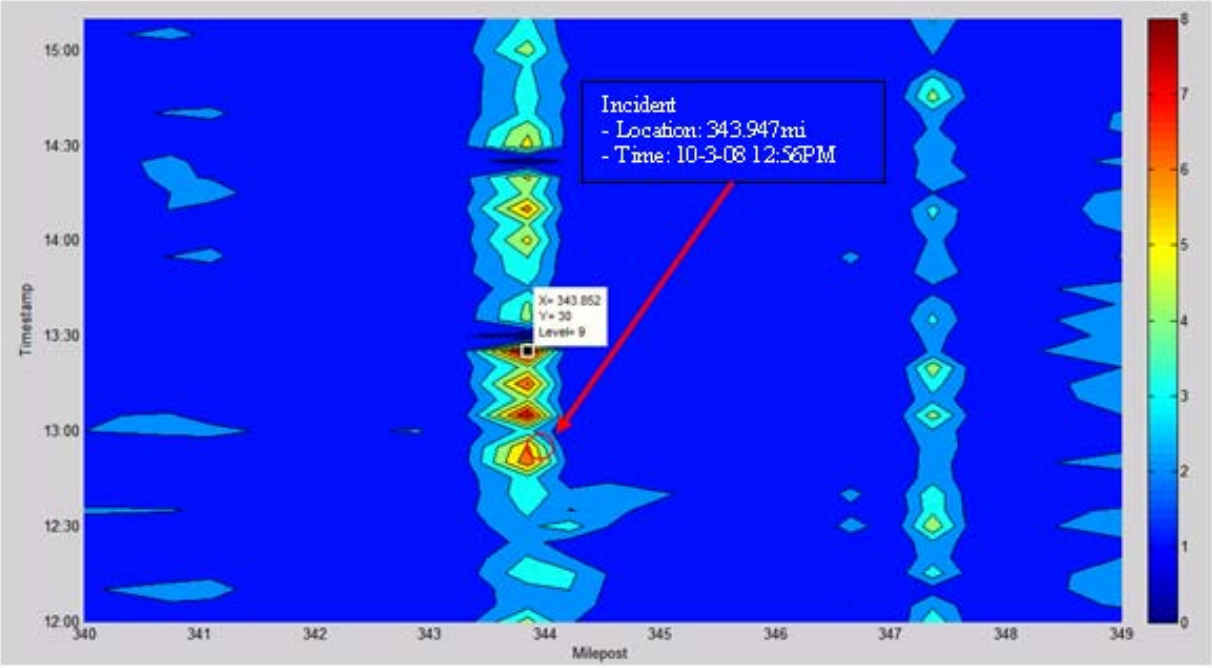


Figure 55: Lane volume balance ratio during the incident

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.

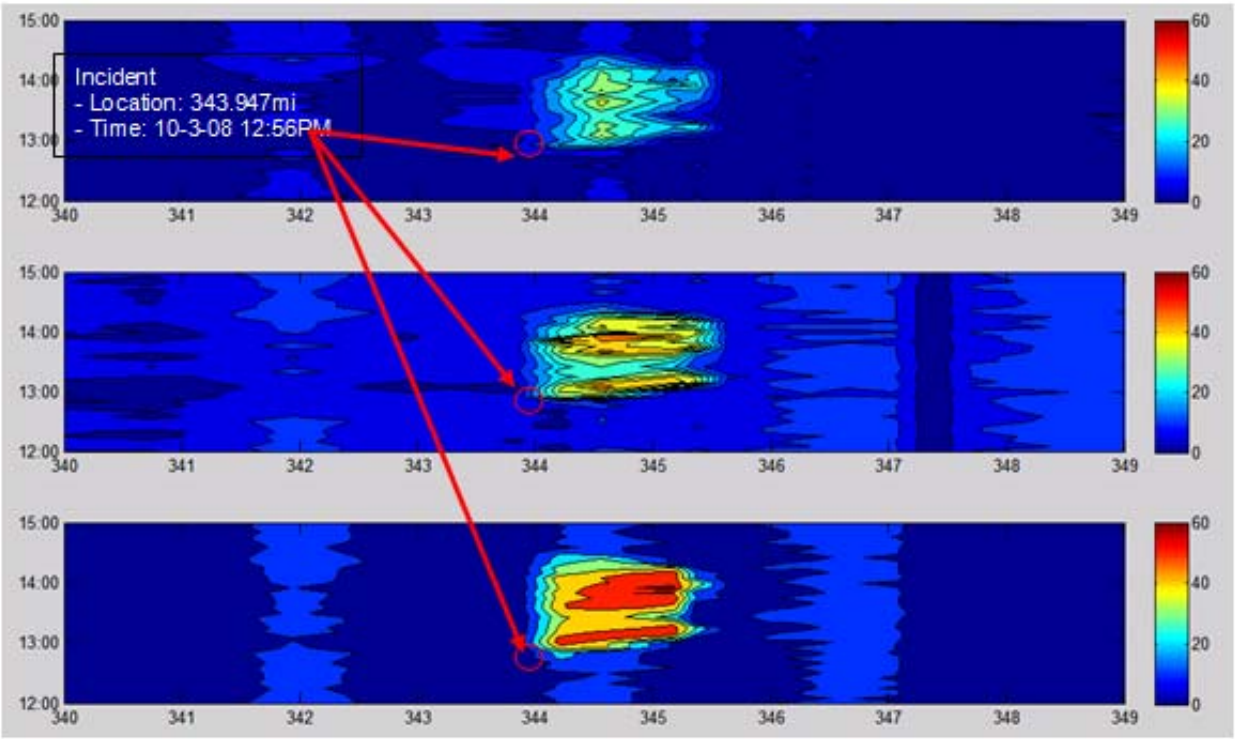


Figure 56: Occupancy contours by lane during the incident

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.

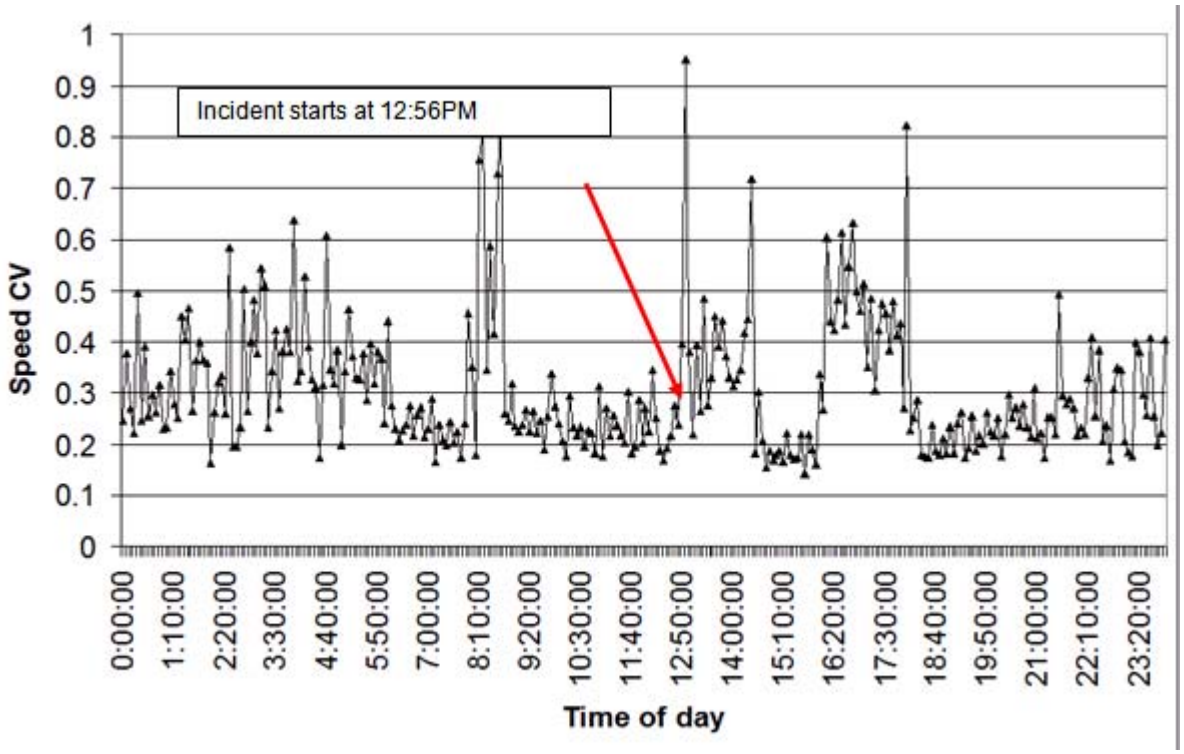


Figure 57: Speed variation by time of day

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 AM and PM 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.

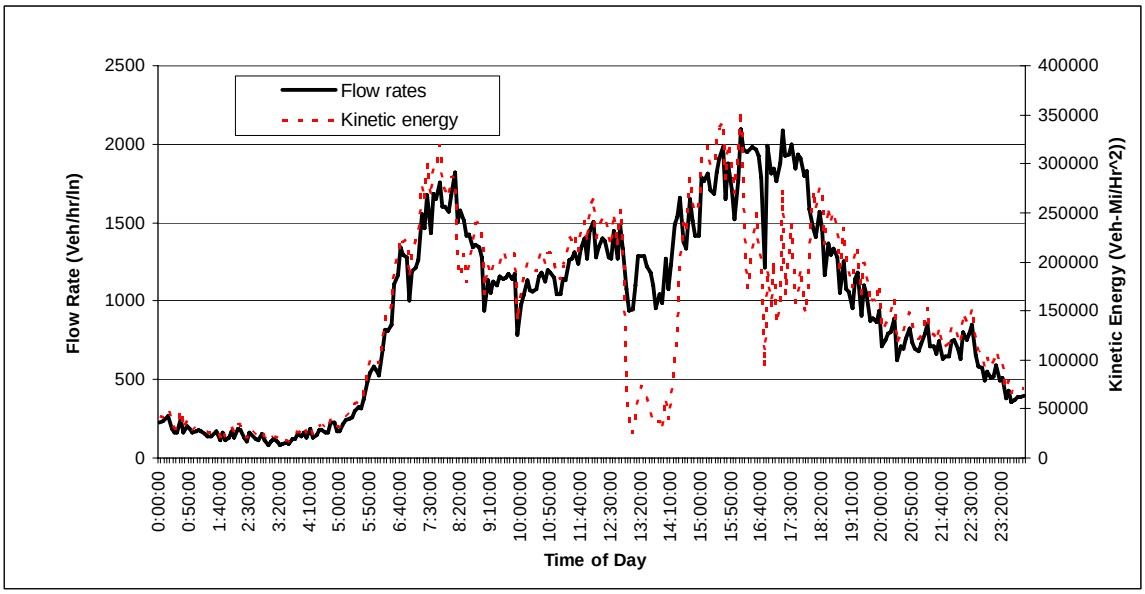


Figure 58: Flow rate and kinetic energy at the incident by time of day

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: 25145
- Vehicle-hours spent in the managed lanes: 372
- Average speed in the managed lanes (mph): 67.5
- Vehicle-miles traveled in the general lanes: 143069
- Vehicle-hours spent in the general lanes: 2374
- Average speed in the general lanes (mph): 60.3
- Vehicle-miles traveled in all lanes: 168214
- Vehicle-hours spent in all lanes: 2746
- 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: 52803
- Passenger miles traveled in the general lanes: 171682
- Total passenger miles traveled: 224486
- Passenger hours spent in the managed lanes: 781
- Passenger hours spent in the general lanes: 2848
- Total passenger hours spent: 3630
- 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 4:00 PM to 6:00 PM 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% to nearly 100% with the higher values at the south end of the section. The average value was 66.1%, indicating that about 1/3 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% 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% 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, 1AM to 5AM. 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 2900 veh/ln/h, which is approximately 20% 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 7AM to 7PM time frame.

- 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 5 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

Appendix 1: STEWARD System Description

Appendix 2: STEWARD Operation

Appendix 3: STEWARD Web Interface

Appendix 4: STEWARD Software Utility Documentation:

Appendix 1
STEWARD System Description

Table of Contents

1 Introduction 1
2 STEWARD System Overview 1
3 STEWARD Database Design and Architecture 2
3.1 External Tables 3
3.1.1 TIME_TABLE 3
3.1.2 TSS_STATION_EX 3
3.1.3 TSS_5MIN_EX, TSS_15MIN_EX and TSS_1HR_EX 3
3.1.4 TSS_5MIN_LANE_EX, TSS_15MIN_LANE_EX and TSS_1HR_LANE_EX 4
3.1.5 TSS_ETL_REPORTS_EX 4
3.2 Dimensions (Dimension Tables) 5
3.2.1 TIME_REF 5
3.2.2 DATE_REF 5
3.2.3 TSS_STATION 6
3.3 Cubes (Fact tables) 7
3.3.1 TSS_5MIN, TSS_15MIN and TSS_1HR 7
3.3.2 TSS_5MIN_LANE, TSS_15MIN_LANE and TSS_1HR_LANE 7
3.4 Materialized views 9
3.4.1 VOLUME_MAP 11
3.4.2 PM_LOS 12
3.4.3 SYSTEM_LANE 12
3.4.4 TRAFFIC_COUNTS 16
3.4.5 MAX_FLOW 18
3.4.6 EFF_VEH_LEN 21
3.5 Functions 24
3.5.1 Function Q_ALLDATAFIELDS 24
3.5.2 Function Q_EFFVEHLEN 30
3.5.3 Function Q_TRAFFICCOUNTS 33
3.5.4 Function Q_TSS 35
3.5.5 Function Q_VolumeMap 37
4 The STEWARD ETL Process 40
4.1 Mappings 40
4.2 Process flow 41
4.2.1 Pre-loading 42
4.2.2 LOADING_TSS, LOADING_TSS_LANE 43
4.3 Schedules 44
5 STEWARD Web Interface 44
5.1 Program Architecture 44
5.2 STEWARD Web Architecture 45

List of Figures

Figure 1 STEWARD system configuration 2
Figure 2 STEWARD mapping example – LOAD_TSS_15MIN 41
Figure 3 STEWARD process flow – stage3 pre-loading 42
Figure 4 STEWARD process flow – stage4 loading 43
Figure 5 STEWARD process flow – substage LOADING_TSS_15MIN 44
Figure 6 STEWARD web architecture 45

List of Tables

Table 1 STEWARD external table – contents of TIME_TABLE 3
Table 2 STEWARD external table – contents of TSS_STATION_EX 3
Table 3 STEWARD external table – contents of TSS_5MIN_EX, TSS_15MIN_EX, and TSS_1HR_EX..... 4
Table 4 STEWARD external table – contents of TSS_5MIN_LANE_EX, TSS_15MIN_LANE_EX, and TSS_1HR_LANE_EX 4
Table 5 STEWARD external table – contents of TSS_ETL_REPORTS_EX..... 5
Table 6 STEWARD dimension table – definitions of TIME_REF 5
Table 7 STEWARD dimension table – definitions of DATE_REF 6
Table 8 STEWARD dimension table – definitions of TSS_STATION 6
Table 9 STEWARD fact tables – definitions of TSS_5MIN, TSS_15MIN and TSS_1HR..... 7
Table 10 STEWARD fact tables – definitions of TSS_5MIN_LANE, TSS_15MIN_LANE and TSS_1HR_LANE 8
Table 11 STEWARD materialized views for TSS reports 9
Table 12 STEWARD web reports and their corresponding materialized views 10
Table 13 STEWARD materialized views – definitions of VOLUME_MAP_5MIN, VOLUME_MAP_15MIN, and VOLUME_MAP_1HR..... 11
Table 14 Query for STEWARD materialized views-VOLUME_MAP_5MIN 11
Table 15 STEWARD materialized views – definitions of PM_LOS_5MIN 12
Table 16 Query for STEWARD materialized views-PM_LOS_5MIN..... 13
Table 17 STEWARD materialized views – definitions of SYSTEM_LANE_5MIN, SYSTEM_LANE_15MIN, and SYSTEM_LANE_1HR 14
Table 18 Query for STEWARD materialized views-SYSTEM_LANE_5MIN..... 15
Table 19 STEWARD materialized views – definitions of TRAFFIC_COUNTS_5MIN, TRAFFIC_COUNTS_15MIN, and TRAFFIC_COUNTS _1HR 16
Table 20 Query for STEWARD materialized views-TRAFFIC_COUNTS_5MIN..... 17
Table 21 STEWARD materialized views – definitions of MAX_FLOW_5MIN, MAX_FLOW_15MIN, and MAX_FLOW_1HR..... 18
Table 22 Query for STEWARD materialized views-MAX_FLOW_5MIN..... 19
Table 23 STEWARD materialized views – definitions of EFF_VEH_LEN_5MIN, EFF_VEH_LEN_15MIN, and EFF_VEH_LEN_1HR 22
Table 24 Query for STEWARD materialized views-EFF_VEH_LEN_5MIN 23
Table 25 STEWARD function – definitions of Q_AllDataFields..... 25
Table 26 STEWARD function - Q_AllDataFields 26
Table 27 STEWARD function – definitions of Q_EffVehLen 30
Table 28 STEWARD function - Q_EffVehLen..... 31
Table 29 STEWARD function – definitions of Q_TrafficCounts..... 33
Table 30 STEWARD function - Q_ TrafficCounts 34
Table 31 STEWARD function – definitions of Q_TSS 36
Table 32 STEWARD function - Q_ TSS..... 36
Table 33 STEWARD function – definitions of Q_VolumeMap 38
Table 34 STEWARD function - Q_ VolumeMap 39
Table 35 Directory architecture of STEWARD web program 46

1 Introduction

The Statewide Transportation Engineering Warehouse for Archived Regional Data (STEWARD) system was designed and developed to accommodate the statewide traffic data in Florida. STEWARD consists of three main parts, the extraction, transformation and loading (ETL) process, the database (DB), and the web user interface. This document will describe the details of this system architecture and implementation with respect to those elements.

The ETL process acquires traffic data from the district SunGuide archive files in Florida, transforms them into the STEWARD data format, and loads them into the STEWARD database. The STEWARD database was developed to archive the TSS traffic data, to create the traffic reports, and to respond to data queries from users. These tasks were developed together using the Oracle Warehouse Builder 10g2. This program was used to design and implement the STEWARD database and to maintain it. Specific operations include loading the traffic data, creating reports and backing up the system.

The user interface for the website was developed for user convenience and speed. Most of the data are processed and archived in the database when the data are loaded. Later when the users request specific a report, the report will be created from the pre-developed data tables and materialized views.

2 STEWARD System Overview

STEWARD was developed using variety of tools to design, develop, deploy and maintain the system efficiently. The Oracle database was selected as a basic requirement from Florida DOT at the beginning of the project. The Windows Server 2003 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, Visual Basic 6.0, VBScript, and ASP/JavaScript were selected for the integrated ETL processes, the database management and the web development by Transportation Research Center (TRC).

Figure 1 shows the overall STEWARD architecture. The front end is the file transfer protocol (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 into the STEWARD database. These data are 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.

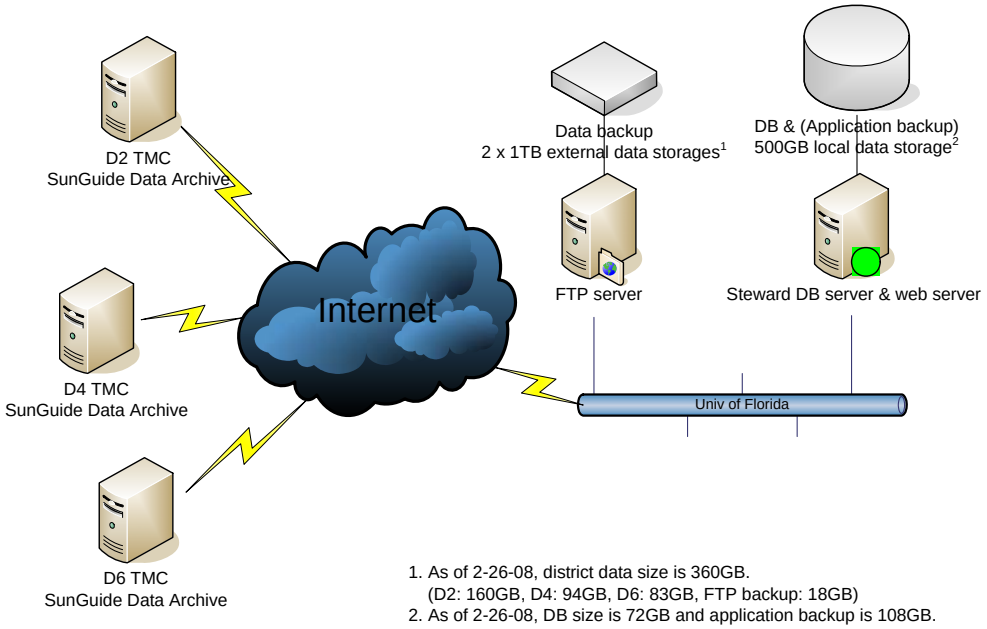


Figure 1 STEWARD system configuration

In this report, STEWARD will be described in the order of STEWARD database design, STEWARD ETL, and web interface.

3 STEWARD 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. For example, cube (fact table) has the dimensions (dimensional tables) as the foreign keys. In this case, cube TSS_5MIN has three dimensions (DATE_REF, TIME_REF and TSS_STATION) as foreign keys.

These dimensions provide the name of the dimension tables and match the content of the surrogate key for each table. The surrogate key is predefined with the name of *DIMENSION_KEY* and the number/data type. This field is filled by Oracle Warehouse Builder automatically with sequential numbers. Users can define the business key as the primary key in the dimension. For example, DATE_REF, TIME_REF and TSS_STATION dimensions have the surrogate key (DIMENSION_KEY) and the business keys, DAY, TIME and STATION_ID for each table.

In this section, the contents of each table will be described. The following table types are used:
External tables

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

3.1 External Tables

These tables are used to import the external text files into the Oracle database. The file name and all the columns attributes (name, data type and length) need to be defined to be loaded into the database. The details of each table used in STEWARD are as follows.

3.1.1 TIME_TABLE

This table is used to import the time text file into the Oracle database. The contents of the time text file is the character string from "00:00:00" to "23:59:00" with 1 minute resolution.

As shown in Table 1, the time field includes the character stings for the timestamps.

Table 1 STEWARD external table – contents of TIME_TABLE

Name	Data Type (size)
TIME	Varchar2(20)

3.1.2 TSS_STATION_EX

As shown in Table 2, this table is used to import the station information from the district facility file into the Oracle database as a dimension table.

Table 2 STEWARD external table – contents of TSS_STATION_EX

Name	Data Type (Length/Precision)
STATION_INDEX	NUMBER
STATIONCDW_NUM	NUMBER
DESCRIPTION	VARCHAR2(200)
STATUS	NUMBER
ROAD	VARCHAR2(40)
DETECTOR_UNIT	NUMBER
LATITUDE	NUMBER
LONGITUDE	NUMBER
STATE_MILEPOST	NUMBER
ROADWAY_ID	VARCHAR2(40)
ROADWAY_MILEPOST	NUMBER
MAX_SPEED	NUMBER
NUM_OF_LANES	NUMBER
UPNODE	NUMBER
LANE_CAPACITY	NUMBER
DETECTOR_TYPE	VARCHAR2(40)

3.1.3 TSS_5MIN_EX, TSS_15MIN_EX and TSS_1HR_EX

As shown in Table 3, these tables are used to import the station-level 5min, 15min and 1hr traffic data files into the Oracle database as fact tables.

Table 3 STEWARD external table – contents of TSS_5MIN_EX, TSS_15MIN_EX, and TSS_1HR_EX

Name	Data Type (Length/Precision)
DATE1	VARCHAR2(255)
TIME	VARCHAR2(255)
CDWID	NUMBER
DIRECTION	NUMBER
FWYSPD	NUMBER
FWYVOL	NUMBER
FWYOCC	NUMBER
SPDCV	NUMBER
VOLRATIO	NUMBER
SPDRATIO	NUMBER
ENTRYVOL	NUMBER
EXITVOL	NUMBER
FWYQA	NUMBER
ENTRYQA	NUMBER
EXITQA	NUMBER
HOVVOL	NUMBER
HOVSPD	NUMBER
HOVOCC	NUMBER
HOVQA	NUMBER

3.1.4 TSS_5MIN_LANE_EX, TSS_15MIN_LANE_EX and TSS_1HR_LANE_EX

As shown in Table 4, these tables are used to import the lane-level 5min, 15min and 1hr traffic data files into the Oracle database as fact tables.

Table 4 STEWARD external table – contents of TSS_5MIN_LANE_EX, TSS_15MIN_LANE_EX, and TSS_1HR_LANE_EX

Name	Data Type (Length/Precision)
DATE1	VARCHAR2(255)
TIME	VARCHAR2(255)
LANE_ID	NUMBER
SPEED	NUMBER
VOLUME	NUMBER
OCCUPANCY	NUMBER
OBSERVATIONS	NUMBER
SPD_CV	NUMBER

3.1.5 TSS_ETL_REPORTS_EX

As shown in Table 5, this table is used to provide the TSS ETL report information to web user. All the report information that web users can access is provided by the STEWARD database for easy maintenance.

Table 5 STEWARD external table – contents of TSS_ETL_REPORTS_EX

Name	Data Type (Length/Precision)
NAME	VARCHAR2(255)
SOURCE	VARCHAR2(255)
PATH	VARCHAR2(255)
CREATION_DATE	DATE
EXPIRATION_DATE	DATE

3.2 Dimensions (Dimension Tables)

Dimension tables are used to archive the attributes used to constrain and group data for the cube (fact table). In the Oracle Warehouse Builder these are named as dimensions.

In a data warehouse, date (day and time) are generally used as the keys in the fact tables. In STEWARD, the date is also used as one of the keys to identify the traffic data. It could be divided into Date and Time fields to minimize the reference table size. For example, five-year traffic data with 5 minute resolution needs 525,600 rows (5 years * 365 days * 24 hrs * 12) in one date table. But if the time and date are divided into two tables, it requires 1825 rows of date table (5 years * 365 days) and 288 rows of time table (24 hrs * 12). This would decrease the table space and improve the query performance.

Facility data and date/time data are used as the dimension tables. All of the tables except DATE_REF have no hierarchies and therefore there is only one level for all dimension tables. DATE_REF is created using the Warehouse Builder time wizard for dimension tables. It has four levels of hierarchy with predefined structures. The details of each table used in STEWARD are as follows.

3.2.1 TIME_REF

As shown in Table 6, this table is used for the time stamp text from “00:00:00” to “23:59:00” with 1 minute resolution.

Table 6 STEWARD dimension table – definitions of TIME_REF

Name	Data Type (Length/Precision)	Key
DIMENSION_KEY	NUMBER	Primary
ID	NUMBER	Surrogate
TIME	VARCHAR2(25)	
DESCRIPTION	VARCHAR2(40)	

3.2.2 DATE_REF

As shown in Table 7, this table is used for the date data from 2006 for 10 years. This table is designed and deployed in the default format in Warehouse Builder program.

Table 7 STEWARD dimension table – definitions of DATE_REF

Name	Data Type (Length/Precision)	Key
DIMENSION_KEY	NUMBER	Primary
ID	NUMBER	Surrogate
DAY	DATE	
CODE	NUMBER	
START_DATE	DATE	
END_DATE	DATE	
TIME_SPAN	NUMBER	
JULIAN_DATE	NUMBER	
DESCRIPTION	VARCHAR2(2000)	
NAME	VARCHAR2(25)	
DAY_OF_CAL_WEEK	NUMBER	
DAY_OF_CAL_MONTH	NUMBER	
DAY_OF_CAL_QUARTER	NUMBER	
DAY_OF_CAL_YEAR	NUMBER	
CAL_MONTH_NUMBER	NUMBER	
MONTH_OF_QUARTER	NUMBER	
MONTH_OF_YEAR	NUMBER	
CAL_QUARTER_NUMBER	NUMBER	
QUARTER_OF_YEAR	NUMBER	
CAL_YEAR_NUMBER	NUMBER	

3.2.3 TSS_STATION

As shown in Table 8, this table is used for the TSS station data from TSS facility data file.

Name	Data Type (Length/Precision)	Key
DIMENSION_KEY	NUMBER	Primary
ID	NUMBER	Surrogate
STATION_ID	NUMBER	
DESCRIPTION	VARCHAR2(200)	
STATUS	NUMBER	
ROAD	VARCHAR2(40)	
DETECTOR_TYPE	VARCHAR2(40)	
DETECTOR_UNIT	NUMBER	
LATITUDE	NUMBER	
LONGITUDE	NUMBER	
STATE_MILEPOST	NUMBER	
ROADWAY_ID	VARCHAR2(40)	
ROADWAY_MILEPOST	NUMBER	
SPEED_LIMIT	NUMBER	
NUM_OF_LANES	NUMBER	
UPNODE	NUMBER	
LANE_CAPACITY	NUMBER	
COUNTY	NUMBER	
COUNT_STATION	VARCHAR2(40)	

Table 8 STEWARD dimension table – definitions of TSS_STATION

3.3 Cubes (Fact tables)

Cubes (Fact tables) archive business facts or measures. It also has the foreign keys which refer to candidate keys (normally primary keys) in the dimension tables. In the Oracle Warehouse Builder, these are called as cubes. In STEWARD system, TSS data tables are cubes which includes 5 min, 15 min and 1 hr detector station and lane data.

3.3.1 TSS_5MIN, TSS_15MIN and TSS_1HR

As shown in Table 9, these tables archive the station-level TSS data with day and time stamp. They have the same foreign keys and attributes. TIME_REF, DATE_REF and TSS_STATION are used as foreign keys.

Table 9 STEWARD fact tables – definitions of TSS_5MIN, TSS_15MIN and TSS_1HR

Name	Data Type (Length/Precision)	Key
DATE_REF	NUMBER	Foreign
TIME_REF	NUMBER	Foreign
TSS_STATION	NUMBER	Foreign
FWY_SPD	NUMBER	
FWY_VOL	NUMBER	
FWY_OCC	NUMBER	
SPD_CV	NUMBER	
VOL_RATIO	NUMBER	
SPD_RATIO	NUMBER	
ENTRY_VOL	NUMBER	
EXIT_VOL	NUMBER	
FWY_QA	NUMBER	
ENTRY_QA	NUMBER	
EXIT_QA	NUMBER	
HOV_SPD	NUMBER	
HOV_VOL	NUMBER	
HOV_OCC	NUMBER	
HOV_QA	NUMBER	

3.3.2 TSS_5MIN_LANE, TSS_15MIN_LANE and TSS_1HR_LANE

As shown in Table 10, these tables archive the lane-level TSS data with day and time stamp. They have the same foreign keys and attributes. TIME_REF, DATE_REF and TSS_STATION are used as foreign keys.

Table 10 STEWARD fact tables – definitions of TSS_5MIN_LANE, TSS_15MIN_LANE and TSS_1HR_LANE

Name	Data Type (Length/Precision)	Key
DATE_REF	NUMBER	Foreign
TIME_REF	NUMBER	Foreign
TSS_STATION	NUMBER	Foreign
LANE1_SPD	NUMBER	
LANE1_VOL	NUMBER	
LANE1_OCC	NUMBER	
LANE1_OBS	NUMBER	
LANE2_SPD	NUMBER	
LANE2_VOL	NUMBER	
LANE2_OCC	NUMBER	
LANE2_OBS	NUMBER	
LANE3_SPD	NUMBER	
LANE3_VOL	NUMBER	
LANE3_OCC	NUMBER	
LANE3_OBS	NUMBER	
LANE4_SPD	NUMBER	
LANE4_VOL	NUMBER	
LANE4_OCC	NUMBER	
LANE4_OBS	NUMBER	
LANE5_SPD	NUMBER	
LANE5_VOL	NUMBER	
LANE5_OCC	NUMBER	
LANE5_OBS	NUMBER	
LANE6_SPD	NUMBER	
LANE6_VOL	NUMBER	
LANE6_OCC	NUMBER	
LANE6_OBS	NUMBER	
ONRAMP1_SPD	NUMBER	
ONRAMP1_VOL	NUMBER	
ONRAMP1_OCC	NUMBER	
ONRAMP1_OBS	NUMBER	
ONRAMP2_SPD	NUMBER	
ONRAMP2_VOL	NUMBER	
ONRAMP2_OCC	NUMBER	
ONRAMP2_OBS	NUMBER	
ONRAMP3_SPD	NUMBER	
ONRAMP3_VOL	NUMBER	
ONRAMP3_OCC	NUMBER	
ONRAMP3_OBS	NUMBER	
OFFRAMP1_SPD	NUMBER	
OFFRAMP1_VOL	NUMBER	
OFFRAMP1_OCC	NUMBER	
OFFRAMP1_OBS	NUMBER	
OFFRAMP2_SPD	NUMBER	
OFFRAMP2_VOL	NUMBER	
OFFRAMP2_OCC	NUMBER	
OFFRAMP2_OBS	NUMBER	
OFFRAMP3_SPD	NUMBER	
OFFRAMP3_VOL	NUMBER	
OFFRAMP3_OCC	NUMBER	
OFFRAMP3_OBS	NUMBER	

3.4 Materialized views

A materialized view in database system is a virtual table that represents the result of a database query. A materialized view takes a different approach in which the query result is cached as a concrete table that may be updated from the original base tables from time to time. This enables much more efficient access, at the cost of some data being potentially out-of-date. It is most useful in data warehousing scenarios, where frequent queries of the actual base tables are extremely time consuming.

In STEWARD, the reports in the web interface are associated with these materialized views to minimize the report creation time and improve the query response time. There are two options to update these tables- *refresh on demand* and *refresh on commit.* All the views here are defined as refresh on demand. There are two reasons to select refresh on demand. In case of refresh on commit, the materialized views are updated when their base tables, such as fact tables are updated. That means the data loading (fact table) and updating (materialized views) occurs at the same time and requires more time. In this case, it is difficult to estimate the execution time of the data loading process and actually it takes more time for the data loading and updating the materialized views. STEWARD expects the system operators to upload the base tables (fact tables) at which point it updates related materialized view.

Second, materialized views cannot refresh on commit if they include the aggregate functions, such as sum or average. These functions increase the reparative calculation for each data insertion and cause extreme execution time delay. One of materialized views, PM_LOS_5MIN uses sum function and cannot be used in the refresh-on-commit mode.

All of these materialized views are partitioned with date and district. The date is categorized by each quarter from 2007 Q3 to 2013 Q4. The district is categorized by District 1 to 7 and the Turnpike. Also the materialized views are indexed with station ID, hour and date to improve the performance. A bitmap index is used for this indexing. Table 11 shows STEWARD materialized views for TSS reports and its dependent tables.

Table 11 STEWARD materialized views for TSS reports

Materialized views	Base tables
VOLUME_MAP_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN
VOLUME_MAP_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN
VOLUME_MAP_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR
PM_LOS_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN
SYSTEM_LANE_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN_LANE
SYSTEM_LANE_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN_LANE
SYSTEM_LANE_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR_LANE
TRAFFIC_COUNTS_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN_LANE
TRAFFIC_COUNTS_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN_LANE
TRAFFIC_COUNTS_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR_LANE
MAX_FLOW_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN_LANE
MAX_FLOW_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN_LANE
MAX_FLOW_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR_LANE
EFF_VEH_LEN_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN_LANE
EFF_VEH_LEN_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN_LANE
EFF_VEH_LEN_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR_LANE
TT_REL_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN

Table 12 shows the STEWARD web reports and their corresponding materialized views.

Table 12 STEWARD web reports and their corresponding materialized views

	STEWARD web report	Materialized views and functions	Base tables
TSS Facility-Level Reports	All Data Fields	SYSTEM_LANE_5MIN SYSTEM_LANE_15MIN SYSTEM_LANE_1HR q_AllDataFields (function)	TSS_5MIN_LANE TSS_15MIN_LANE TSS_1HR_LANE TSS_5MIN_LANE
	Volume map and I/O Balance	VOLUME_MAP_5MIN VOLUME_MAP_15MIN VOLUME_MAP_1HR q_VolumeMap (function)	TSS_5MIN TSS_15MIN TSS_1HR TSS_5MIN
	Traffic Counts	TRAFFIC_COUNTS_5MIN TRAFFIC_COUNTS_15MIN TRAFFIC_COUNTS_1HR Q_TrafficCounts (function)	TSS_5MIN_LANE TSS_15MIN_LANE TSS_1HR_LANE
TSS Section-Level Reports	Performance Measure	PM_5MIN	TSS_5MIN
	Travel Time Relaibility	TT_REL_5MIN	TSS_5MIN
TSS Station-Level Reports	All Data Fields	TSS_5MIN TSS_15MIN TSS_1HR Q_TSS (function)	
	Traffic Counts	TRAFFIC_COUNTS_5MIN TRAFFIC_COUNTS_15MIN TRAFFIC_COUNTS_1HR Q_TrafficCounts (function)	TSS_5MIN_LANE TSS_15MIN_LANE TSS_1HR_LANE TSS_5MIN_LANE
	Max Flow Rates	MAX_FLOW_5MIN MAX_FLOW_15MIN MAX_FLOW_1HR	TSS_5MIN_LANE TSS_15MIN_LANE TSS_1HR_LANE
	Effective Vehicle Length	EFF_VEH_LEN_5MIN EFF_VEH_LEN_15MIN EFF_VEH_LEN_1HR Q_EffVehLen (function)	TSS_5MIN_LANE TSS_15MIN_LANE TSS_1HR_LANE TSS_5MIN_LANE

3.4.1 VOLUME_MAP

These tables provide volume map for web user queries. Table 13 shows the materialized views definition for VOLUME_MAP_5MIN, VOLUME_MAP_15MIN, and VOLUME_MAP_1HR.

Table 13 STEWARD materialized views – definitions of VOLUME_MAP_5MIN, VOLUME_MAP_15MIN, and VOLUME_MAP_1HR

Name	Data Type (Length/Precision)
DATE1	DATE
TIME	VARCHAR2(25)
FACILITY	NUMBER
STATION_ID	NUMBER
STATION_MP	NUMBER
UPNODE_ID	NUMBER
ENTRY_VOLUME	NUMBER
FWY_VOLUME	NUMBER
EXIT_VOLUME	NUMBER
LINK_INPUT	NUMBER
LINK_OUTPUT	NUMBER
DIFFERENCE	NUMBER
PCNT_DIFF	NUMBER

Table 14 shows the query to generate the materialized view- VOLUME_MAP_5MIN.

Table 14 Query for STEWARD materialized views-VOLUME_MAP_5MIN

```
SELECT
      DATE_TBL.DAY,
      TIME_TBL.TIME,
      TO_NUMBER(SUBSTR(TIME_TBL.TIME, 1, 2)),     /* Hour */
      FLOOR(STATION.STATION_ID/100000),       /* District */
      FLOOR(STATION.STATION_ID/10000)- 10*FLOOR(STATION.STATION_ID/100000),
      STATION.STATION_ID,
      MOD(STATION.STATION_ID,10),
      STATION.STATE_MILEPOST,
      UPNODE.STATION_ID,
      TSS_5MIN.ENTRY_VOL,
      TSS_5MIN.FWY_VOL,
      TSS_5MIN.EXIT_VOL,
      UPNODE_TSS_5MIN.ENTRY_VOL + UPNODE_TSS_5MIN.FWY_VOL,
      TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL,
      UPNODE_TSS_5MIN.FWY_VOL + UPNODE_TSS_5MIN.ENTRY_VOL - TSS_5MIN.EXIT_VOL -
                  TSS_5MIN.FWY_VOL,
      CASE
            WHEN (UPNODE_TSS_5MIN.FWY_VOL + UPNODE_TSS_5MIN.ENTRY_VOL + TSS_5MIN.EXIT_VOL
                  + TSS_5MIN.FWY_VOL) > 0
           THEN (100 * 2*(UPNODE_TSS_5MIN.FWY_VOL + UPNODE_TSS_5MIN.ENTRY_VOL -
                  TSS_5MIN.EXIT_VOL - TSS_5MIN.FWY_VOL) / (UPNODE_TSS_5MIN.FWY_VOL +
                  UPNODE_TSS_5MIN.ENTRY_VOL + TSS_5MIN.EXIT_VOL + TSS_5MIN.FWY_VOL))
           ELSE NULL
      END
FROM GATOR.TSS_STATION STATION, GATOR.TSS_STATION UPNODE, GATOR.TSS_5MIN TSS_5MIN,
     GATOR.TSS_5MIN UPNODE_TSS_5MIN, GATOR.DATE_REF DATE_TBL, GATOR.TIME_REF TIME_TBL
WHERE TSS_5MIN.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND STATION.UPNODE = UPNODE.STATION_ID
            AND UPNODE_TSS_5MIN.TSS_STATION = UPNODE.DIMENSION_KEY
            AND UPNODE_TSS_5MIN.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND UPNODE_TSS_5MIN.TIME_REF = TIME_TBL.DIMENSION_KEY
```

3.4.2 PM_LOS

These tables provide performance measure and LOS for web user queries. Table 15 shows the materialized views definitions for PM_LOS_5MIN.

Table 15 STEWARD materialized views – definitions of PM_LOS_5MIN

Name	Data Type (Length/Precision)
DATE1	DATE
TIME	VARCHAR2(25)
HOUR	NUMBER
STATION_ID	NUMBER
DIRECTION	NUMBER
STATION_DESC	VARCHAR2(200)
UPNODE_ID	NUMBER
UPNODE_DESC	VARCHAR2(200)
STATION_MP	NUMBER
LENGTH	NUMBER
LINK_VOL	NUMBER
NUM_OF_LANES	NUMBER
VOL_PER_LANE	NUMBER
VEH_MI	NUMBER
VEH_HR	NUMBER
SPEED	NUMBER
DELAY	NUMBER
KENERGY	NUMBER
DENSITY	NUMBER
VOL_TO_CAP	NUMBER

Table 16 shows the query to generate the materialized view- PM_LOS_5MIN.

3.4.3 SYSTEM_LANE

These tables provide system lane volume for web user queries. Table 17 shows the materialized views definitions for SYSTEM_LANE_5MIN, SYSTEM_LANE_15MIN, and SYSTEM_LANE_1HR.

Table 16 Query for STEWARD materialized views-PM_LOS_5MIN

```
SELECT
      DATE_TBL.DAY,             /* Date */
      TIME_TBL.TIME,            /* Time */
      TO_NUMBER(SUBSTR(TIME_TBL.TIME, 1, 2)),     /* Hour */
      STATION.STATION_ID,       /* Destination */
      FLOOR(STATION.STATION_ID/100000),       /* District */
      MOD(STATION.STATION_ID,10),       /* Direction */
      STATION.DESCRIPTION,      /* Destination description */
      UPNODE.STATION_ID,        /* Origin */
      UPNODE.DESCRIPTION,       /* Origin description */
      STATION.STATE_MILEPOST,   /* Destination milepost */
      ABS(STATION.STATE_MILEPOST- UPNODE.STATE_MILEPOST),   /* Length */

      ((TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL)+ (UPNODE.IN_VOLUME))/2,  /*LinkVol*/
      STATION.NUM_OF_LANES, /*Main Lane*/
      12*((TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL)+ (UPNODE.IN_VOLUME))/2
                  /STATION.NUM_OF_LANES,  /*Vol/Lane */
      ((TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL)+ (UPNODE.IN_VOLUME))/2*
                  ABS(STATION.STATE_MILEPOST- UPNODE.STATE_MILEPOST),    /*Veh-Mi*/
      ((TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL)+ (UPNODE.IN_VOLUME))/2*
                  ABS(STATION.STATE_MILEPOST- UPNODE.STATE_MILEPOST) / TSS_5MIN.FWY_SPD,
                  /*Veh-Hr*/
      TSS_5MIN.FWY_SPD,   /*AvSpeed */
     CASE    /*Delay*/
         WHEN (1.5*TSS_5MIN.FWY_SPD) >= (STATION.SPEED_LIMIT) then 0
         ELSE ((TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL)+ (UPNODE.IN_VOLUME))/2 *
                  ABS(STATION.STATE_MILEPOST- UPNODE.STATE_MILEPOST)*(1/ TSS_5MIN.FWY_SPD
                  -1.5/STATION.SPEED_LIMIT)
      END,

      ((TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL)+ (UPNODE.IN_VOLUME))/2* TSS_5MIN.FWY_SPD,
                  /*KEnergy*/
     /*Density- Analytic function Sum-Over is used to calculate density for 15min.  */
     SUM((12/3) * (TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL+UPNODE.IN_VOLUME)/2/
                  (TSS_5MIN.FWY_SPD*STATION.NUM_OF_LANES))
                    OVER(PARTITION BY TSS_5MIN.TSS_STATION, TSS_5MIN.DATE_REF ORDER BY
                  TSS_5MIN.TIME_REF RANGE BETWEEN 11 PRECEDING and CURRENT ROW),
     /*V/C- Analytic function Sum-Over is used to calculate density for 15min.  */
     SUM( (12/3) * (TSS_5MIN.FWY_VOL
                  +TSS_5MIN.EXIT_VOL+UPNODE.IN_VOLUME)/2/(STATION.LANE_CAPACITY*
                  STATION.NUM_OF_LANES))
                    OVER(PARTITION BY TSS_5MIN.TSS_STATION, TSS_5MIN.DATE_REF ORDER BY
                  TSS_5MIN.TIME_REF RANGE BETWEEN 11 PRECEDING and CURRENT ROW)

FROM GATOR.TSS_STATION STATION, (SELECT STATION_ID, DESCRIPTION, STATE_MILEPOST,
                  NVL((FWY_VOL+ENTRY_VOL),0) AS IN_VOLUME, DATE_REF, TIME_REF FROM
                  TSS_STATION, TSS_5MIN  WHERE TSS_5MIN.TSS_STATION =
                  TSS_STATION.DIMENSION_KEY) UPNODE,
          GATOR.TSS_5MIN TSS_5MIN, GATOR.DATE_REF DATE_TBL, GATOR.TIME_REF TIME_TBL

WHERE TSS_5MIN.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND STATION.UPNODE = UPNODE.STATION_ID
            AND UPNODE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND UPNODE.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND TSS_5MIN.FWY_SPD >0
            AND STATION.NUM_OF_LANES >0
            AND STATION.LANE_CAPACITY >0
```

Table 17 STEWARD materialized views – definitions of SYSTEM_LANE_5MIN, SYSTEM_LANE_15MIN, and SYSTEM_LANE_1HR

Name	Data Type (Length/Precision)
DATE1	DATE
TIME	VARCHAR2(30)
HOUR	NUMBER
DISTRICT	NUMBER
FACILITY	NUMBER
STATION_ID	NUMBER
DIRECTION	NUMBER
STATION_DESC	VARCHAR2(240)
STATION_MP	NUMBER
LANE1_VOL	NUMBER
LANE1_SPD	NUMBER
LANE1_OCC	NUMBER
LANE2_VOL	NUMBER
LANE2_SPD	NUMBER
LANE2_OCC	NUMBER
LANE3_VOL	NUMBER
LANE3_SPD	NUMBER
LANE3_OCC	NUMBER
LANE4_VOL	NUMBER
LANE4_SPD	NUMBER
LANE4_OCC	NUMBER
LANE5_VOL	NUMBER
LANE5_SPD	NUMBER
LANE5_OCC	NUMBER
LANE6_VOL	NUMBER
LANE6_SPD	NUMBER
LANE6_OCC	NUMBER
ONRAMP1_VOL	NUMBER
ONRAMP1_SPD	NUMBER
ONRAMP1_OCC	NUMBER
ONRAMP2_VOL	NUMBER
ONRAMP2_SPD	NUMBER
ONRAMP2_OCC	NUMBER
ONRAMP3_VOL	NUMBER
ONRAMP3_SPD	NUMBER
ONRAMP3_OCC	NUMBER
OFFRAMP1_VOL	NUMBER
OFFRAMP1_SPD	NUMBER
OFFRAMP1_OCC	NUMBER
OFFRAMP2_VOL	NUMBER
OFFRAMP2_SPD	NUMBER
OFFRAMP2_OCC	NUMBER
OFFRAMP3_VOL	NUMBER
OFFRAMP3_SPD	NUMBER
OFFRAMP3_OCC	NUMBER

Table 18 shows the query to generate the materialized view- SYSTEM_LANE_5MIN.

Table 18 Query for STEWARD materialized views-SYSTEM_LANE_5MIN

```
SELECT
      DATE_TBL.DAY,
      TIME_TBL.TIME,
      TO_NUMBER(SUBSTR(TIME_TBL.TIME, 1, 2)),     /* Hour */
      FLOOR(STATION.STATION_ID/100000),       /* District */
      FLOOR(STATION.STATION_ID/10000)- 10*FLOOR(STATION.STATION_ID/100000),
      STATION.STATION_ID,
      MOD(STATION.STATION_ID,10),
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      TSS_5MIN_LANE.LANE1_VOL,
      TSS_5MIN_LANE.LANE1_SPD,
      TSS_5MIN_LANE.LANE1_OCC,

      TSS_5MIN_LANE.LANE2_VOL,
      TSS_5MIN_LANE.LANE2_SPD,
      TSS_5MIN_LANE.LANE2_OCC,

      TSS_5MIN_LANE.LANE3_VOL,
      TSS_5MIN_LANE.LANE3_SPD,
      TSS_5MIN_LANE.LANE3_OCC,

      TSS_5MIN_LANE.LANE4_VOL,
      TSS_5MIN_LANE.LANE4_SPD,
      TSS_5MIN_LANE.LANE4_OCC,

      TSS_5MIN_LANE.LANE5_VOL,
      TSS_5MIN_LANE.LANE5_SPD,
      TSS_5MIN_LANE.LANE5_OCC,

      TSS_5MIN_LANE.LANE6_VOL,
      TSS_5MIN_LANE.LANE6_SPD,
      TSS_5MIN_LANE.LANE6_OCC,

      TSS_5MIN_LANE.ONRAMP1_VOL,
      TSS_5MIN_LANE.ONRAMP1_SPD,
      TSS_5MIN_LANE.ONRAMP1_OCC,

      TSS_5MIN_LANE.ONRAMP2_VOL,
      TSS_5MIN_LANE.ONRAMP2_SPD,
      TSS_5MIN_LANE.ONRAMP2_OCC,

      TSS_5MIN_LANE.ONRAMP3_VOL,
      TSS_5MIN_LANE.ONRAMP3_SPD,
      TSS_5MIN_LANE.ONRAMP3_OCC,

      TSS_5MIN_LANE.OFFRAMP1_VOL,
      TSS_5MIN_LANE.OFFRAMP1_SPD,
      TSS_5MIN_LANE.OFFRAMP1_OCC,

      TSS_5MIN_LANE.OFFRAMP2_VOL,
      TSS_5MIN_LANE.OFFRAMP2_SPD,
      TSS_5MIN_LANE.OFFRAMP2_OCC,

      TSS_5MIN_LANE.OFFRAMP3_VOL,
      TSS_5MIN_LANE.OFFRAMP3_SPD,
      TSS_5MIN_LANE.OFFRAMP3_OCC

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE,
     GATOR.DATE_REF DATE_TBL, GATOR.TIME_REF TIME_TBL

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
```

3.4.4 TRAFFIC_COUNTS

These tables provide station lane distribution for web user queries. Table 19 shows the materialized views definitions for TRAFFIC_COUNTS_5MIN, TRAFFIC_COUNTS_15MIN, and TRAFFIC_COUNTS_1HR

Table 19 STEWARD materialized views – definitions of TRAFFIC_COUNTS_5MIN, TRAFFIC_COUNTS_15MIN, and TRAFFIC_COUNTS _1HR

Name	Data Type (Length/Precision)
DATE1	DATE
TIME	VARCHAR2(30)
HOUR	NUMBER
DISTRICT	NUMBER
FACILITY	NUMBER
STATION_ID	NUMBER
NUM_OF_LANES	NUMBER
DIRECTION	VARCHAR2(5)
STATION_DESC	VARCHAR2(240)
STATION_MP	NUMBER
TOTAL	NUMBER
LANE1_VOL	NUMBER
LANE2_VOL	NUMBER
LANE3_VOL	NUMBER
LANE4_VOL	NUMBER
LANE5_VOL	NUMBER
LANE6_VOL	NUMBER
BALANCE	NUMBER
FWY_QA	NUMBER
ON_RAMP1	NUMBER
ON_RAMP2	NUMBER
ON_RAMP3	NUMBER
ON_RAMP_QA	NUMBER
OFF_RAMP1	NUMBER
OFF_RAMP2	NUMBER
OFF_RAMP3	NUMBER
OFF_RAMP_QA	NUMBER
COUNTY	NUMBER
COUNT_STATION	VARCHAR2(40)

Table 20 shows the query to generate the materialized view- TRAFFIC_COUNTS_5MIN.

Table 20 Query for STEWARD materialized views-TRAFFIC_COUNTS_5MIN

```
SELECT
      DATE_TBL.DAY,
      TIME_TBL.TIME,
      TO_NUMBER(SUBSTR(TIME_TBL.TIME, 1, 2)),     /* Hour */
      FLOOR(STATION.STATION_ID/100000),       /* District */
      FLOOR(STATION.STATION_ID/10000)- 10*FLOOR(STATION.STATION_ID/100000),
      STATION.STATION_ID,
      STATION.NUM_OF_LANES,

      CASE     /*District 4 Facility 2(I-595) and District 6 Facility 4(I-195) has E-W direction */
          WHEN ((FLOOR(STATION.STATION_ID/10000) =42) or (FLOOR(STATION.STATION_ID/10000) =64)) THEN
              CASE
                  WHEN MOD(STATION.STATION_ID,10) = 1 THEN 'E'
                  ELSE 'W'
                  END
          ELSE
              CASE
                  WHEN MOD(STATION.STATION_ID,10) = 1 THEN 'N'
                  ELSE 'S'
                  END
      END,

      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      (COALESCE(TSS_5MIN_LANE.LANE1_VOL,0) + COALESCE(TSS_5MIN_LANE.LANE2_VOL,0) +
                  COALESCE(TSS_5MIN_LANE.LANE3_VOL,0) + COALESCE(TSS_5MIN_LANE.LANE4_VOL,0) +
                  COALESCE(TSS_5MIN_LANE.LANE5_VOL,0) + COALESCE(TSS_5MIN_LANE.LANE6_VOL,0)) AS
                  TOTAL,

      TSS_5MIN_LANE.LANE1_VOL,
      TSS_5MIN_LANE.LANE2_VOL,
      TSS_5MIN_LANE.LANE3_VOL,
      TSS_5MIN_LANE.LANE4_VOL,
      TSS_5MIN_LANE.LANE5_VOL,
      TSS_5MIN_LANE.LANE6_VOL,
      CASE
          WHEN TSS_5MIN_LANE.LANE1_VOL IS NULL THEN NULL

          WHEN (TSS_5MIN_LANE.LANE2_VOL IS NULL) AND (TSS_5MIN_LANE.LANE1_VOL > 0) THEN 1

          WHEN (TSS_5MIN_LANE.LANE3_VOL IS NULL) AND (LEAST(TSS_5MIN_LANE.LANE1_VOL,
                  TSS_5MIN_LANE.LANE2_VOL) >0 ) THEN GREATEST (TSS_5MIN_LANE.LANE1_VOL,
                  TSS_5MIN_LANE.LANE2_VOL)/ LEAST(TSS_5MIN_LANE.LANE1_VOL, TSS_5MIN_LANE.LANE2_VOL)

          WHEN (TSS_5MIN_LANE.LANE4_VOL IS NULL) AND (LEAST(TSS_5MIN_LANE.LANE1_VOL,
                  TSS_5MIN_LANE.LANE2_VOL, TSS_5MIN_LANE.LANE3_VOL) >0 ) THEN GREATEST
                  (TSS_5MIN_LANE.LANE1_VOL, TSS_5MIN_LANE.LANE2_VOL, TSS_5MIN_LANE.LANE3_VOL)/
                  LEAST(TSS_5MIN_LANE.LANE1_VOL, TSS_5MIN_LANE.LANE2_VOL, TSS_5MIN_LANE.LANE3_VOL)

          WHEN (TSS_5MIN_LANE.LANE5_VOL IS NULL) AND (LEAST(TSS_5MIN_LANE.LANE1_VOL,
                  TSS_5MIN_LANE.LANE2_VOL, TSS_5MIN_LANE.LANE3_VOL, TSS_5MIN_LANE.LANE4_VOL) >0 )
                  THEN GREATEST (TSS_5MIN_LANE.LANE1_VOL, TSS_5MIN_LANE.LANE2_VOL,
                  TSS_5MIN_LANE.LANE3_VOL, TSS_5MIN_LANE.LANE4_VOL)/ LEAST(TSS_5MIN_LANE.LANE1_VOL,
                  TSS_5MIN_LANE.LANE2_VOL, TSS_5MIN_LANE.LANE3_VOL, TSS_5MIN_LANE.LANE4_VOL)

          WHEN (TSS_5MIN_LANE.LANE6_VOL IS NULL) AND (LEAST(TSS_5MIN_LANE.LANE1_VOL,
                  TSS_5MIN_LANE.LANE2_VOL, TSS_5MIN_LANE.LANE3_VOL, TSS_5MIN_LANE.LANE4_VOL,
                  TSS_5MIN_LANE.LANE5_VOL) >0 ) THEN GREATEST (TSS_5MIN_LANE.LANE1_VOL,
                  TSS_5MIN_LANE.LANE2_VOL, TSS_5MIN_LANE.LANE3_VOL, TSS_5MIN_LANE.LANE4_VOL,
                  TSS_5MIN_LANE.LANE5_VOL)/ LEAST(TSS_5MIN_LANE.LANE1_VOL, TSS_5MIN_LANE.LANE2_VOL,
                  TSS_5MIN_LANE.LANE3_VOL, TSS_5MIN_LANE.LANE4_VOL, TSS_5MIN_LANE.LANE5_VOL)
```

Table 20 (Continued)

```
          WHEN LEAST(TSS_5MIN_LANE.LANE1_VOL, TSS_5MIN_LANE.LANE2_VOL, TSS_5MIN_LANE.LANE3_VOL,
                   TSS_5MIN_LANE.LANE4_VOL, TSS_5MIN_LANE.LANE5_VOL, TSS_5MIN_LANE.LANE6_VOL) >0
                   THEN GREATEST (TSS_5MIN_LANE.LANE1_VOL, TSS_5MIN_LANE.LANE2_VOL,
                   TSS_5MIN_LANE.LANE3_VOL, TSS_5MIN_LANE.LANE4_VOL, TSS_5MIN_LANE.LANE5_VOL,
                   TSS_5MIN_LANE.LANE6_VOL)/ LEAST(TSS_5MIN_LANE.LANE1_VOL, TSS_5MIN_LANE.LANE2_VOL,
                   TSS_5MIN_LANE.LANE3_VOL, TSS_5MIN_LANE.LANE4_VOL, TSS_5MIN_LANE.LANE5_VOL,
                   TSS_5MIN_LANE.LANE6_VOL)

          ELSE NULL
      END,

      CASE
          WHEN TSS_5MIN_LANE.LANE1_OBS IS NULL THEN NULL
          WHEN TSS_5MIN_LANE.LANE1_OBS >15 THEN 100
          ELSE ROUND(TSS_5MIN_LANE.LANE1_OBS/15*100,2)
      END,

      TSS_5MIN_LANE.ONRAMP1_VOL,
      TSS_5MIN_LANE.ONRAMP2_VOL,
      TSS_5MIN_LANE.ONRAMP3_VOL,
      0,

      TSS_5MIN_LANE.OFFRAMP1_VOL,
      TSS_5MIN_LANE.OFFRAMP2_VOL,
      TSS_5MIN_LANE.OFFRAMP3_VOL,
      0,

     STATION.COUNTY,
     STATION.COUNT_STATION

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE,
     GATOR.DATE_REF DATE_TBL, GATOR.TIME_REF TIME_TBL

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
```

3.4.5 MAX_FLOW

These tables provide max flow for web user queries. Table 21 shows the materialized views definitions for MAX_FLOW_5MIN, MAX_FLOW_15MIN, and MAX_FLOW_1HR.

Table 21 STEWARD materialized views – definitions of MAX_FLOW_5MIN, MAX_FLOW_15MIN, and MAX_FLOW_1HR

Name	Data Type (Length/Precision)
DATE1	DATE
DISTRICT	NUMBER
FACILITY	NUMBER
STATION_ID	NUMBER
STATION_DESC	VARCHAR2(240)
STATION_MP	NUMBER
LANE_NUM	NUMBER
MAX_FLOW	NUMBER
MAX_TIME	VARCHAR2(10)
MAX_GAP	NUMBER

Table 22 shows the query to generate the materialized view- MAX_FLOW_5MIN.

Table 22 Query for STEWARD materialized views-MAX_FLOW_5MIN

```
SELECT     /*Lane1*/
      DATE_TBL.DAY,
      FLOOR(STATION.STATION_ID/100000),       /* District */
      FLOOR(STATION.STATION_ID/10000)- 10*FLOOR(STATION.STATION_ID/100000),
      STATION.STATION_ID,
      MOD(STATION.STATION_ID,10),
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      1,
      12*TSS_5MIN_LANE.LANE1_VOL,
      TIME_TBL.TIME,
      0

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE,
     GATOR.DATE_REF DATE_TBL, GATOR.TIME_REF TIME_TBL,
     (SELECT MAX(TSS_5MIN_LANE.LANE1_VOL) as MAX_VOL, TSS_STATION, DATE_REF
      FROM GATOR.TSS_5MIN_LANE GROUP BY TSS_STATION, DATE_REF) MAX_FLOW

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.LANE1_VOL = MAX_FLOW.MAX_VOL
            AND TSS_5MIN_LANE.TSS_STATION = MAX_FLOW.TSS_STATION
            AND TSS_5MIN_LANE.DATE_REF = MAX_FLOW.DATE_REF
            AND MAX_FLOW.MAX_VOL IS NOT NULL
            AND MAX_FLOW.MAX_VOL >0
UNION

SELECT     /*Lane2*/
      DATE_TBL.DAY,
      FLOOR(STATION.STATION_ID/100000),       /* District */
      FLOOR(STATION.STATION_ID/10000)- 10*FLOOR(STATION.STATION_ID/100000),
      STATION.STATION_ID,
      MOD(STATION.STATION_ID,10),
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      2,
      12*TSS_5MIN_LANE.LANE2_VOL,
      TIME_TBL.TIME,
      0

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE,
     GATOR.DATE_REF DATE_TBL, GATOR.TIME_REF TIME_TBL,
     (SELECT MAX(TSS_5MIN_LANE.LANE2_VOL) as MAX_VOL, TSS_STATION, DATE_REF
      FROM GATOR.TSS_5MIN_LANE GROUP BY TSS_STATION, DATE_REF) MAX_FLOW

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.LANE2_VOL = MAX_FLOW.MAX_VOL
            AND TSS_5MIN_LANE.TSS_STATION = MAX_FLOW.TSS_STATION
            AND TSS_5MIN_LANE.DATE_REF = MAX_FLOW.DATE_REF
            AND MAX_FLOW.MAX_VOL IS NOT NULL
            AND MAX_FLOW.MAX_VOL >0
UNION

SELECT     /*Lane3*/
      DATE_TBL.DAY,
      FLOOR(STATION.STATION_ID/100000),       /* District */
      FLOOR(STATION.STATION_ID/10000)- 10*FLOOR(STATION.STATION_ID/100000),
      STATION.STATION_ID,
      MOD(STATION.STATION_ID,10),
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,
```

Table 22 (continued)

```
      3,
      12*TSS_5MIN_LANE.LANE3_VOL,
      TIME_TBL.TIME,
      0

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE,
     GATOR.DATE_REF DATE_TBL, GATOR.TIME_REF TIME_TBL,
     (SELECT MAX(TSS_5MIN_LANE.LANE3_VOL) as MAX_VOL, TSS_STATION, DATE_REF
     FROM GATOR.TSS_5MIN_LANE GROUP BY TSS_STATION, DATE_REF) MAX_FLOW

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.LANE3_VOL = MAX_FLOW.MAX_VOL
            AND TSS_5MIN_LANE.TSS_STATION = MAX_FLOW.TSS_STATION
            AND TSS_5MIN_LANE.DATE_REF = MAX_FLOW.DATE_REF
            AND MAX_FLOW.MAX_VOL IS NOT NULL
            AND MAX_FLOW.MAX_VOL >0
UNION

SELECT     /*LANE4*/
      DATE_TBL.DAY,
      FLOOR(STATION.STATION_ID/100000),       /* District */
      FLOOR(STATION.STATION_ID/10000)- 10*FLOOR(STATION.STATION_ID/100000),
      STATION.STATION_ID,
      MOD(STATION.STATION_ID,10),
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      4,
      12*TSS_5MIN_LANE.LANE4_VOL,
      TIME_TBL.TIME,
      0

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE,
     GATOR.DATE_REF DATE_TBL, GATOR.TIME_REF TIME_TBL,
     (SELECT MAX(TSS_5MIN_LANE.LANE4_VOL) as MAX_VOL, TSS_STATION, DATE_REF
      FROM GATOR.TSS_5MIN_LANE GROUP BY TSS_STATION, DATE_REF) MAX_FLOW

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.LANE4_VOL = MAX_FLOW.MAX_VOL
            AND TSS_5MIN_LANE.TSS_STATION = MAX_FLOW.TSS_STATION
            AND TSS_5MIN_LANE.DATE_REF = MAX_FLOW.DATE_REF
            AND MAX_FLOW.MAX_VOL IS NOT NULL
            AND MAX_FLOW.MAX_VOL >0
UNION

SELECT     /*Lane5*/
      DATE_TBL.DAY,
      FLOOR(STATION.STATION_ID/100000),       /* District */
      FLOOR(STATION.STATION_ID/10000)- 10*FLOOR(STATION.STATION_ID/100000),
      STATION.STATION_ID,
      MOD(STATION.STATION_ID,10),
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      5,
      12*TSS_5MIN_LANE.LANE5_VOL,
      TIME_TBL.TIME,
      0

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE,
     GATOR.DATE_REF DATE_TBL, GATOR.TIME_REF TIME_TBL,
     (SELECT MAX(TSS_5MIN_LANE.LANE5_VOL) as MAX_VOL, TSS_STATION, DATE_REF
      FROM GATOR.TSS_5MIN_LANE GROUP BY TSS_STATION, DATE_REF) MAX_FLOW
```

Table 22 (continued)

```
WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.LANE5_VOL = MAX_FLOW.MAX_VOL
            AND TSS_5MIN_LANE.TSS_STATION = MAX_FLOW.TSS_STATION
            AND TSS_5MIN_LANE.DATE_REF = MAX_FLOW.DATE_REF
            AND MAX_FLOW.MAX_VOL IS NOT NULL
            AND MAX_FLOW.MAX_VOL >0
UNION

SELECT     /*Lane6*/
      DATE_TBL.DAY,
      FLOOR(STATION.STATION_ID/100000),       /* District */
      FLOOR(STATION.STATION_ID/10000)- 10*FLOOR(STATION.STATION_ID/100000),
      STATION.STATION_ID,
      MOD(STATION.STATION_ID,10),
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      6,
      12*TSS_5MIN_LANE.LANE6_VOL,
      TIME_TBL.TIME,
      0

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE,
     GATOR.DATE_REF DATE_TBL, GATOR.TIME_REF TIME_TBL,
    (SELECT MAX(TSS_5MIN_LANE.LANE6_VOL) as MAX_VOL, TSS_STATION, DATE_REF
     FROM GATOR.TSS_5MIN_LANE GROUP BY TSS_STATION, DATE_REF) MAX_FLOW

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.LANE6_VOL = MAX_FLOW.MAX_VOL
            AND TSS_5MIN_LANE.TSS_STATION = MAX_FLOW.TSS_STATION
            AND TSS_5MIN_LANE.DATE_REF = MAX_FLOW.DATE_REF
            AND MAX_FLOW.MAX_VOL IS NOT NULL
            AND MAX_FLOW.MAX_VOL >0
```

3.4.6 EFF_VEH_LEN

These tables provide effective vehicle length for web user queries. Table 23 shows the materialized views definitions for EFF_VEH_LEN_5MIN, EFF_VEH_LEN_15MIN, and EFF_VEH_LEN_1HR.

Table 23 STEWARD materialized views – definitions of EFF_VEH_LEN_5MIN, EFF_VEH_LEN_15MIN, and EFF_VEH_LEN_1HR

Name	Data Type (Length/Precision)
DATE1	DATE
TIME	VARCHAR2(30)
HOUR	NUMBER
DISTRICT	NUMBER
FACILITY	NUMBER
STATION_ID	NUMBER
DIRECTION	NUMBER
STATION_DESC	VARCHAR2(240)
STATION_MP	NUMBER
LANE1_VOL	NUMBER
LANE1_SPD	NUMBER
LANE1_OCC	NUMBER
LANE1_EFF_VEH_LENGTH	NUMBER
LANE2_VOL	NUMBER
LANE2_SPD	NUMBER
LANE2_OCC	NUMBER
LANE2_EFF_VEH_LENGTH	NUMBER
LANE3_VOL	NUMBER
LANE3_SPD	NUMBER
LANE3_OCC	NUMBER
LANE3_EFF_VEH_LENGTH	NUMBER
LANE4_VOL	NUMBER
LANE4_SPD	NUMBER
LANE4_OCC	NUMBER
LANE4_EFF_VEH_LENGTH	NUMBER
LANE5_VOL	NUMBER
LANE5_SPD	NUMBER
LANE5_OCC	NUMBER
LANE5_EFF_VEH_LENGTH	NUMBER
LANE6_VOL	NUMBER
LANE6_SPD	NUMBER
LANE6_OCC	NUMBER
LANE6_EFF_VEH_LENGTH	NUMBER

Table 24 shows the query to generate the materialized view- EFF_VEH_LEN_5MIN.

Table 24 Query for STEWARD materialized views-EFF_VEH_LEN_5MIN

```
SELECT
      DATE_TBL.DAY,
      TIME_TBL.TIME,
      TO_NUMBER(SUBSTR(TIME_TBL.TIME, 1, 2)),     /* Hour */
      FLOOR(STATION.STATION_ID/100000),       /* District */
      FLOOR(STATION.STATION_ID/10000)- 10*FLOOR(STATION.STATION_ID/100000),
      STATION.STATION_ID,
      MOD(STATION.STATION_ID,10),
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      TSS_5MIN_LANE.LANE1_VOL,
      TSS_5MIN_LANE.LANE1_SPD,
      TSS_5MIN_LANE.LANE1_OCC,
      /* vehicle length = velocity*5280*occupancy/100/flow */
      CASE
            WHEN TSS_5MIN_LANE.LANE1_VOL > 0 THEN (TSS_5MIN_LANE.LANE1_SPD * 5280)
                 * (TSS_5MIN_LANE.LANE1_OCC /100) / (12*TSS_5MIN_LANE.LANE1_VOL)
            WHEN TSS_5MIN_LANE.LANE1_VOL = 0 THEN 0
            ELSE NULL
      END,

      TSS_5MIN_LANE.LANE2_VOL,
      TSS_5MIN_LANE.LANE2_SPD,
      TSS_5MIN_LANE.LANE2_OCC,
      /* vehicle length = velocity*5280*occupancy/100/flow */
      CASE
            WHEN TSS_5MIN_LANE.LANE2_VOL > 0 THEN (TSS_5MIN_LANE.LANE2_SPD * 5280)
                 * (TSS_5MIN_LANE.LANE2_OCC /100) / (12*TSS_5MIN_LANE.LANE2_VOL)
            WHEN TSS_5MIN_LANE.LANE1_VOL = 0 THEN 0
            ELSE NULL
      END,

      TSS_5MIN_LANE.LANE3_VOL,
      TSS_5MIN_LANE.LANE3_SPD,
      TSS_5MIN_LANE.LANE3_OCC,
      /* vehicle length = velocity*5280*occupancy/100/flow */
      CASE
            WHEN TSS_5MIN_LANE.LANE3_VOL > 0 THEN (TSS_5MIN_LANE.LANE3_SPD * 5280)
                 * (TSS_5MIN_LANE.LANE3_OCC /100) / (12*TSS_5MIN_LANE.LANE3_VOL)
            WHEN TSS_5MIN_LANE.LANE1_VOL = 0 THEN 0
            ELSE NULL
      END,

      TSS_5MIN_LANE.LANE4_VOL,
      TSS_5MIN_LANE.LANE4_SPD,
      TSS_5MIN_LANE.LANE4_OCC,
      /* vehicle length = velocity*5280*occupancy/100/flow */
      CASE
            WHEN TSS_5MIN_LANE.LANE4_VOL > 0 THEN (TSS_5MIN_LANE.LANE4_SPD * 5280)
                 * (TSS_5MIN_LANE.LANE4_OCC /100) / (12*TSS_5MIN_LANE.LANE4_VOL)
            WHEN TSS_5MIN_LANE.LANE1_VOL = 0 THEN 0
            ELSE NULL
      END,

      TSS_5MIN_LANE.LANE5_VOL,
      TSS_5MIN_LANE.LANE5_SPD,
      TSS_5MIN_LANE.LANE5_OCC,
      /* vehicle length = velocity*5280*occupancy/100/flow */
      CASE
            WHEN TSS_5MIN_LANE.LANE5_VOL > 0 THEN (TSS_5MIN_LANE.LANE5_SPD * 5280)
                 * (TSS_5MIN_LANE.LANE5_OCC /100) / (12*TSS_5MIN_LANE.LANE5_VOL)
            WHEN TSS_5MIN_LANE.LANE1_VOL = 0 THEN 0
            ELSE NULL
      END,
```

Table 24 (Continued)

```
      TSS_5MIN_LANE.LANE6_VOL,
      TSS_5MIN_LANE.LANE6_SPD,
      TSS_5MIN_LANE.LANE6_OCC,
      /* vehicle length = velocity*5280*occupancy/100/flow */
      CASE
            WHEN TSS_5MIN_LANE.LANE6_VOL > 0 THEN (TSS_5MIN_LANE.LANE6_SPD * 5280)
                 * (TSS_5MIN_LANE.LANE6_OCC /100) / (12*TSS_5MIN_LANE.LANE6_VOL)
            WHEN TSS_5MIN_LANE.LANE1_VOL = 0 THEN 0
            ELSE NULL
      END

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE,
     GATOR.DATE_REF DATE_TBL, GATOR.TIME_REF TIME_TBL

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
```

3.5 Functions

When STEWARD Phase 3 started, there was a new requirement for a *full period* aggregation level. In this level, the report is created as a summary of the day for the specific daily time range. For example, if the user selects the traffic count report with daily time range from 7:00AM to 10:00AM and with the full period aggregation level for a week, STEWARD would generate the traffic count report with the aggregation level of three hours (7:00AM to 10:00AM) per each day on the specific facility.

Because of this flexibility, full period report cannot be generated in advance like the materialized views. STEWARD has to aggregate the results on every user's request. Two main principles were adopted to implement these requirements:

- Process as much data as possible on server side. This will minimize the need for large data transfer from the server to client side.
- Make the output of this process as a table. This will give the same interface to the web development.

Both principles were established to give the STEWARD web developers and users a consistent STEWARD interface. For these purposes, Oracle functions are used to provide the full period functions. These functions read the internal table and process the data internally at the server side. Also, they return the Oracle table to the function caller in the same format as the materialized views.

Table 12, presented preciously, shows the STEWARD web reports and their corresponding functions. There are 5 functions for web reports, Q_ALLDATAFIELDS, Q_EFFVEHLEN, Q_TRAFFICCOUNTS, Q_TSS, and Q_VOLUMEMAP. These functions will now be described.

3.5.1 Function Q_ALLDATAFIELDS

This function receives the input parameters of start date, start time, end date, end time, and district, and creates the return variables as a table of daily average values for all data fields. Table 25 shows the return table for Q_AllDataFields.

Table 25 STEWARD function – definitions of Q_AllDataFields

Name	Data Type (Length/Precision)
DATE1	DATE
DISTRICT	VARCHAR2(7)
FACILITY	NUMBER
STATION_ID	NUMBER
DIRECTION	NUMBER
STATION_DESC	VARCHAR2(240)
STATION_MP	NUMBER
LANE1_VOL	NUMBER
LANE1_SPD	NUMBER
LANE1_OCC	NUMBER
LANE2_VOL	NUMBER
LANE2_SPD	NUMBER
LANE2_OCC	NUMBER
LANE3_VOL	NUMBER
LANE3_SPD	NUMBER
LANE3_OCC	NUMBER
LANE4_VOL	NUMBER
LANE4_SPD	NUMBER
LANE4_OCC	NUMBER
LANE5_VOL	NUMBER
LANE5_SPD	NUMBER
LANE5_OCC	NUMBER
LANE6_VOL	NUMBER
LANE6_SPD	NUMBER
LANE6_OCC	NUMBER
ONRAMP1_VOL	NUMBER
ONRAMP1_SPD	NUMBER
ONRAMP1_OCC	NUMBER
ONRAMP2_VOL	NUMBER
ONRAMP2_SPD	NUMBER
ONRAMP2_OCC	NUMBER
ONRAMP3_VOL	NUMBER
ONRAMP3_SPD	NUMBER
ONRAMP3_OCC	NUMBER
OFFRAMP1_VOL	NUMBER
OFFRAMP1_SPD	NUMBER
OFFRAMP1_OCC	NUMBER
OFFRAMP2_VOL	NUMBER
OFFRAMP2_SPD	NUMBER
OFFRAMP2_OCC	NUMBER
OFFRAMP3_VOL	NUMBER
OFFRAMP3_SPD	NUMBER
OFFRAMP3_OCC	NUMBER

Table 26 shows the function Q_AllDataFields.

Table 26 STEWARD function - Q_AllDataFields

```
--initialize variables here
      CURSOR cur (c_startdate1 date, c_enddate1 date, c_starttime number, c_endtime
number, c_district number)
      IS
         SELECT
            DATE1,
            DISTRICT,
            FACILITY,
            STATION_ID,
            DIRECTION,
            STATION_DESC,
            STATION_MP,

            SUM(LANE1_VOL) AS SUM_OF_LANE1_VOL,
            (CASE
            WHEN SUM(LANE1_VOL)> 0 THEN SUM(LANE1_VOL*LANE1_SPD)/ SUM(LANE1_VOL)
            WHEN COUNT(LANE1_VOL)= 0 THEN NULL
            ELSE 0
            END) AVG_OF_LANE1_SPD,
            AVG(LANE1_OCC) AS AVG_OF_LANE1_OCC,

            SUM(LANE2_VOL) AS SUM_OF_LANE2_VOL,
            (CASE
                  WHEN SUM(LANE2_VOL) > 0 THEN SUM(LANE2_SPD*LANE2_VOL)/SUM(LANE2_VOL)
                  WHEN SUM(LANE2_VOL) = 0 THEN 0
                  ELSE NULL
            END) AVG_OF_LANE2_SPD,
            AVG(LANE2_OCC) AS AVG_OF_LANE2_OCC,

            SUM(LANE3_VOL) AS SUM_OF_LANE3_VOL,
            (CASE
                  WHEN SUM(LANE3_VOL) > 0 THEN SUM(LANE3_SPD*LANE3_VOL)/SUM(LANE3_VOL)
                  WHEN SUM(LANE3_VOL) = 0 THEN 0
                  ELSE NULL
            END) AVG_OF_LANE3_SPD,
            AVG(LANE3_OCC) AS AVG_OF_LANE3_OCC,

            SUM(LANE4_VOL) AS SUM_OF_LANE4_VOL,
            (CASE
                  WHEN SUM(LANE4_VOL) > 0 THEN SUM(LANE4_SPD*LANE4_VOL)/SUM(LANE4_VOL)
                  WHEN SUM(LANE4_VOL) = 0 THEN 0
                  ELSE NULL
            END) AVG_OF_LANE4_SPD,
            AVG(LANE4_OCC) AS AVG_OF_LANE4_OCC,

            SUM(LANE5_VOL) AS SUM_OF_LANE5_VOL,
            (CASE
                  WHEN SUM(LANE5_VOL) > 0 THEN SUM(LANE5_SPD*LANE5_VOL)/SUM(LANE5_VOL)
                  WHEN SUM(LANE5_VOL) = 0 THEN 0
                  ELSE NULL
            END) AVG_OF_LANE5_SPD,
            AVG(LANE5_OCC) AS AVG_OF_LANE5_OCC,

            SUM(LANE6_VOL) AS SUM_OF_LANE6_VOL,
            (CASE
                  WHEN SUM(LANE6_VOL) > 0 THEN SUM(LANE6_SPD*LANE6_VOL)/SUM(LANE6_VOL)
                  WHEN SUM(LANE6_VOL) = 0 THEN 0
                  ELSE NULL
            END) AVG_OF_LANE6_SPD,
            AVG(LANE6_OCC) AS AVG_OF_LANE6_OCC,

            SUM(ONRAMP1_VOL) AS SUM_OF_ONRAMP1_VOL,
            (CASE
            WHEN SUM(ONRAMP1_VOL)> 0 THEN SUM(ONRAMP1_VOL*ONRAMP1_SPD)/ SUM(ONRAMP1_VOL)
            WHEN COUNT(ONRAMP1_VOL)= 0 THEN NULL
            ELSE 0
            END) AVG_OF_ONRAMP1_SPD,
            AVG(ONRAMP1_OCC) AS AVG_OF_ONRAMP1_OCC,

            SUM(ONRAMP2_VOL) AS SUM_OF_ONRAMP2_VOL,
```

Table 26 (Continued)

```
            (CASE
            WHEN SUM(ONRAMP2_VOL)> 0 THEN SUM(ONRAMP2_VOL*ONRAMP2_SPD)/ SUM(ONRAMP2_VOL)
            WHEN COUNT(ONRAMP2_VOL)= 0 THEN NULL
            ELSE 0
            END) AVG_OF_ONRAMP2_SPD,
            AVG(ONRAMP2_OCC) AS AVG_OF_ONRAMP2_OCC,

            SUM(ONRAMP3_VOL) AS SUM_OF_ONRAMP3_VOL,
            (CASE
            WHEN SUM(ONRAMP3_VOL)> 0 THEN SUM(ONRAMP3_VOL*ONRAMP3_SPD)/ SUM(ONRAMP3_VOL)
            WHEN COUNT(ONRAMP3_VOL)= 0 THEN NULL
            ELSE 0
            END) AVG_OF_ONRAMP3_SPD,
            AVG(ONRAMP3_OCC) AS AVG_OF_ONRAMP3_OCC,

            SUM(OFFRAMP1_VOL) AS SUM_OF_OFFRAMP1_VOL,
            (CASE
            WHEN SUM(OFFRAMP1_VOL)> 0 THEN
                 SUM(OFFRAMP1_VOL*OFFRAMP1_SPD)/ SUM(OFFRAMP1_VOL)
            WHEN COUNT(OFFRAMP1_VOL)= 0 THEN NULL
            ELSE 0
            END) AVG_OF_OFFRAMP1_SPD,
            AVG(OFFRAMP1_OCC) AS AVG_OF_OFFRAMP1_OCC,

            SUM(OFFRAMP2_VOL) AS SUM_OF_OFFRAMP2_VOL,
            (CASE
            WHEN SUM(OFFRAMP2_VOL)> 0 THEN
                 SUM(OFFRAMP2_VOL*OFFRAMP2_SPD)/ SUM(OFFRAMP2_VOL)
            WHEN COUNT(OFFRAMP2_VOL)= 0 THEN NULL
            ELSE 0
            END) AVG_OF_OFFRAMP2_SPD,
            AVG(OFFRAMP2_OCC) AS AVG_OF_OFFRAMP2_OCC,

            SUM(OFFRAMP3_VOL) AS SUM_OF_OFFRAMP3_VOL,
            (CASE
            WHEN SUM(OFFRAMP3_VOL)> 0 THEN
                 SUM(OFFRAMP3_VOL*OFFRAMP3_SPD)/ SUM(OFFRAMP3_VOL)
            WHEN COUNT(OFFRAMP3_VOL)= 0 THEN NULL
            ELSE 0
            END) AVG_OF_OFFRAMP3_SPD,
            AVG(OFFRAMP3_OCC) AS AVG_OF_OFFRAMP3_OCC

         FROM SYSTEM_LANE_1HR
         WHERE date1>=c_startdate1 AND date1<=c_enddate1
               AND TO_NUMBER(SUBSTR(TIME, 1, 2))>= c_starttime
               AND TO_NUMBER(SUBSTR(TIME, 1, 2))<= c_endtime
               AND DISTRICT = c_district
         GROUP BY DATE1, DISTRICT, FACILITY, STATION_ID, DIRECTION,
                  STATION_DESC, STATION_MP;

         ty             AllDataFields_fullperiod_ty;
         tbl            AllDataFields_fullperiod := AllDataFields_fullperiod ();
-- main body
BEGIN
      FOR rec IN cur (p_startdate, p_enddate, p_starttime, p_endtime, p_district)
      LOOP
         ty := Alldatafields_fullperiod_ty (rec.DATE1,
                                            rec.DISTRICT,
                                            rec.FACILITY,
                                            rec.STATION_ID,
                                            rec.DIRECTION,
                                            rec.STATION_DESC,
                                            rec.STATION_MP,
                                            rec.SUM_OF_LANE1_VOL,
                                            rec.AVG_OF_LANE1_SPD,
                                            rec.AVG_OF_LANE1_OCC,
                                            rec.SUM_OF_LANE2_VOL,
                                            rec.AVG_OF_LANE2_SPD,
                                            rec.AVG_OF_LANE2_OCC,
                                            rec.SUM_OF_LANE3_VOL,
```

Table 26 (Continued)

```
                                        rec.AVG_OF_LANE3_SPD,
                                        rec.AVG_OF_LANE3_OCC,
                                        rec.SUM_OF_LANE4_VOL,
                                        rec.AVG_OF_LANE4_SPD,
                                        rec.AVG_OF_LANE4_OCC,
                                        rec.SUM_OF_LANE5_VOL,
                                        rec.AVG_OF_LANE5_SPD,
                                        rec.AVG_OF_LANE5_OCC,
                                        rec.SUM_OF_LANE6_VOL,
                                        rec.AVG_OF_LANE6_SPD,
                                        rec.AVG_OF_LANE6_OCC,

                                        rec.SUM_OF_ONRAMP1_VOL,
                                        rec.AVG_OF_ONRAMP1_SPD,
                                        rec.AVG_OF_ONRAMP1_OCC,
                                        rec.SUM_OF_ONRAMP2_VOL,
                                        rec.AVG_OF_ONRAMP2_SPD,
                                        rec.AVG_OF_ONRAMP2_OCC,
                                        rec.SUM_OF_ONRAMP3_VOL,
                                        rec.AVG_OF_ONRAMP3_SPD,
                                        rec.AVG_OF_ONRAMP3_OCC,

                                        rec.SUM_OF_OFFRAMP1_VOL,
                                        rec.AVG_OF_OFFRAMP1_SPD,
                                        rec.AVG_OF_OFFRAMP1_OCC,
                                        rec.SUM_OF_OFFRAMP2_VOL,
                                        rec.AVG_OF_OFFRAMP2_SPD,
                                        rec.AVG_OF_OFFRAMP2_OCC,
                                        rec.SUM_OF_OFFRAMP3_VOL,
                                        rec.AVG_OF_OFFRAMP3_SPD,
                                        rec.AVG_OF_OFFRAMP3_OCC);
         tbl.EXTEND;
         tbl (tbl.LAST) := ty;
      END LOOP;
      RETURN tbl;
   END;
FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE, GATOR.DATE_REF
DATE_TBL, GATOR.TIME_REF TIME_TBL,
   (SELECT MAX(TSS_5MIN_LANE.LANE3_VOL) as MAX_VOL, TSS_STATION, DATE_REF
    FROM GATOR.TSS_5MIN_LANE GROUP BY TSS_STATION, DATE_REF) MAX_FLOW

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.LANE3_VOL = MAX_FLOW.MAX_VOL
            AND TSS_5MIN_LANE.TSS_STATION = MAX_FLOW.TSS_STATION
            AND TSS_5MIN_LANE.DATE_REF = MAX_FLOW.DATE_REF
            AND MAX_FLOW.MAX_VOL IS NOT NULL
            AND MAX_FLOW.MAX_VOL >0
UNION

SELECT     /*LANE4*/
      DATE_TBL.DAY,
      FLOOR(STATION.STATION_ID/100000),       /* District */
      FLOOR(STATION.STATION_ID/10000)- 10*FLOOR(STATION.STATION_ID/100000),
      STATION.STATION_ID,
      MOD(STATION.STATION_ID,10),
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      4,
      12*TSS_5MIN_LANE.LANE4_VOL,
      TIME_TBL.TIME,
      0



FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE, GATOR.DATE_REF
DATE_TBL, GATOR.TIME_REF TIME_TBL,
   (SELECT MAX(TSS_5MIN_LANE.LANE4_VOL) as MAX_VOL, TSS_STATION, DATE_REF
```

Table 26 (Continued)

```
FROM GATOR.TSS_5MIN_LANE GROUP BY TSS_STATION, DATE_REF) MAX_FLOW

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.LANE4_VOL = MAX_FLOW.MAX_VOL
            AND TSS_5MIN_LANE.TSS_STATION = MAX_FLOW.TSS_STATION
            AND TSS_5MIN_LANE.DATE_REF = MAX_FLOW.DATE_REF
            AND MAX_FLOW.MAX_VOL IS NOT NULL
            AND MAX_FLOW.MAX_VOL >0
UNION

SELECT     /*Lane5*/
      DATE_TBL.DAY,
      FLOOR(STATION.STATION_ID/100000),       /* District */
      FLOOR(STATION.STATION_ID/10000)- 10*FLOOR(STATION.STATION_ID/100000),
      STATION.STATION_ID,
      MOD(STATION.STATION_ID,10),
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      5,
      12*TSS_5MIN_LANE.LANE5_VOL,
      TIME_TBL.TIME,
      0

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE, GATOR.DATE_REF
DATE_TBL, GATOR.TIME_REF TIME_TBL,
   (SELECT MAX(TSS_5MIN_LANE.LANE5_VOL) as MAX_VOL, TSS_STATION, DATE_REF
FROM GATOR.TSS_5MIN_LANE GROUP BY TSS_STATION, DATE_REF) MAX_FLOW

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.LANE5_VOL = MAX_FLOW.MAX_VOL
            AND TSS_5MIN_LANE.TSS_STATION = MAX_FLOW.TSS_STATION
            AND TSS_5MIN_LANE.DATE_REF = MAX_FLOW.DATE_REF
            AND MAX_FLOW.MAX_VOL IS NOT NULL
            AND MAX_FLOW.MAX_VOL >0
UNION

SELECT     /*Lane6*/
      DATE_TBL.DAY,
      FLOOR(STATION.STATION_ID/100000),       /* District */
      FLOOR(STATION.STATION_ID/10000)- 10*FLOOR(STATION.STATION_ID/100000),
      STATION.STATION_ID,
      MOD(STATION.STATION_ID,10),
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      6,
      12*TSS_5MIN_LANE.LANE6_VOL,
      TIME_TBL.TIME,
      0

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE, GATOR.DATE_REF
DATE_TBL, GATOR.TIME_REF TIME_TBL,
   (SELECT MAX(TSS_5MIN_LANE.LANE6_VOL) as MAX_VOL, TSS_STATION, DATE_REF
FROM GATOR.TSS_5MIN_LANE GROUP BY TSS_STATION, DATE_REF) MAX_FLOW

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.LANE6_VOL = MAX_FLOW.MAX_VOL
            AND TSS_5MIN_LANE.TSS_STATION = MAX_FLOW.TSS_STATION
            AND TSS_5MIN_LANE.DATE_REF = MAX_FLOW.DATE_REF
            AND MAX_FLOW.MAX_VOL IS NOT NULL
            AND MAX_FLOW.MAX_VOL >0
```

3.5.2 Function Q_EFFVEHLEN

This function receives the input parameters of start date, start time, end date, end time, and district, and creates the return variables as a table of daily average values for effective vehicle lengths. Table 27 shows the return table for STEWARD Q_EffVehLen.

Table 27 STEWARD function – definitions of Q_EffVehLen

Name	Data Type (Length/Precision)
DATE1	DATE
FACILITY	NUMBER
STATION_ID	NUMBER
DIRECTION	NUMBER
STATION_DESC	VARCHAR2(240)
STATION_MP	NUMBER
LANE1_VOL	NUMBER
LANE1_SPD	NUMBER
LANE1_OCC	NUMBER
LANE1_EFF_VEH_LENGTH	NUMBER
LANE2_VOL	NUMBER
LANE2_SPD	NUMBER
LANE2_OCC	NUMBER
LANE2_EFF_VEH_LENGTH	NUMBER
LANE3_VOL	NUMBER
LANE3_SPD	NUMBER
LANE3_OCC	NUMBER
LANE3_EFF_VEH_LENGTH	NUMBER
LANE4_VOL	NUMBER
LANE4_SPD	NUMBER
LANE4_OCC	NUMBER
LANE4_EFF_VEH_LENGTH	NUMBER
LANE5_VOL	NUMBER
LANE5_SPD	NUMBER
LANE5_OCC	NUMBER
LANE5_EFF_VEH_LENGTH	NUMBER
LANE6_VOL	NUMBER
LANE6_SPD	NUMBER
LANE6_OCC	NUMBER
LANE6_EFF_VEH_LENGTH	NUMBER

Table 28 shows the function Q_EffVehLen.

Table 28 STEWARD function - Q_EffVehLen

```
--initialize variables here
      CURSOR cur (c_startdate1 date, c_enddate1 date, c_starttime number, c_endtime
number, c_district number)
      IS
         SELECT
            DATE1,
            FACILITY,
            STATION_ID,
            DIRECTION,
            STATION_DESC,
            STATION_MP,

            SUM(LANE1_VOL) AS SUM_OF_LANE1_VOL,
            (CASE
                  WHEN SUM(LANE1_VOL) > 0 THEN SUM(LANE1_SPD*LANE1_VOL)/SUM(LANE1_VOL)
                  WHEN SUM(LANE1_VOL) = 0 THEN 0
                  ELSE NULL
            END) AVG_OF_LANE1_SPD,
            AVG(LANE1_OCC) AS AVG_OF_LANE1_OCC,
            /* vehicle length = velocity*5280*occupancy/100/flow */
            (CASE
                  WHEN SUM(LANE1_VOL) > 0 THEN
                      ((SUM(LANE1_SPD*LANE1_VOL)/SUM(LANE1_VOL)) * 5280)
                       * (AVG(LANE1_OCC) /100) /(SUM(LANE1_VOL)/24)
                  WHEN SUM(LANE1_VOL) = 0 THEN 0
                  ELSE NULL
            END) LANE1_EFFVEHLEN,

            SUM(LANE2_VOL) AS SUM_OF_LANE2_VOL,
            (CASE
                  WHEN SUM(LANE2_VOL) > 0 THEN
                       SUM(LANE2_SPD*LANE2_VOL)/SUM(LANE2_VOL)
                  WHEN SUM(LANE2_VOL) = 0 THEN 0
                  ELSE NULL
            END) AVG_OF_LANE2_SPD,
            AVG(LANE2_OCC) AS AVG_OF_LANE2_OCC,
            /* vehicle length = velocity*5280*occupancy/100/flow */
            (CASE
                  WHEN SUM(LANE2_VOL) > 0 THEN
                      (SUM(LANE2_SPD*LANE2_VOL)/SUM(LANE2_VOL) * 5280)
                       * (AVG(LANE2_OCC) /100) /(SUM(LANE2_VOL)/24)
                  WHEN SUM(LANE2_VOL) = 0 THEN 0
                  ELSE NULL
            END) LANE2_EFFDETLEN,

            SUM(LANE3_VOL) AS SUM_OF_LANE3_VOL,
            (CASE
                  WHEN SUM(LANE3_VOL) > 0 THEN SUM(LANE3_SPD*LANE3_VOL)/SUM(LANE3_VOL)
                  WHEN SUM(LANE3_VOL) = 0 THEN 0
                  ELSE NULL
            END) AVG_OF_LANE3_SPD,
            AVG(LANE3_OCC) AS AVG_OF_LANE3_OCC,
            /* vehicle length = velocity*5280*occupancy/100/flow */
            (CASE
                  WHEN SUM(LANE3_VOL) > 0 THEN
                      (SUM(LANE3_SPD*LANE3_VOL)/SUM(LANE3_VOL) * 5280)
                       * (AVG(LANE3_OCC) /100) /(SUM(LANE3_VOL)/24)
                  WHEN SUM(LANE3_VOL) = 0 THEN 0
                  ELSE NULL
            END) LANE3_EFFVEHLEN,

            SUM(LANE4_VOL) AS SUM_OF_LANE4_VOL,
            (CASE
                  WHEN SUM(LANE4_VOL) > 0 THEN SUM(LANE4_SPD*LANE4_VOL)/SUM(LANE4_VOL)
                  WHEN SUM(LANE4_VOL) = 0 THEN 0
                  ELSE NULL
            END) AVG_OF_LANE4_SPD,
            AVG(LANE4_OCC) AS AVG_OF_LANE4_OCC,
            /* vehicle length = velocity*5280*occupancy/100/flow */
```

Table 28 (Continued)

```
          (CASE
                WHEN SUM(LANE4_VOL) > 0 THEN
                    (SUM(LANE4_SPD*LANE4_VOL)/SUM(LANE4_VOL) * 5280)
                    * (AVG(LANE4_OCC) /100) / (SUM(LANE4_VOL)/24)
                WHEN SUM(LANE4_VOL) = 0 THEN 0
                ELSE NULL
          END) LANE4_EFFVEHLEN,

          SUM(LANE5_VOL) AS SUM_OF_LANE5_VOL,
          (CASE
                WHEN SUM(LANE5_VOL) > 0 THEN SUM(LANE5_SPD*LANE5_VOL)/SUM(LANE5_VOL)
                WHEN SUM(LANE5_VOL) = 0 THEN 0
                ELSE NULL
          END) AVG_OF_LANE5_SPD,
          AVG(LANE5_OCC) AS AVG_OF_LANE5_OCC,
          /* vehicle length = velocity*5280*occupancy/100/flow */
          (CASE
                WHEN SUM(LANE5_VOL) > 0 THEN
                    (SUM(LANE5_SPD*LANE5_VOL)/SUM(LANE5_VOL) * 5280)
                    * (AVG(LANE5_OCC) /100) / (SUM(LANE5_VOL)/24)
                WHEN SUM(LANE5_VOL) = 0 THEN 0
                ELSE NULL
          END) LANE5_EFFVEHLEN,

          SUM(LANE6_VOL) AS SUM_OF_LANE6_VOL,
          (CASE
                WHEN SUM(LANE6_VOL) > 0 THEN SUM(LANE6_SPD*LANE6_VOL)/SUM(LANE6_VOL)
                WHEN SUM(LANE6_VOL) = 0 THEN 0
                ELSE NULL
          END) AVG_OF_LANE6_SPD,
          AVG(LANE6_OCC) AS AVG_OF_LANE6_OCC,
          /* vehicle length = velocity*5280*occupancy/100/flow */
          (CASE
                WHEN SUM(LANE6_VOL) > 0 THEN
                    (SUM(LANE6_SPD*LANE6_VOL)/SUM(LANE6_VOL) * 5280)
                    * (AVG(LANE6_OCC) /100) / (SUM(LANE6_VOL)/24)
                WHEN SUM(LANE6_VOL) = 0 THEN 0
                ELSE NULL
          END) LANE6_EFFVEHLEN

        FROM EFF_VEH_LEN_1HR
        WHERE date1>=c_startdate1 AND date1<=c_enddate1
            AND TO_NUMBER(SUBSTR(TIME, 1, 2))>= c_starttime
            AND TO_NUMBER(SUBSTR(TIME, 1, 2))<= c_endtime
            AND DISTRICT = c_district
        GROUP BY DATE1, FACILITY, STATION_ID, DIRECTION, STATION_DESC, STATION_MP;

        ty             EffVehLen_fullperiod_ty;
        tbl            EffVehLen_fullperiod := EffVehLen_fullperiod ();
-- main body
BEGIN
      FOR rec IN cur (p_startdate, p_enddate, p_starttime, p_endtime, p_district)
      LOOP
         ty := EffVehLen_fullperiod_ty (rec.DATE1,
                                        rec.FACILITY,
                                        rec.STATION_ID,
                                        rec.DIRECTION,
                                        rec.STATION_DESC,
                                        rec.STATION_MP,
                                        rec.SUM_OF_LANE1_VOL,
                                        rec.AVG_OF_LANE1_SPD,
                                        rec.AVG_OF_LANE1_OCC,
                                        rec.LANE1_EFFVEHLEN,
                                        rec.SUM_OF_LANE2_VOL,
                                        rec.AVG_OF_LANE2_SPD,
                                        rec.AVG_OF_LANE2_OCC,
                                        rec.LANE2_EFFVEHLEN,
                                        rec.SUM_OF_LANE3_VOL,
                                        rec.AVG_OF_LANE3_SPD,
```

Table 28 (Continued)

```
                                        rec.AVG_OF_LANE3_OCC,
                                        rec.LANE3_EFFVEHLEN,
                                        rec.SUM_OF_LANE4_VOL,
                                        rec.AVG_OF_LANE4_SPD,
                                        rec.AVG_OF_LANE4_OCC,
                                        rec.LANE4_EFFVEHLEN,
                                        rec.SUM_OF_LANE5_VOL,
                                        rec.AVG_OF_LANE5_SPD,
                                        rec.AVG_OF_LANE5_OCC,
                                        rec.LANE5_EFFVEHLEN,
                                        rec.SUM_OF_LANE6_VOL,
                                        rec.AVG_OF_LANE6_SPD,
                                        rec.AVG_OF_LANE6_OCC,
                                        rec.LANE6_EFFVEHLEN);
        tbl.EXTEND;
        tbl (tbl.LAST) := ty;
     END LOOP;
     RETURN tbl;
  END;
```

3.5.3 Function Q_TRAFFICCOUNTS

This function receives the input parameters of start date, start time, end date, end time, and district, and creates the return variables as a table of daily average values for all data fields. Table 29 shows the return table for STEWARD Q_TrafficCounts.

Table 29 STEWARD function – definitions of Q_TrafficCounts

Name	Data Type (Length/Precision)
DATE1	DATE
DISTRICT	NUMBER
FACILITY	NUMBER
STATION_ID	NUMBER
NUM_OF_LANES	NUMBER
DIRECTION	VARCHAR2(1)
STATION_DESC	VARCHAR2(200)
STATION_MP	NUMBER
TOTAL	NUMBER
LANE1_VOL	NUMBER
LANE2_VOL	NUMBER
LANE3_VOL	NUMBER
LANE4_VOL	NUMBER
LANE5_VOL	NUMBER
LANE6_VOL	NUMBER
BALANCE	NUMBER
FWY_QA	NUMBER
ON_RAMP1	NUMBER
ON_RAMP2	NUMBER
ON_RAMP3	NUMBER
ON_RAMP_QA	NUMBER
OFF_RAMP1	NUMBER
OFF_RAMP2	NUMBER
OFF_RAMP3	NUMBER
OFF_RAMP_QA	NUMBER
COUNTY	NUMBER
COUNT_STATION	VARCHAR2(40)

Table 30 shows the function Q_TrafficCounts.

Table 30 STEWARD function - Q_ TrafficCounts

```
--initialize variables here
      CURSOR cur (c_startdate1 date, c_enddate1 date, c_starttime number, c_endtime
number, c_district number)
      IS
         SELECT
               DATE1,
               DISTRICT,
               FACILITY,
               STATION_ID,
               NUM_OF_LANES,
               DIRECTION,
               STATION_DESC,
               STATION_MP,
               (COALESCE(SUM(LANE1_VOL),0) + COALESCE(SUM(LANE2_VOL),0)
                 + COALESCE(SUM(LANE3_VOL),0) + COALESCE(SUM(LANE4_VOL),0)
                 + COALESCE(SUM(LANE5_VOL),0) + COALESCE(SUM(LANE6_VOL),0)) AS TOTAL,
               SUM(LANE1_VOL) AS SUM_LANE1_VOL,
               SUM(LANE2_VOL) AS SUM_LANE2_VOL,
               SUM(LANE3_VOL) AS SUM_LANE3_VOL,
               SUM(LANE4_VOL) AS SUM_LANE4_VOL,
               SUM(LANE5_VOL) AS SUM_LANE5_VOL,
               SUM(LANE6_VOL) AS SUM_LANE6_VOL,
               (CASE
                   WHEN SUM(LANE1_VOL) IS NULL THEN NULL
                   WHEN (SUM(LANE2_VOL) IS NULL) AND (SUM(LANE1_VOL) > 0) THEN 1
                   WHEN (SUM(LANE3_VOL) IS NULL)
                       AND (LEAST(SUM(LANE1_VOL), SUM(LANE2_VOL)) >0 )
                   THEN GREATEST (SUM(LANE1_VOL), SUM(LANE2_VOL))
                        / LEAST(SUM(LANE1_VOL), SUM(LANE2_VOL))
                   WHEN (SUM(LANE4_VOL) IS NULL)
                       AND (LEAST(SUM(LANE1_VOL), SUM(LANE2_VOL), SUM(LANE3_VOL)) >0 )
                   THEN GREATEST (SUM(LANE1_VOL), SUM(LANE2_VOL), SUM(LANE3_VOL))
                        / LEAST(SUM(LANE1_VOL), SUM(LANE2_VOL), SUM(LANE3_VOL))
                   WHEN (SUM(LANE5_VOL) IS NULL)
                       AND (LEAST(SUM(LANE1_VOL), SUM(LANE2_VOL), SUM(LANE3_VOL),
                          SUM(LANE4_VOL)) >0 )
                   THEN GREATEST (SUM(LANE1_VOL), SUM(LANE2_VOL), SUM(LANE3_VOL),
                                  SUM(LANE4_VOL))
                        / LEAST(SUM(LANE1_VOL), SUM(LANE2_VOL), SUM(LANE3_VOL),
                                SUM(LANE4_VOL))
                   WHEN (SUM(LANE6_VOL) IS NULL) AND (LEAST(SUM(LANE1_VOL),
                         SUM(LANE2_VOL), SUM(LANE3_VOL), SUM(LANE4_VOL),
                         SUM(LANE5_VOL)) >0 )
                   THEN GREATEST (SUM(LANE1_VOL), SUM(LANE2_VOL), SUM(LANE3_VOL),
                                  SUM(LANE4_VOL), SUM(LANE5_VOL))
                        / LEAST(SUM(LANE1_VOL), SUM(LANE2_VOL), SUM(LANE3_VOL),
                                SUM(LANE4_VOL), SUM(LANE5_VOL))
                   WHEN LEAST(SUM(LANE1_VOL), SUM(LANE2_VOL), SUM(LANE3_VOL),
                              SUM(LANE4_VOL), SUM(LANE5_VOL)) >0
                   THEN GREATEST (SUM(LANE1_VOL), SUM(LANE2_VOL), SUM(LANE3_VOL),
                                  SUM(LANE4_VOL), SUM(LANE5_VOL), SUM(LANE6_VOL))
                      / LEAST(SUM(LANE1_VOL), SUM(LANE2_VOL), SUM(LANE3_VOL),
                              SUM(LANE4_VOL), SUM(LANE5_VOL), SUM(LANE6_VOL))
                   ELSE NULL
               END) BALANCE,
               ROUND(AVG(FWY_QA),2) AVG_FWY_QA,
               SUM(ON_RAMP1) AS SUM_ON_RAMP1,
               SUM(ON_RAMP2) AS SUM_ON_RAMP2,
               SUM(ON_RAMP3) AS SUM_ON_RAMP3,
               ROUND(AVG(ON_RAMP_QA),2) AVG_ON_RAMP_QA,
               SUM(OFF_RAMP1) AS SUM_OFF_RAMP1,
               SUM(OFF_RAMP2) AS SUM_OFF_RAMP2,
               SUM(OFF_RAMP3) AS SUM_OFF_RAMP3,
               ROUND(AVG(OFF_RAMP_QA) ,2) AS AVG_OFF_RAMP_QA,
               COUNTY,
               COUNT_STATION
```

Table 30 (Continued)

```
        FROM GATOR.TRAFFIC_COUNTS_1HR
        WHERE date1>=c_startdate1 AND date1<=c_enddate1
             AND TO_NUMBER(SUBSTR(TIME, 1, 2))>= c_starttime
             AND TO_NUMBER(SUBSTR(TIME, 1, 2))<= c_endtime AND DISTRICT = c_district
        GROUP BY DATE1, DISTRICT, FACILITY, STATION_ID, NUM_OF_LANES, DIRECTION,
                 STATION_DESC, STATION_MP, COUNTY, COUNT_STATION
        ORDER BY DATE1 ASC;
        ty             TrafficCounts_fullperiod_ty;
        tbl            TrafficCounts_fullperiod := TrafficCounts_fullperiod ();
-- main body
BEGIN
      FOR rec IN cur (p_startdate, p_enddate, p_starttime, p_endtime, p_district)
      LOOP
         ty := TrafficCounts_fullperiod_ty (rec.DATE1,
                                            rec.DISTRICT,
                                            rec.FACILITY,
                                            rec.STATION_ID,
                                            rec.NUM_OF_LANES,
                                            rec.DIRECTION,
                                            rec.STATION_DESC,
                                            rec.STATION_MP,
                                            rec.TOTAL,
                                            rec.SUM_LANE1_VOL,
                                            rec.SUM_LANE2_VOL,
                                            rec.SUM_LANE3_VOL,
                                            rec.SUM_LANE4_VOL,
                                            rec.SUM_LANE5_VOL,
                                            rec.SUM_LANE6_VOL,
                                            rec.BALANCE,
                                            rec.AVG_FWY_QA,
                                            rec.SUM_ON_RAMP1,
                                            rec.SUM_ON_RAMP2,
                                            rec.SUM_ON_RAMP3,
                                            rec.AVG_ON_RAMP_QA,
                                            rec.SUM_OFF_RAMP1,
                                            rec.SUM_OFF_RAMP2,
                                            rec.SUM_OFF_RAMP3,
                                            rec.AVG_OFF_RAMP_QA,
                                            rec.COUNTY,
                                            rec.COUNT_STATION);
         tbl.EXTEND;
         tbl (tbl.LAST) := ty;
      END LOOP;
      RETURN tbl;
   END;
```

3.5.4 Function Q_TSS

This function receives the input parameters of start date, start time, end date, end time, and district, and creates the return variables as a table of daily average values for all data fields. Table 31 shows the return table for STEWARD Q_TSS.

Table 31 STEWARD function – definitions of Q_TSS

Name	Data Type (Length/Precision)
DATE_REF	NUMBER
TIME_REF	NUMBER
TSS_STATION	NUMBER
FWY_SPD	NUMBER
FWY_VOL	NUMBER
FWY_OCC	NUMBER
SPD_CV	NUMBER
VOL_RATIO	NUMBER
SPD_RATIO	NUMBER
ENTRY_VOL	NUMBER
EXIT_VOL	NUMBER
FWY_QA	NUMBER
ENTRY_QA	NUMBER
EXIT_QA	NUMBER
HOV_SPD	NUMBER
HOV_VOL	NUMBER
HOV_OCC	NUMBER
HOV_QA	NUMBER

Table 32 shows the function Q_TSS.

Table 32 STEWARD function - Q_ TSS

```
--initialize variables here

      CURSOR cur (c_startdate date, c_enddate date, c_starttime number, c_endtime number,
c_district number)
      IS
         SELECT
             TSS_1HR.DATE_REF,
             TSS_1HR.TIME_REF,
             TSS_1HR.TSS_STATION,
             (CASE
                 WHEN SUM(FWY_VOL)>0 THEN SUM(FWY_VOL*FWY_SPD)/SUM(FWY_VOL)
                 ELSE NULL
             END) AVG_FWY_SPD,
             SUM(FWY_VOL) AS SUM_FWY_VOL,
             AVG(FWY_OCC) AS AVG_FWY_OCC,
             NULL AS SPD_CV,    --SPD_CV
             MAX(VOL_RATIO) AS MAX_VOL_RATIO,
             MAX(SPD_RATIO) AS MAX_SPD_RATIO,
             SUM(ENTRY_VOL) AS SUM_ENTRY_VOL,
             SUM(EXIT_VOL) AS SUM_EXIT_VOL,
             AVG(FWY_QA) AS AVG_FWY_QA,
             AVG(ENTRY_QA) AS AVG_ENTRY_QA,
             AVG(EXIT_QA) AS AVG_EXIT_QA,
             (CASE
                 WHEN SUM(HOV_VOL)>0 THEN SUM(HOV_VOL*HOV_SPD)/SUM(HOV_VOL)
                 ELSE NULL
             END) AVG_HOV_SPD,
             SUM(HOV_VOL) AS SUM_HOV_VOL,
             AVG(HOV_OCC) AS AVG_HOV_OCC,
             AVG(HOV_QA) AS AVG_HOV_QA
```

Table 32 (Continued)

```
         FROM TSS_1HR,
              (SELECT DIMENSION_KEY AS STARTDATE FROM DATE_REF
                   WHERE DAY=c_startdate) DATE_REF1,
              (SELECT DIMENSION_KEY AS ENDDATE FROM DATE_REF
                   WHERE DAY=c_enddate) DATE_REF2,
              (SELECT DIMENSION_KEY AS STARTTIME FROM TIME_REF
                   WHERE TO_NUMBER(SUBSTR(TIME, 1, 2))=c_starttime) TIME_REF1,
              (SELECT DIMENSION_KEY AS ENDTIME FROM TIME_REF
                   WHERE TO_NUMBER(SUBSTR(TIME, 1, 2))=c_endtime) TIME_REF2,
              (SELECT DIMENSION_KEY AS DISTRICT FROM TSS_STATION
                   WHERE TO_NUMBER(SUBSTR(STATION_ID, 1, 1))=c_district) TSS_STATION1
         WHERE TSS_1HR.DATE_REF >= DATE_REF1.STARTDATE
              AND TSS_1HR.DATE_REF <= DATE_REF2.ENDDATE
              AND TSS_1HR.TIME_REF >= TIME_REF1.STARTTIME
              AND TSS_1HR.TIME_REF <= TIME_REF2.ENDTIME
              AND TSS_1HR.TSS_STATION <= TSS_STATION1.DISTRICT
         GROUP BY DATE_REF, TIME_REF, TSS_STATION;
      ty             TSS_fullperiod_ty;
      tbl            TSS_fullperiod := TSS_fullperiod ();

-- main body
BEGIN
      FOR rec IN cur (p_startdate, p_enddate, p_starttime, p_endtime, p_district)
      LOOP
         ty := TSS_fullperiod_ty (rec.DATE_REF,
                                          rec.TIME_REF,
                                          rec.TSS_STATION,
                                          rec.AVG_FWY_SPD,
                                          rec.SUM_FWY_VOL,
                                          rec.AVG_FWY_OCC,
                                          rec.SPD_CV,
                                          rec.MAX_VOL_RATIO,
                                          rec.MAX_SPD_RATIO,
                                          rec.SUM_ENTRY_VOL,
                                          rec.SUM_EXIT_VOL,
                                          rec.AVG_FWY_QA,
                                          rec.AVG_ENTRY_QA,
                                          rec.AVG_EXIT_QA,
                                          rec.AVG_HOV_SPD,
                                          rec.SUM_HOV_VOL,
                                          rec.AVG_HOV_OCC,
                                          rec.AVG_HOV_QA);
         tbl.EXTEND;
         tbl (tbl.LAST) := ty;
      END LOOP;
      RETURN tbl;
   END;
```

3.5.5 Function Q_VolumeMap

This function receives the input parameters of start date, start time, end date, end time, and district, and creates the return variables as a table of daily average values for all data fields. Table 33 shows the return table for STEWARD Q_VolumeMap.

Table 33 STEWARD function – definitions of Q_VolumeMap

Name	Data Type (Length/Precision)
DATE1	DATE
FACILITY	NUMBER
STATION_ID	NUMBER
DIRECTION	NUMBER
STATION_MP	NUMBER
UPNODE_ID	NUMBER
ENTRY_VOLUME	NUMBER
FWY_VOLUME	NUMBER
EXIT_VOLUME	NUMBER
LINK_INPUT	NUMBER
LINK_OUTPUT	NUMBER
DIFFERENCE	NUMBER
PCNT_DIFF	NUMBER

Table 34 shows the function Q_VolumeMap.

Table 34 STEWARD function - Q_ VolumeMap

```
--initialize variables here

      CURSOR cur (c_startdate1 date, c_enddate1 date, c_starttime number, c_endtime
number, c_district number)
      IS
          SELECT
              VOLUME_MAP_5MIN_SUB.DATE1 AS DATE1_VM,             /* Date */
              VOLUME_MAP_5MIN_SUB.FACILITY AS FACILITY_VM,         /* Facility */
              VOLUME_MAP_5MIN_SUB.STATION_ID AS STATION_ID_VM,       /* Destination */
              VOLUME_MAP_5MIN_SUB.DIRECTION AS DIRECTION_VM,         /* DIRECTION */
              VOLUME_MAP_5MIN_SUB.STATION_MP AS STATION_MP_VM,   /*Destination milepost*/
              VOLUME_MAP_5MIN_SUB.UPNODE_ID AS UPNODE_ID_VM,        /* Origin */
              SUM(VOLUME_MAP_5MIN_SUB.ENTRY_VOLUME) AS ENTRY_VOL_VM,  /*ENTRY Vol*/
              SUM(VOLUME_MAP_5MIN_SUB.FWY_VOLUME) AS FWY_VOL_VM,  /*Main Vol*/
              SUM(VOLUME_MAP_5MIN_SUB.EXIT_VOLUME) AS EXIT_VOL_VM,  /*Exit Vol*/
              SUM(VOLUME_MAP_5MIN_SUB.LINK_INPUT) AS LINK_INPUT_VM,  /*LINK INPUT*/
              SUM(VOLUME_MAP_5MIN_SUB.LINK_OUTPUT) AS LINK_OUTPUT_VM,  /*LINK OUNPUT*/
              (SUM(VOLUME_MAP_5MIN_SUB.LINK_INPUT)
                 -SUM(VOLUME_MAP_5MIN_SUB.LINK_OUTPUT)) AS VOLUME_DIFF_VM,
                 /*Volume Difference*/
             (CASE
                 WHEN (SUM(LINK_INPUT) + SUM(LINK_OUTPUT)) > 0
                     THEN (100 * 2*(SUM(LINK_INPUT) - SUM(LINK_OUTPUT))
                          / (SUM(LINK_INPUT) + SUM(LINK_OUTPUT)))
                 ELSE NULL
              END) PCNT_VOLUME_DIFF_VM   /*Percentage Volume Difference*/

           FROM (SELECT * FROM VOLUME_MAP_5MIN
                    WHERE date1>=c_startdate1 AND date1<=c_enddate1
                          AND TO_NUMBER(SUBSTR(TIME, 1, 2))>= c_starttime
                          AND TO_NUMBER(SUBSTR(TIME, 1, 2))<= c_endtime
                          AND DISTRICT = c_district) VOLUME_MAP_5MIN_SUB
           GROUP BY VOLUME_MAP_5MIN_SUB.DATE1,  VOLUME_MAP_5MIN_SUB.FACILITY,
                    VOLUME_MAP_5MIN_SUB.STATION_ID, VOLUME_MAP_5MIN_SUB.DIRECTION,
                       VOLUME_MAP_5MIN_SUB.STATION_MP, VOLUME_MAP_5MIN_SUB.UPNODE_ID;

         ty            VOLUMEMAP_fullperiod_ty;
         tbl           VOLUMEMAP_fullperiod := VOLUMEMAP_fullperiod ();

-- main body

BEGIN
      FOR rec IN cur (p_startdate, p_enddate, p_starttime, p_endtime, p_district)
      LOOP
         ty := VOLUMEMAP_fullperiod_ty (rec.DATE1_VM,
                                        rec.FACILITY_VM,
                                        rec.STATION_ID_VM,
                                        rec.DIRECTION_VM,
                                        rec.STATION_MP_VM,
                                        rec.UPNODE_ID_VM,
                                        rec.ENTRY_VOL_VM,
                                        rec.FWY_VOL_VM,
                                        rec.EXIT_VOL_VM,
                                        rec.LINK_INPUT_VM,
                                        rec.LINK_OUTPUT_VM,
                                        rec.VOLUME_DIFF_VM,
                                        rec.PCNT_VOLUME_DIFF_VM);
         tbl.EXTEND;
         tbl (tbl.LAST) := ty;
      END LOOP;
      RETURN tbl;
   END;
```

4 The STEWARD ETL Process

The overall ETL process and its operations are described in Appendix 2 of this report.

The brief summary of STEWRD ETL process is provided here. The ETL functions are categorized as follows:

- *Extract:* STEWARD manages the FTP server and expects each district send the archive data everyday. Currently, District 2,4,5,6, and 7 create the archived files daily and send them to district-specific FTP sites automatically as a scheduled task.
- *Transform:* As soon as STEWARD receives the data from districts, it archives, verifies and processes them. Raw data and processed data are all archived for backup purposes. The quality of the data is checked and verified. Then the data are converted to the STEWARD data format and aggregated at the station or lane-levels. The SunETLUtility program was developed using the VB6 platform to customize the process.
- *Load:* STEWARD uses Oracle database processing functions to load the data into the STEWARD database. After the data loading, the materialized views are updated to reflect the changes.

The Oracle Warehouse Builder is involved in the STEWARD ETL process in the following three modules, mappings, process flows and schedules. The details of each ETL process are described in the following sections.

4.1 Mappings

Mappings are defined in the Oracle Warehouse Builder as a series of operations that extract data from sources, transform it, and load it into targets. In STEWARD, the mappings were developed to load the data from the external processed files (outputs of SunETLUtility) into STEWARD database tables. Fifteen mappings are defined in the STEWARD system. All of the mappings except the DATE_REF_MAP have a similar architecture, mapping from the external files into the dimension and fact tables. In the case of DATE_REF_MAP, the mapping was generated automatically using the time wizard provided by the Oracle Warehouse Builder for dimensional tables.

Mappings for the dimension tables include

- DATE_REF_MAP
- TIME_REF_MAP
- LOAD_TSS_STATION

Mappings for the fact tables include

- LOADING_TSS_5MIN
- LOADING_TSS_15MIN
- LOADING_TSS_1HR
- LOADING_TSS_5MIN_LANE
- LOADING_TSS_15MIN_LANE
- LOADING_TSS_1HR_LANE

Figure 2 shows one example of STEWARD mapping. The external file (tss-15min_data.csv) is shown as the Oracle external table and treated as the table (TSS_15MIN_EX). The first column (DATE1) in the external table (TSS_15MIN_EX) is transformed using the expression utility. In this case, the column type is changed from VARCHAR2 to DATE with the following format ('MM/DD/YYYY' → 'YYYYMMDD'). The following PL/SQL code is used for this purpose.

```
>> TO_NUMBER(TO_CHAR(TO_DATE(INGRP1.DATE1,'MM/DD/YYYY'),'YYYYMMDD'),'99999999')
```

The second column (Time) in the external table (TSS_15MIN_EX) is transformed using the expression utility. In this case, the column data is simply trimmed to eliminate the redundant space. The following PL/SQL code is used for this purpose.

```
>> TRIM( INGRP1.TIME )
```

The other columns in the external file are mapped into the fact table without any data conversion.

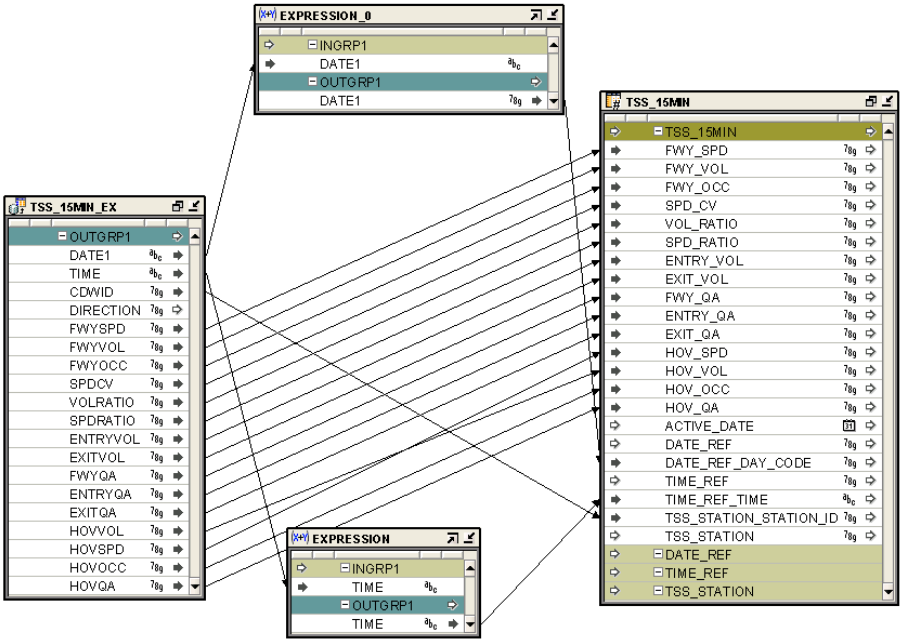


Figure 2 STEWARD mapping example – LOAD_TSS_15MIN

4.2 Process flow

Process flow modules were used to design and execute the STEWARD ETL process in the STEWARD Database server. Process flows can be executed by an Oracle Schedule or manually, using the Oracle Warehouse Builder control center. The STEWARD transform process was developed using an external program (SunETLUtility) but most of the loading process was developed using this process flow.

Four main process and nine sub-process flow modules are defined for STEWARD. The three main process flow modules are as follows

- PRELOADING
- LOADING_TSS
- LOADING_TSS_LANE

4.2.1 Pre-loading

This process moves the outputs of transform process (output of SunETLUtility) into the designated file locations for STEWARD external data files. SunETLUtility creates eight output files:

- Four Station-level files: 1min, 5min, 15min, and 60min aggregated data file
- Four Lane-level files: 1min, 5min, 15min, and 60min aggregated data file

Six of these files (5min, 15min and 60min data file for each level) are moved into the designated file locations for load process. The 1 minute aggregations are not accessible to the user from the website.

As shown in Figure 3, the outputs of SunETLUtility are determined from its file name and moved to the STEWARD loading folders. The default input folder (output folder of SunETLUtility) is c:\progra~1\TRC\SunETLUtility.

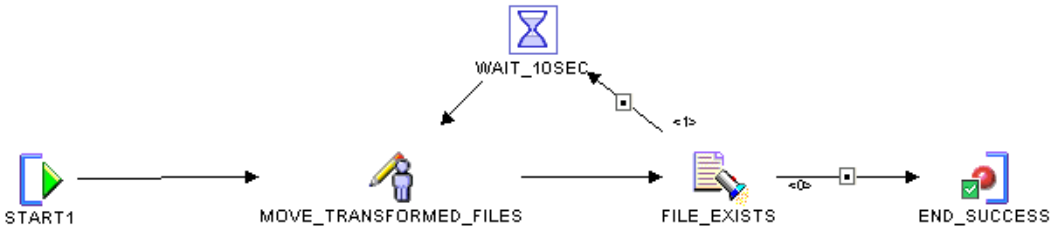


Figure 3 STEWARD process flow – stage3 pre-loading

The target data folders for data loading are as follows:

- C:\Steward\TSS_STATION_DATA\TSS_5MIN_Data
- C:\Steward\TSS_STATION_DATA\TSS_15MIN_Data
- C:\Steward\TSS_STATION_DATA\TSS_1Hr_Data
- C:\Steward\TSS_LANE_DATA\TSS_5MIN_Data
- C:\Steward\TSS_LANE_DATA\TSS_15MIN_Data
- C:\Steward\TSS_LANE_DATA\TSS_1Hr_Data

A custom-built program (Move2StewardFolder.exe) is executed within the MOVE_TRNSFORMED_FILES object.

4.2.2 LOADING_TSS, LOADING_TSS_LANE

As shown in Figure 4, there are two main process flows at this stage. These are for TSS station data and TSS lane data. Each main process includes three sub-process flows for 5min, 15min and 1hr data loading. The main processes are named as LOADG_TSS and LOADG_TSS_LANE

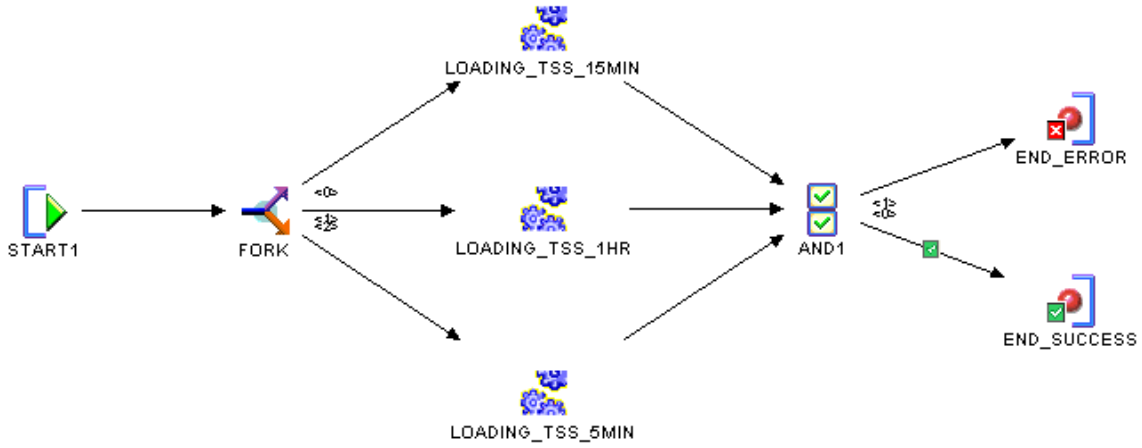


Figure 4 STEWARD process flow – stage4 loading

Six sub process flows are included:

- LOADING_TSS_5MIN
- LOADING_TSS_15MIN
- LOADING_TSS_1HR
- LOADING_TSS_5MIN_LANE
- LOADING_TSS_15MIN_LANE
- LOADING_TSS_1HR_LANE

As shown in Figure 5, each process loads one type of files into STEWARD database. Oracle external file loader allows only one predefined file name for the data loading. This means every input data file needs to change its file name as predefined file name before it loads.

PREPARE_FILE_LOADING object does the following steps

- Check and delete the existence of the predefined file in the designated folder.
- Rename one of the input data file to the predefined input data file.

LOAD_FILE_EXISTS object checks there are predefined file in the designated folder.
LOAD_TSS_15MIN object loads the predefined file into database.

These three object s are executed repeatedly until there are no input data files left.

Figure 5 shows one of six sub process (LOADING_TSS_15MIN). The other processes have the same scheme except the “LOAD_TSS_5MIN” object.

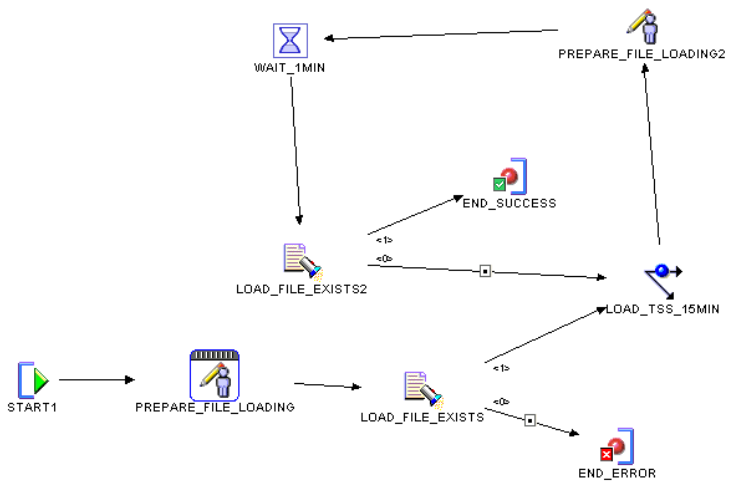


Figure 5 STEWARD process flow – substage LOADING_TSS_15MIN

A custom-built program (PrepareFileLoading.exe) is executed within the PREPARE_FILE_LOADING object.

4.3 Schedules

The STEWARD operator can use the scheduler to run the process flows at a predetermined time. The details of how to use the scheduler are described in Appendix 2.

5 STEWARD Web Interface

The STEWARD web program was developed to provide access to the data for general users. It communicates with the STEWARD database, retrieves the traffic data and generates the reports from dynamic queries. It was developed with Microsoft Internet Information Services (IIS) 5.1 using Active Server Pages (ASP), JavaScript and HTML.

5.1 Program Architecture

The STEWARD web program has a tree architecture. The site map is shown in Figure 6.

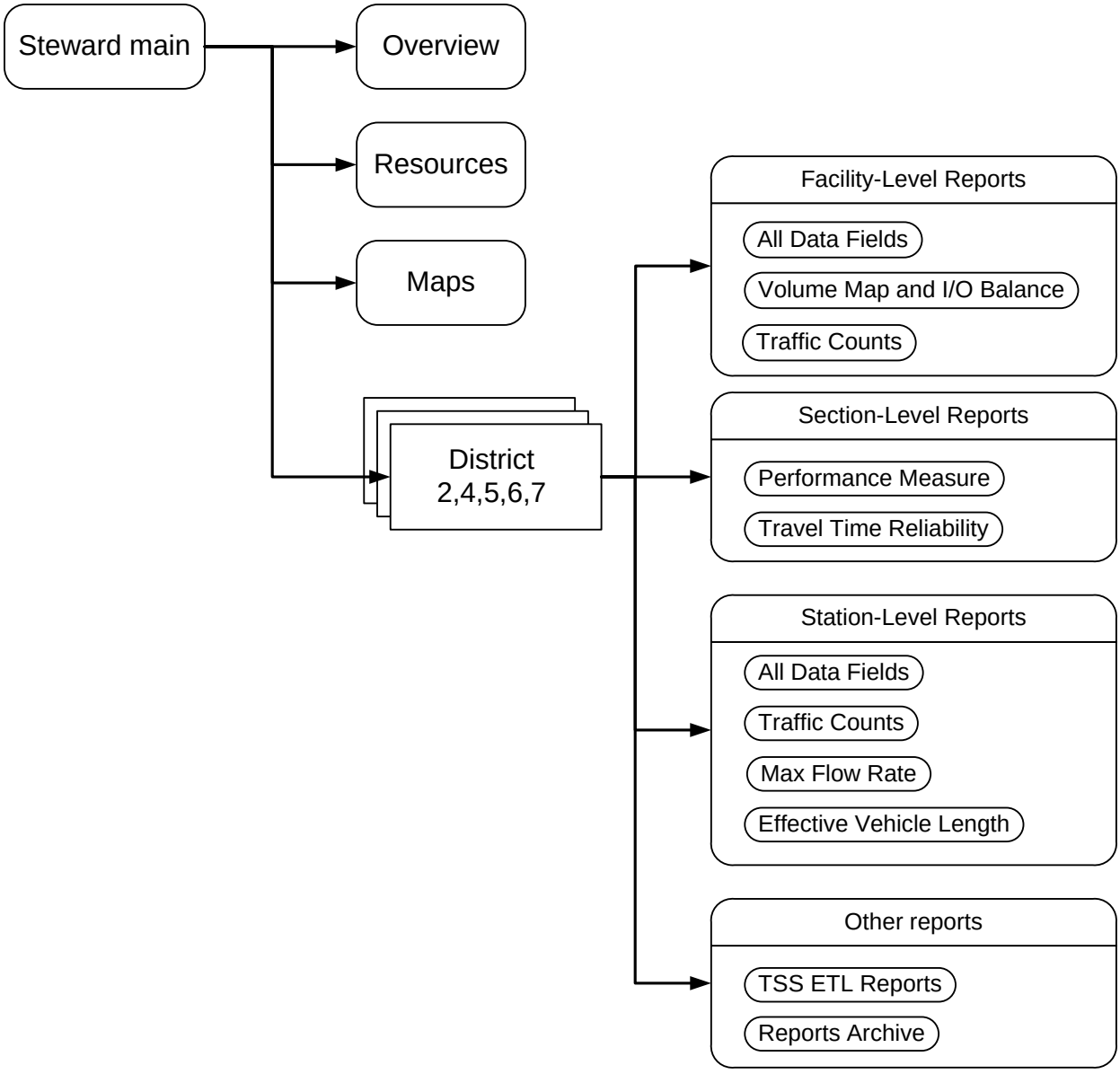


Figure 6 STEWARD web architecture

5.2 STEWARD Web Architecture

The STEWARD web program consists of one main folder and six sub folders. The main folder includes the program main and resource pages. The Index_main folder includes subfolders for district data and reports. The *WebFiles* and *Files* folders include SunETLUtility output and STEWARD web report outputs. Other folders (*images, common* and *css*) include utility files used in the STEWARD web program. The details are described in Table 35.

Table 35 Directory architecture of STEWARD web program

Folder name	Descriptions
wwwroot\steward	Includes STEWARD main page and three sub pages (STEWARD Overview, Resources and Contact Info)
~\files	Includes CSV files generated by the web users
~\maps	Includes map files: GIS files and google map interface
~\resouces	Includes resource files for the resource web pages
~\images ~\css ~\common	Utility files for STEWARD web programs
~\index_main	Main folder for STEWARD reports
~\index_main\tss_facility_level_reports	Main folder for STEWARD facility-level reports
~\index_main\tss_facility_level_reports\lane_volumes	All data fields report folder
~\index_main\tss_facility_level_reports\traffic_counts	Traffic counts report folder
~\index_main\tss_facility_level_reports\volume_map	Volume map report folder
~\index_main\tss_section_level_reports	Main folder for STEWARD section-level reports
~\index_main\tss_section_level_reports\performance_measure	Performance measure report folder
~\index_main\tss_section_level_reports\TT_Reliability	Travel time reliability report folder
~\index_main\tss_station_level_reports	Main folder for STEWARD station-level reports
~\index_main\tss_station_level_reports\volume_data	All data fields report folder
~\index_main\tss_station_level_reports\ traffic_counts	Traffic counts report folder
~\index_main\tss_station_level_reports\ max_flow_rate	Max flow rate report folder
~\index_main\tss_station_level_reports\Effective_vehicle_length	Effective vehicle length report folder
~\index_main\TSS_ETL_Reports	TSS ETL reports folder
~\index_main\csv_reports	CSV report archive folder

Appendix 2
STEWARD Operation

Table of Contents

1 INTRODUCTION 1

2 ORACLE DATABASE PROGRAM INSTALLATION 1

2.1 Install Oracle 10*g* Release 2 1
2.1.1 Run the Setup.exe File 1
2.1.2 Select Advanced Installation 1
2.1.3 Select Enterprise Edition 2
2.1.4 Select Default Location 2
2.1.5 Run the Prerequisite Checks 3
2.1.6 Select Database Software Only Option for Configuration Option 3
2.1.7 Oracle Database Installation 4

2.2 Install Oracle Workflow 5
2.2.1 Run the Setup.exe File 5
2.2.2 Select the Oracle Database 10g Product 6
2.2.3 Select the Home Name and Address 6
2.2.4 Run the Prerequisite Checks 7
2.2.5 Oracle Workflow Installation 7
2.2.6 Oracle Workflow Installation Completion 8

2.3 Installing the Oracle Warehouse Builder (OWB) 8
2.3.1 Run the Setup.exe File 8
2.3.2 Specify Home Details 9
2.3.3 Preparation to install the OWB 9
2.3.4 Install the OWB 9
2.3.5 Installation Completed (OWB) 10

2.4 Oracle Database Configuration Assistant 11
2.4.1 Oracle Database Configuration Assistant -Welcome 12
2.4.2 Oracle Database Configuration Assistant - Operations 12
2.4.3 Oracle Database Configuration Assistant -Database Templates 13
2.4.4 Oracle Database Configuration Assistant - Database Identification 13
2.4.5 Oracle Database Configuration Assistant -Management Options 14
2.4.6 Oracle Database Configuration Assistant - Database Credentials 14
2.4.7 Oracle Database Configuration Assistant -Storage Options 15
2.4.8 Oracle Database Configuration Assistant - Database File Locations 15
2.4.9 Oracle Database Configuration Assistant -Recovery Configuration 16
2.4.10 Oracle Database Configuration Assistant - Database Contents 16
2.4.11 Oracle Database Configuration Assistant -Initialization Parameters 17
2.4.12 Oracle Database Configuration Assistant - Database Storage 17
2.4.13 Oracle Database Configuration Assistant -Creation Options 18
2.4.14 Oracle Database Configuration Assistant - Confirmation 18
2.4.15 Oracle Database Configuration Assistant -Creating Database 19
2.4.16 Oracle Database Configuration Assistant - End of Creating Database 19
2.4.17 Oracle Database Configuration Assistant -Password Management 20

2.5 Oracle Net Configuration Assistant 21
2.5.1 Oracle Net Configuration Assistant - Listener Configuration 22
2.5.2 Oracle Net Configuration Assistant -Local Net Service Name Configuration 25
2.5.3 System Variables Setting 31

2.6 Oracle Enterprise Manager 32
2.6.1 Oracle Enterprise Manager - Login 32
2.6.2 Oracle Enterprise Manager - Database Control Main Page 33

2.7 Oracle Warehouse Builder Setup 34
2.7.1 Create the Design Center User and its Repository 35

2.8 Oracle Workflow Configuration 39
2.8.1 Create the owf_mgr Workflow Schema using the Workflow Configuration Assistant. 39
2.8.2 Set the Workflow Configuration 40
2.8.3 Unlock the owf_mgr Account 40

3 STEWARD DEPLOYMENT 42

3.1 First Step - Login 42

3.2 Prerequisites- Create a Target User 42

3.3 Prerequisites- Create New Files 47
3.3.1 Oracle Warehouse Builder - Create New File 47
3.3.2 Oracle Warehouse Builder - Create Module Wizard 49
3.3.3 Oracle Warehouse Builder - Name and Description 49
3.3.4 Oracle Warehouse Builder - Select Location 50
3.3.5 Oracle Warehouse Builder - Select the Location Path 50
3.3.6 Oracle Warehouse Builder - Finish Connection Information 51
3.3.7 Oracle Warehouse Builder - Module is Created 51
3.3.8 Oracle Warehouse Builder – Importing from the New Module 52
3.3.9 Oracle Warehouse Builder - Import Metadata Wizard 52
3.3.10 Oracle Warehouse Builder - Filter Information 53
3.3.11 Oracle Warehouse Builder - Object Selection 53
3.3.12 Oracle Warehouse Builder - Run the Sample 54
3.3.13 Oracle Warehouse Builder - Flat File Sample Wizard 54
3.3.14 Oracle Warehouse Builder - Check the File Name 55
3.3.15 Oracle Warehouse Builder - Record Organization, File Format and File Layout 55
3.3.16 Oracle Warehouse Builder - Field Properties 56
3.3.17 Oracle Warehouse Builder - Summary 58
3.3.18 Oracle Warehouse Builder - Import Data Wizard Summary 58

3.4 Importing Metadata 59
3.4.1 Prerequisite- Export Warehouse Builder Metadata (Step 1) 59
3.4.2 Prerequisite- Export Warehouse Builder Metadata (Step 2) 59
3.4.3 Importing Metadata - Import Warehouse Builder Metadata (Step 1) 60
3.4.4 Importing Metadata - Import Warehouse Builder Metadata (Step 2) 61
3.4.5 Importing Metadata - Import Warehouse Builder Metadata (Step 3) 61

3.5 Registration of the Control Center Manager 62

3.6 Data Deployment Process 66
3.6.1 External Table Deployment 66
3.6.2 Dimension Table Deployment 66
3.6.3 Fact Table Deployment 67
3.6.4 Mappings Deployment 67

3.7 Data loading process 69
3.7.1 Dimension Table Data loading 69

3.7.2 Loading Fact Table Data 69

4 STEWARD WEB INSTALLATION 71

4.1 Net Configuration Assistant Steps 71

4.2 STEWARD Web Program Installation 76

4.3 System Configuration 76
4.3.1 Firewall Setting 76
4.3.2 Permission for File Sharing 78

4.4 Web Program Configuration 79
4.4.1 Web Server Address 80
4.4.2 Google Map API Key 80
4.4.3 STEWARD DB Login Information 80

5 STEWARD MANAGEMENT 81

5.1 STEWARD Daily Operation 81
5.1.1 Data Transfer from District SunGuide Systems 81
5.1.2 Data Backup and Transformation in the STEWARD FTP Server 82
5.1.3 Data Loading into the STEWARD Database 83
5.1.4 Refresh Configuration for Materialized Views in the STEWARD Database 87
5.1.5 Update the Materialized Views in the STEWARD Database 89
5.1.6 Backup Plan and Procedure for the STEWARD Database and Web 90

5.2 STEWARD Operation- Adding a New District Facility 92
5.2.1 Changes in ETL Process 92
5.2.2 Updates to the STEWARD Database 93
5.2.3 Updates to the STEWARD Web Interface 93

5.3 Known Issues 94

List of Figures

Figure 1 Oracle DB installation - installation method 1
Figure 2 Oracle DB installation - Installation type 2
Figure 3 Oracle DB installation - installation home details 2
Figure 4 Oracle DB installation - prerequisite checks 3
Figure 5 Oracle DB installation - configuration option 3
Figure 6 Oracle DB installation - installation Summary 4
Figure 7 Installation Summary - installation 4
Figure 8 Oracle DB installation - end of installation 5
Figure 9 Oracle Workflow installation 5
Figure 10 Oracle Workflow installation - select a product to install 6
Figure 11 Oracle Workflow installation - installation home details 6
Figure 12 Oracle Workflow installation - prerequisite checks 7
Figure 13 Oracle Workflow installation - installation Summary 7
Figure 14 Oracle Workflow installation - end of installation 8
Figure 15 Oracle Warehouse Builder (OWB) installation - welcome 8
Figure 16 OWB installation - specify home details 9
Figure 17 OWB installation - summary information for installation 9
Figure 18 OWB installation 10
Figure 19 OWB installation - end of installation 10
Figure 20 Oracle Database Configuration Assistant - run program and icon 11
Figure 21 Oracle Database Configuration Assistant - Welcome 12
Figure 22 Oracle Database Configuration Assistant - operations 12
Figure 23 Oracle Database Configuration Assistant - database templates 13
Figure 24 Oracle Database Configuration Assistant - database identification 13
Figure 25 Oracle Database Configuration Assistant - management options 14
Figure 26 Oracle Database Configuration Assistant - database credentials 14
Figure 27 Oracle Database Configuration Assistant - storage options 15
Figure 28 Oracle Database Configuration Assistant - database file locations 15
Figure 29 Oracle Database Configuration Assistant - recovery configuration 16
Figure 30 Oracle Database Configuration Assistant - database contents 16
Figure 31 Oracle Database Configuration Assistant - initialization parameters 17
Figure 32 Oracle Database Configuration Assistant - database storage 17
Figure 33 Oracle Database Configuration Assistant - creation options 18
Figure 34 Oracle Database Configuration Assistant - confirmation 18
Figure 35 Oracle Database Configuration Assistant - creating the database 19
Figure 36 Oracle Database Configuration Assistant - end of creating database 19
Figure 37 Oracle Database Configuration Assistant - password management 20
Figure 38 Oracle Net Configuration Assistant - run the ‘Net Configuration Assistant’ 21
Figure 39 Oracle Net Configuration Assistant - listener configuration 22
Figure 40 Oracle Net Configuration Assistant - listener configuration 22
Figure 41 Oracle Net Configuration Assistant - listener name 23
Figure 42 Oracle Net Configuration Assistant - Select Protocols 23
Figure 43 Oracle Net Configuration Assistant - TCP/IP protocols and port number 24
Figure 44 Oracle Net Configuration Assistant - listener configuration 24

Figure 45 Oracle Net Configuration Assistant -listeners configuration done..... 25
Figure 46 Oracle Net Configuration Assistant -local net service name configuration..... 25
Figure 47 Oracle Net Configuration Assistant -local net service name configuration (Add) 26
Figure 48 Oracle Net Configuration Assistant -local net service name configuration (Service Name)..... 26
Figure 49 Oracle Net Configuration Assistant -local net service name configuration (Select Protocols)..... 27
Figure 50 Oracle Net Configuration Assistant -local net service name configuration (TCP/IP Protocols)..... 27
Figure 51 Oracle Net Configuration Assistant -local net service name configuration (Connection Test)..... 28
Figure 52 Oracle Net Configuration Assistant -local net service name configuration (Failure).. 28
Figure 53 Oracle Net Configuration Assistant -change login..... 29
Figure 54 Oracle Net Configuration Assistant -connection succeeded..... 29
Figure 55 Oracle Net Configuration Assistant -net service name..... 30
Figure 56 Oracle Net Configuration Assistant - another net service name?..... 30
Figure 57 Oracle Net Configuration Assistant -net service name (configuration done)..... 31
Figure 58 System variables setting..... 32
Figure 59 Oracle Enterprise Manager - login..... 33
Figure 60 Oracle Enterprise Manager - main page..... 33
Figure 61 Oracle WarehouseBuilder configuration..... 34
Figure 62 Oracle Warehouse Builder login window..... 35
Figure 63 Oracle Warehouse Builder- installation type..... 35
Figure 64 Oracle Warehouse Builder- repository user and connection information..... 36
Figure 65 Oracle Warehouse Builder- password confirmation..... 37
Figure 66 Oracle Warehouse Builder- repository owner information..... 37
Figure 67 Oracle Warehouse Builder- installation progress..... 38
Figure 68 Oracle Warehouse Builder- installation summary..... 38
Figure 69 Oracle workflow configuration..... 39
Figure 70 Oracle Workflow Configuration Assistant..... 40
Figure 71 Oracle workflow configuration - unlock the owf_mgr..... 41
Figure 72 Oracle workflow configuration - unlock the owf_mgr..... 42
Figure 73 Oracle Warehouse Builder - create DB user..... 43
Figure 74 Oracle Warehouse Builder - create database user..... 43
Figure 75 Oracle Warehouse Builder - check user as target schema..... 44
Figure 76 Oracle Warehouse Builder - target user password..... 44
Figure 77 Oracle Warehouse Builder - check user summary..... 45
Figure 78 Oracle Warehouse Builder - register users progress..... 45
Figure 79 Oracle Warehouse Builder - Connection Explorer..... 46
Figure 80 Oracle Warehouse Builder - delete Gator_Location..... 46
Figure 81 Oracle Warehouse Builder - create new file..... 48
Figure 82 Oracle Warehouse Builder - Create Module Wizard..... 49
Figure 83 Oracle Warehouse Builder - name and description..... 49
Figure 84 Oracle Warehouse Builder - connection information..... 50
Figure 85 Oracle Warehouse Builder - select location path..... 50
Figure 86 Oracle Warehouse Builder - finish connection information..... 51

Figure 87 Oracle Warehouse Builder - finishing module creation 51
Figure 88 Oracle Warehouse Builder - importing files 52
Figure 89 Oracle Warehouse Builder - Import Metadata Wizard 52
Figure 90 Oracle Warehouse Builder - filter information 53
Figure 91 Oracle Warehouse Builder - select the right object for the module 53
Figure 92 Oracle Warehouse Builder - run the sample 54
Figure 93 Oracle Warehouse Builder - Flat File Sample Wizard 54
Figure 94 Oracle Warehouse Builder - check the file name 55
Figure 95 Oracle Warehouse Builder - report organization 55
Figure 96 Oracle Warehouse Builder - file format 56
Figure 97 Oracle Warehouse Builder - file layout 56
Figure 98 Oracle Warehouse Builder - field properties 57
Figure 99 Oracle Warehouse Builder - updated field properties 57
Figure 100 Oracle Warehouse Builder – flat file summary 58
Figure 101 Oracle Warehouse Builder - Import Data Wizard summary 58
Figure 102 Oracle Warehouse Builder -export warehouse builder metadata 59
Figure 103 Oracle Warehouse Builder - metadata export window 60
Figure 104 Oracle Warehouse Builder - metadata import 60
Figure 105 Oracle Warehouse Builder - edit database location 61
Figure 106 Oracle Warehouse Builder - file locations in the control center 62
Figure 107 Oracle Warehouse Builder - edit *Connection Explorer* 63
Figure 108 Oracle Warehouse Builder - edit *file system location* 64
Figure 109 Oracle Warehouse Builder - edit process flow and schedule location 65
Figure 110 Oracle Warehouse Builder - updated Control Center 65
Figure 111 Oracle Warehouse Builder - external table deployment 66
Figure 112 Oracle Warehouse Builder - Dimension Table Deployment 67
Figure 113 Oracle Warehouse Builder - fact table deployment 67
Figure 114 Oracle Warehouse Builder - mappings deployment 68
Figure 115 Oracle Warehouse Builder - loading dimension table data 69
Figure 116 Oracle Warehouse Builder - loading fact table data 70
Figure 117 STEWARD web installation - Web Configuration Assistant 71
Figure 118 STEWARD web installation - Oracle net configuration (Step 1) 72
Figure 119 STEWARD web installation - Oracle net configuration (Step 2) 72
Figure 120 STEWARD web installation - Oracle net configuration(Step 3) 73
Figure 121 STEWARD web installation - Oracle net configuration(Step 4) 73
Figure 122 STEWARD web installation - Oracle net configuration(Step 5) 74
Figure 123 STEWARD web installation - Oracle net configuration(Step 6) 74
Figure 124 STEWARD web installation - Oracle net configuration(Step 7) 75
Figure 125 STEWARD web installation - Oracle net configuration(Step 8) 75
Figure 126 STEWARD web installation - XP firewall setting 1 76
Figure 127 STEWARD web installation - XP firewall setting 2 77
Figure 128 STEWARD web installation - XP firewall setting 3 77
Figure 129 STEWARD web installation - file sharing permission 1 78
Figure 130 STEWARD web installation - File sharing permission 2 79
Figure 131 STEWARD web installation - File sharing permission 3 79
Figure 132 STEWARD system architecture 81

Figure 133 Mapping modules for STEWARD 84
Figure 134 Process flow modules for STEWARD 85
Figure 135 Oracle schedule sample for STEWARD 86
Figure 136 Oracle Warehouse Builder Control Center 87
Figure 137 Refresh settings for materialized views 88
Figure 138 Creating STEWARD materialized views 89
Figure 139 STEWARD database backup reports 90
Figure 140 STEWARD warehouse builder metadata export 91

1 Introduction

A description of the steps required to install the Statewide Traffic Engineering Warehouse for Regionally Archived Data (STEWARD) software and databases is presented in this document as an appendix to the STEWARD Final Report for Phase 2. This material has been developed to assist FDOT and ITS contract personnel in setting up STEWARD in the Traffic Engineering Research Laboratory (TERL) facility. The topics include Oracle database program installation, STEWARD deployment and the STEWARD web site installation. A working knowledge of the Oracle data base manager and internet site management is assumed in the discussion that follows.

2 Oracle Database Program Installation

Oracle 10gR2 and the Oracle Workflow Server 2.6.4 are required for STEWARD database installation. The Oracle Workflow Server is included in the Oracle Database 10g companion CD. The installation steps are as follows.

2.1 Install Oracle 10*g* Release 2

Oracle 10*g* Release 2 (10.2) for Microsoft Windows Enterprise edition (32-Bit) is installed for STEWARD.

2.1.1 Run the Setup.exe File

This is a standard Windows procedure

2.1.2 Select Advanced Installation

As shown in Figure 1, select *Advanced Installation* for the installation method.

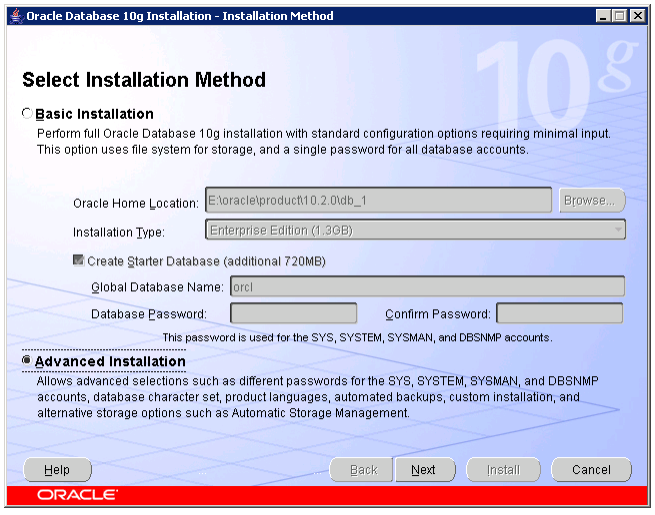


Figure 1 Oracle DB installation - installation method

2.1.3 Select Enterprise Edition

As shown in Figure 2, select Enterprise Edition for the installation type.

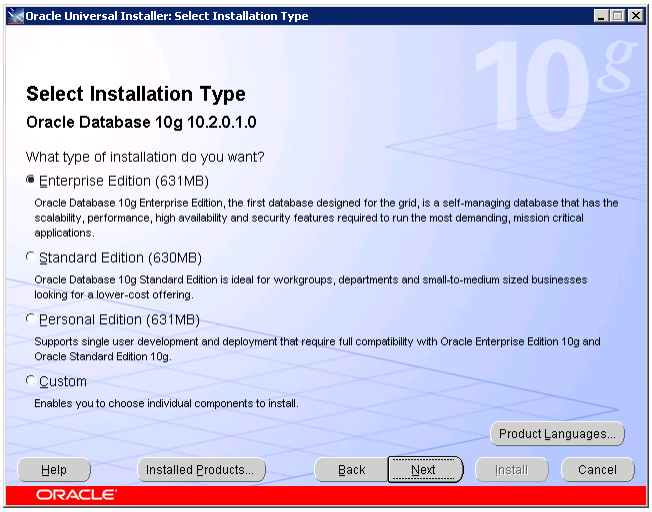


Figure 2 Oracle DB installation - Installation type

2.1.4 Select Default Location

As shown in Figure 3, select the home name and path.
Select the name as OraDb10g_home1 and the path as C:\oracle\product\10.2.0\db_1

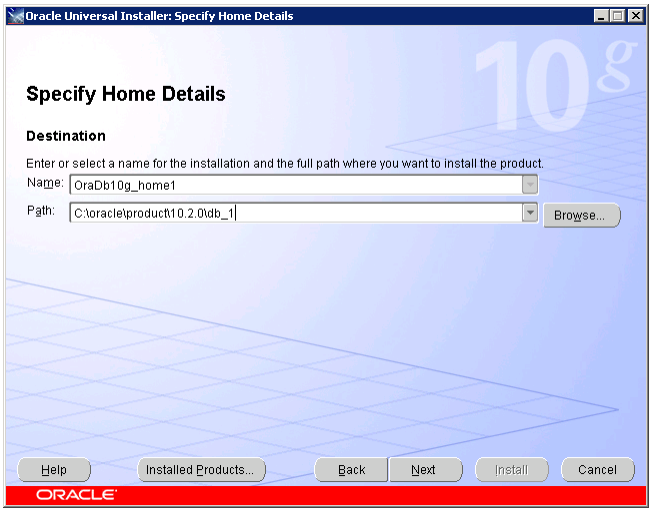


Figure 3 Oracle DB installation - installation home details

2.1.5 Run the Prerequisite Checks

As shown in Figure 4, *prerequisite* will be checked.

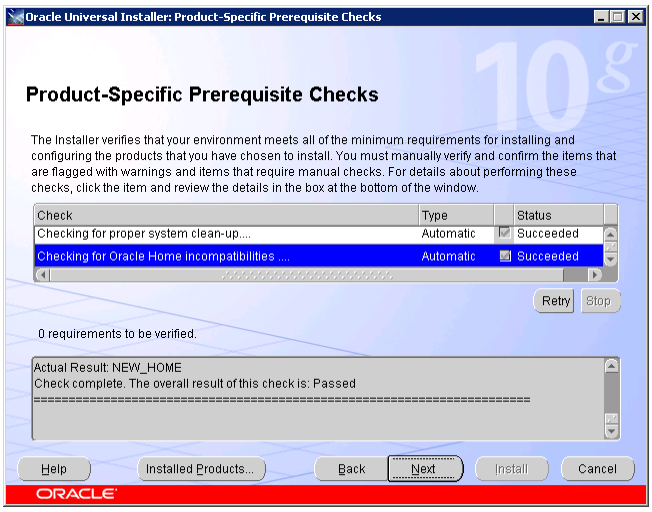


Figure 4 Oracle DB installation - prerequisite checks

2.1.6 Select Database Software Only Option for Configuration Option

As shown in Figure 5, select the *Database Software Only* option.

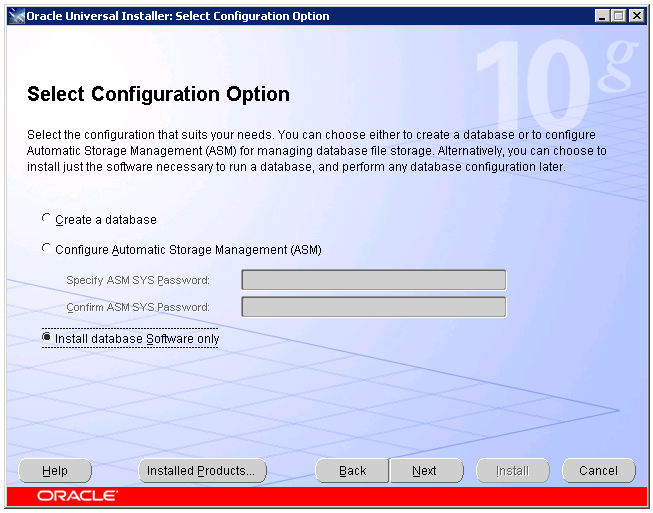


Figure 5 Oracle DB installation - configuration option

2.1.7 Oracle Database Installation

As shown in Figure 6, the Oracle Database installation summary is displayed before installation.

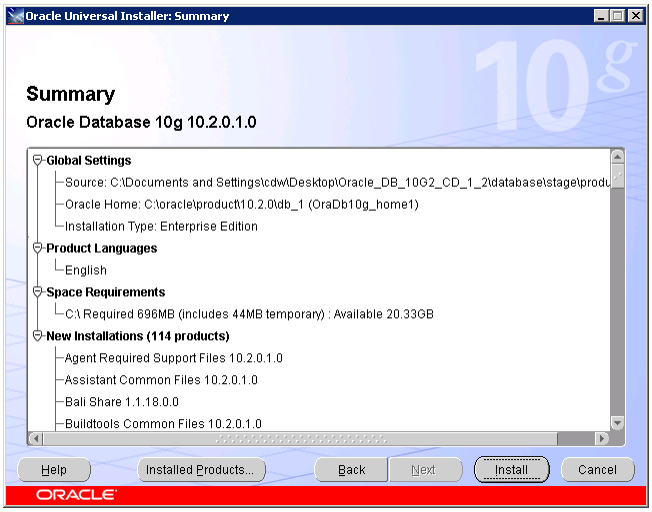


Figure 6 Oracle DB installation - installation Summary

As shown in Figure 7, the Oracle database is installed.

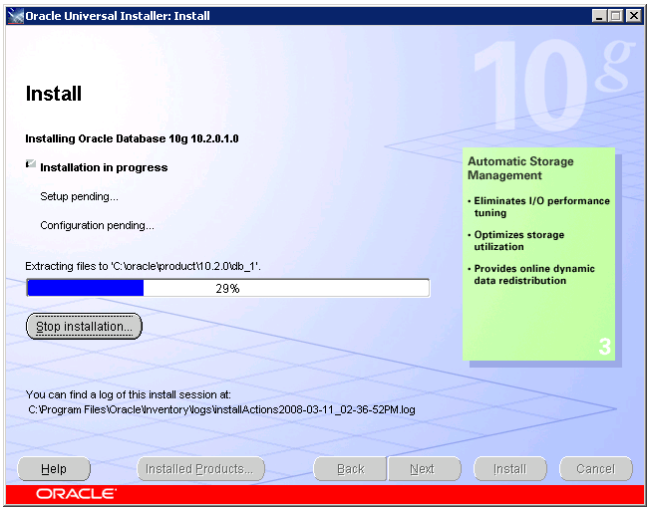


Figure 7 Installation Summary - installation

As shown in Figure 8, the Oracle database installation is done.

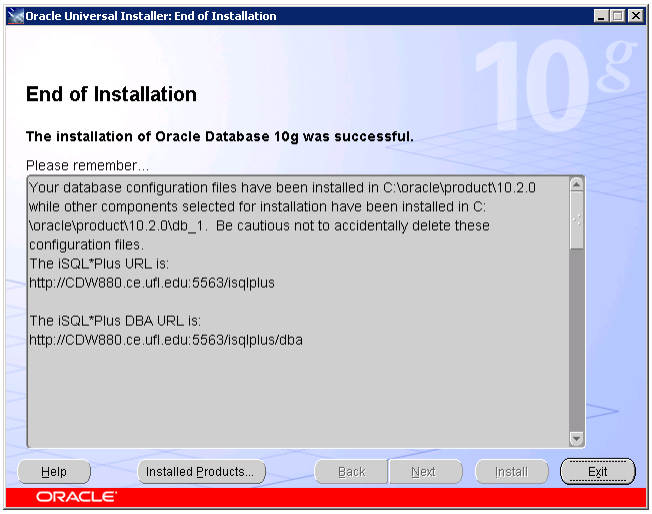


Figure 8 Oracle DB installation - end of installation

2.2 Install Oracle Workflow

Install the Oracle Workflow from the companion CD for Oracle 10*g* Release 2 (10.2) for Microsoft Windows Enterprise Edition (32-Bit)

2.2.1 Run the Setup.exe File

As shown in Figure 9, Oracle Workflow is installed with Oracle Universal Installer.

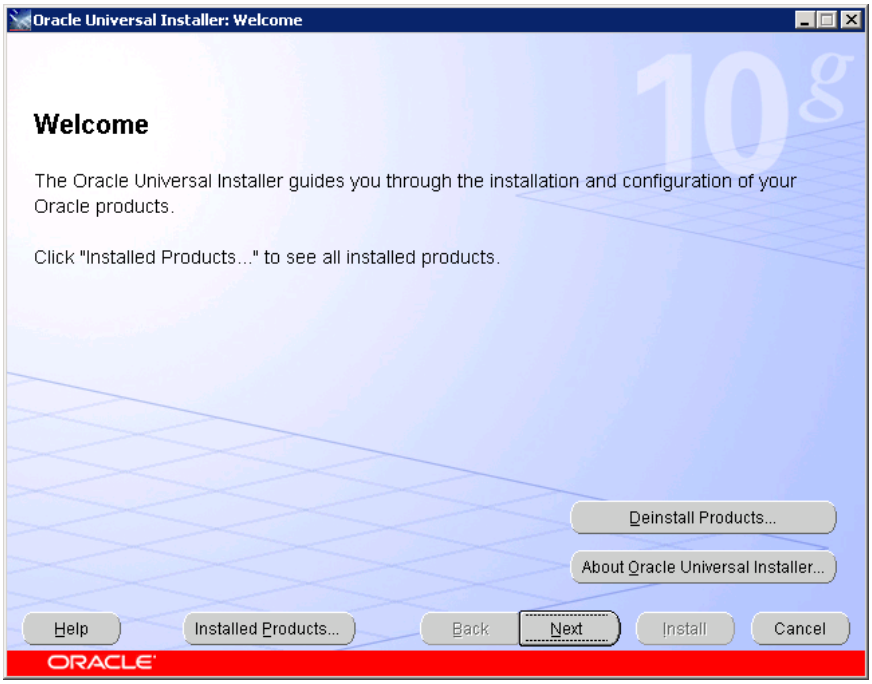


Figure 9 Oracle Workflow installation

2.2.2 Select the Oracle Database 10g Product

As shown in Figure 10, select the Oracle Database 10g Product for Oracle Workflow installation.

Figure 10 Oracle Workflow installation - select a product to install

2.2.3 Select the Home Name and Address

As shown in Figure 11, select the Home name and path.
Select the name as OraDb10g_home1 and the path as "C:\oracle\product\10.2.0\db_1"

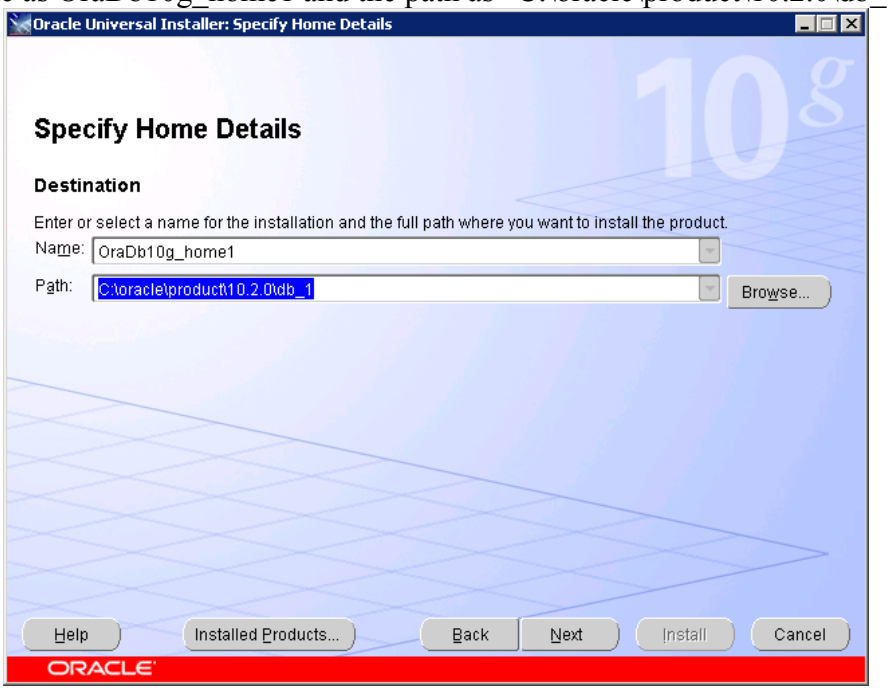


Figure 11 Oracle Workflow installation - installation home details

2.2.4 Run the Prerequisite Checks

As shown in Figure 12, prerequisites will be checked.

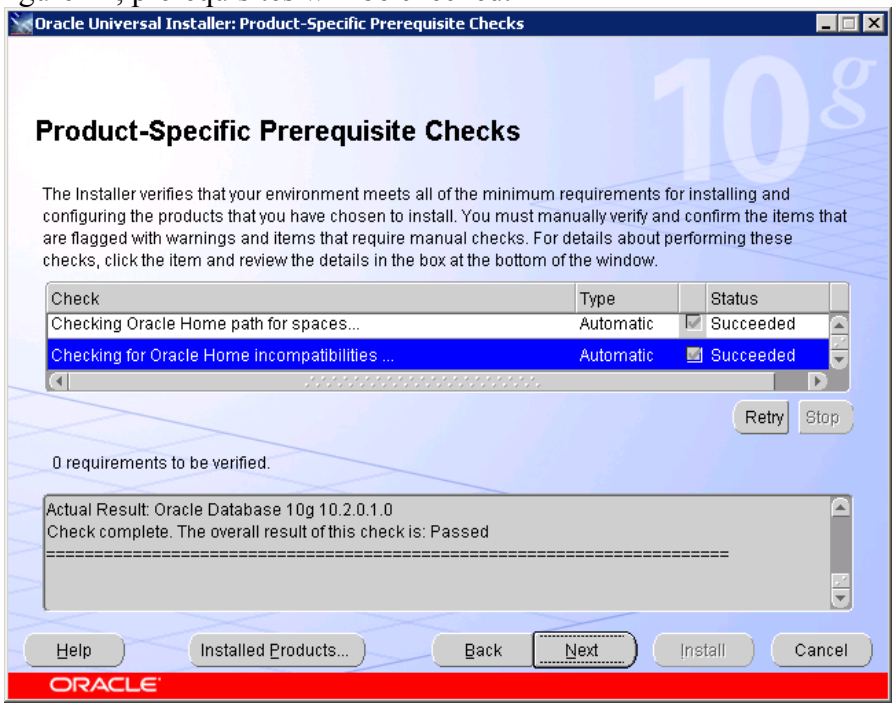


Figure 12 Oracle Workflow installation - prerequisite checks

2.2.5 Oracle Workflow Installation

As shown in Figure 13, the Oracle Workflow installation summary is displayed before installation.

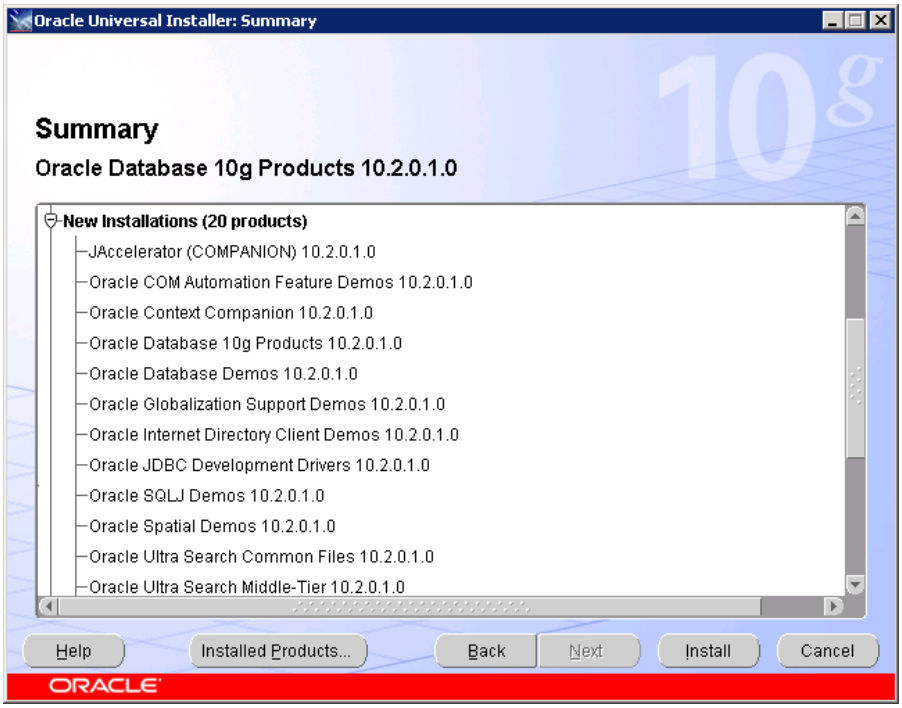


Figure 13 Oracle Workflow installation - installation Summary

2.2.6 Oracle Workflow Installation Completion

As shown in Figure 14, the Oracle Database installation is done.

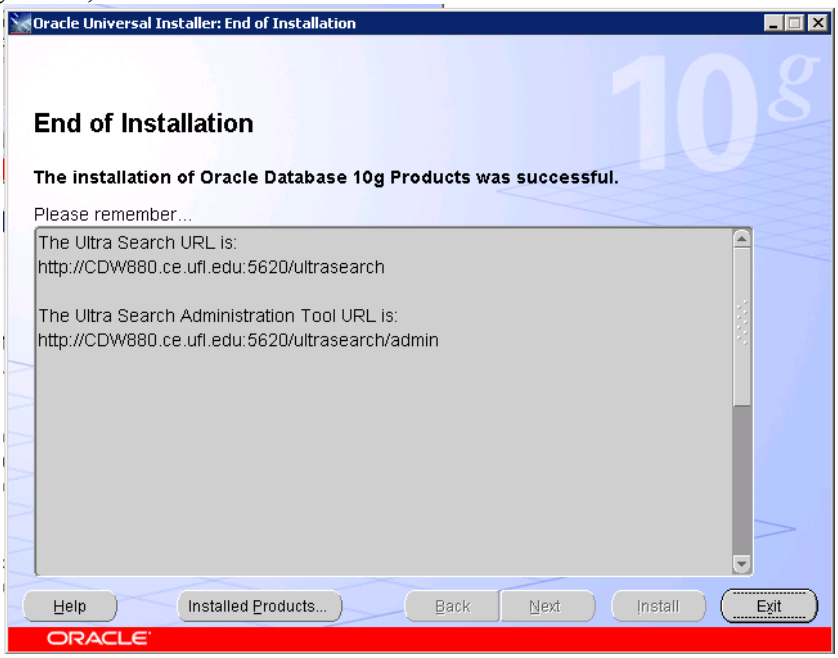


Figure 14 Oracle Workflow installation - end of installation

2.3 Installing the Oracle Warehouse Builder (OWB)

Install the OWB from the Oracle Warehouse Builder 10*g* Release 2 (10.2) for Microsoft Windows Enterprise edition (32-Bit)

2.3.1 Run the Setup.exe File

As shown in Figure 15, the OWBr is installed with the Oracle Universal Installer.

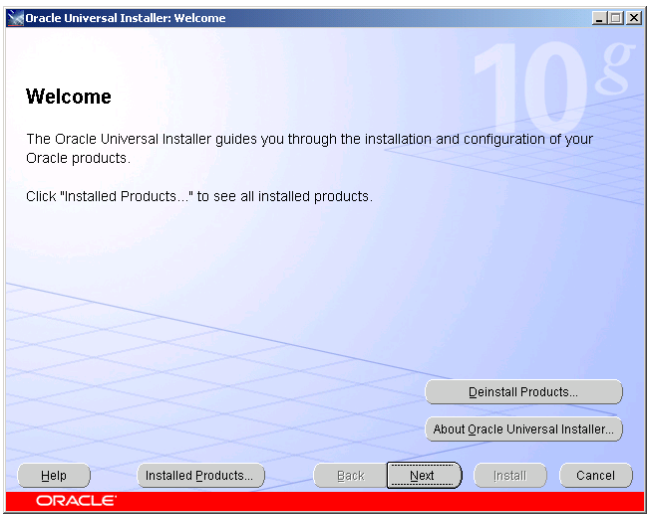


Figure 15 Oracle Warehouse Builder (OWB) installation - welcome

2.3.2 Specify Home Details

As shown in Figure 16, set the home name and path where the OWB will be installed. Select the name as OUIHome1 and the path as "C:\OraHome_1." Then click 'Next.'

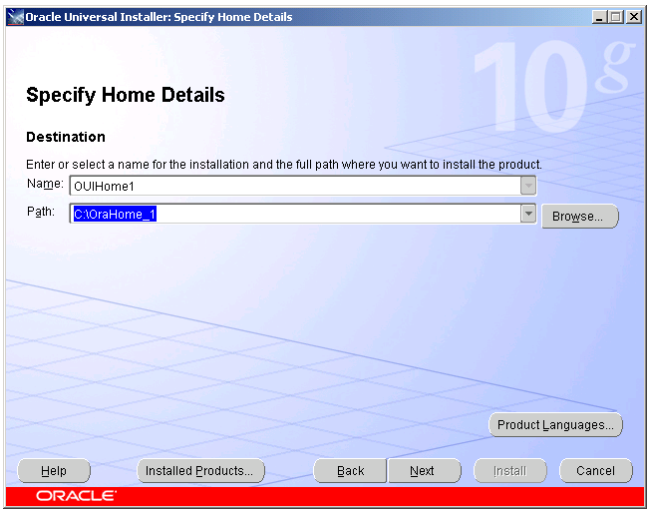


Figure 16 OWB installation - specify home details

2.3.3 Preparation to install the OWB

As shown in Figure 17, the OWB installation summary is displayed before installation. Click the 'install' button.

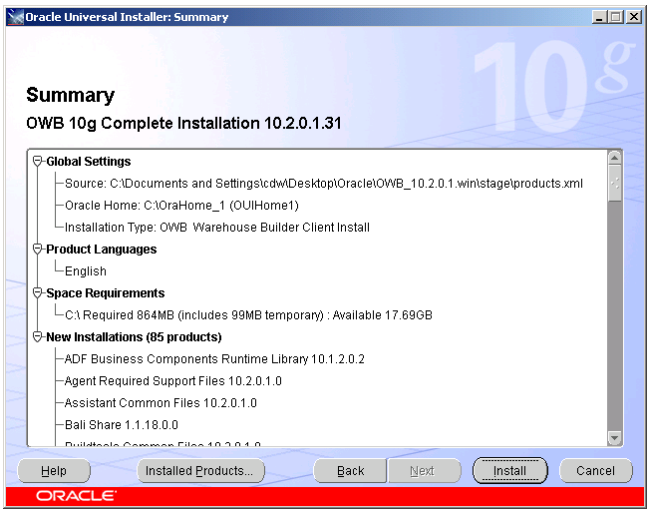


Figure 17 OWB installation - summary information for installation

2.3.4 Install the OWB

As shown in Figure 18, Oracle Warehouse Builder is being installed.

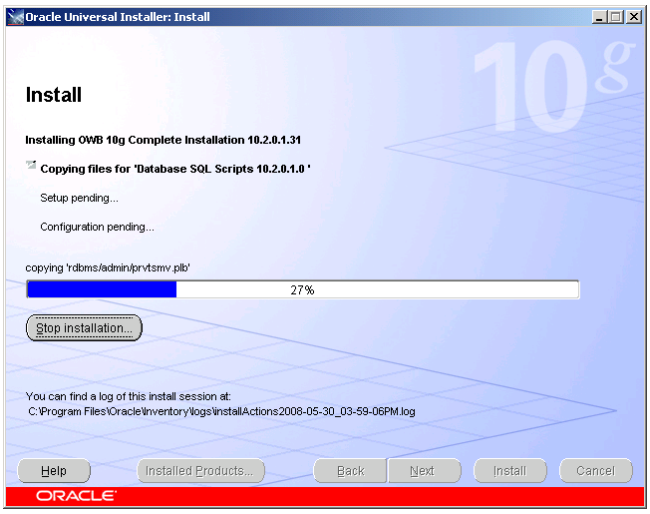


Figure 18 OWB installation

2.3.5 Installation Completed (OWB)

As shown in Figure 19, installation of Oracle Warehouse Builder has been completed.

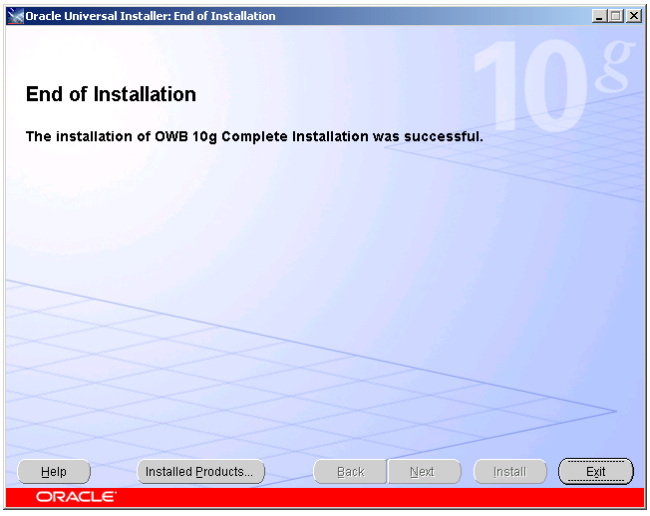


Figure 19 OWB installation - end of installation

2.4 Oracle Database Configuration Assistant

Before setting up the Oracle Warehouse Builder using the Design Center and Oracle Net Configuration (see 2.5), the Database Configuration Assistant should be run. Run the Oracle Database Configuration Assistant (Start → All Programs → Oracle-OraDb10g_home1 → Configuration and Migration Tools → Database Configuration Assistant) as shown in Figure 20.

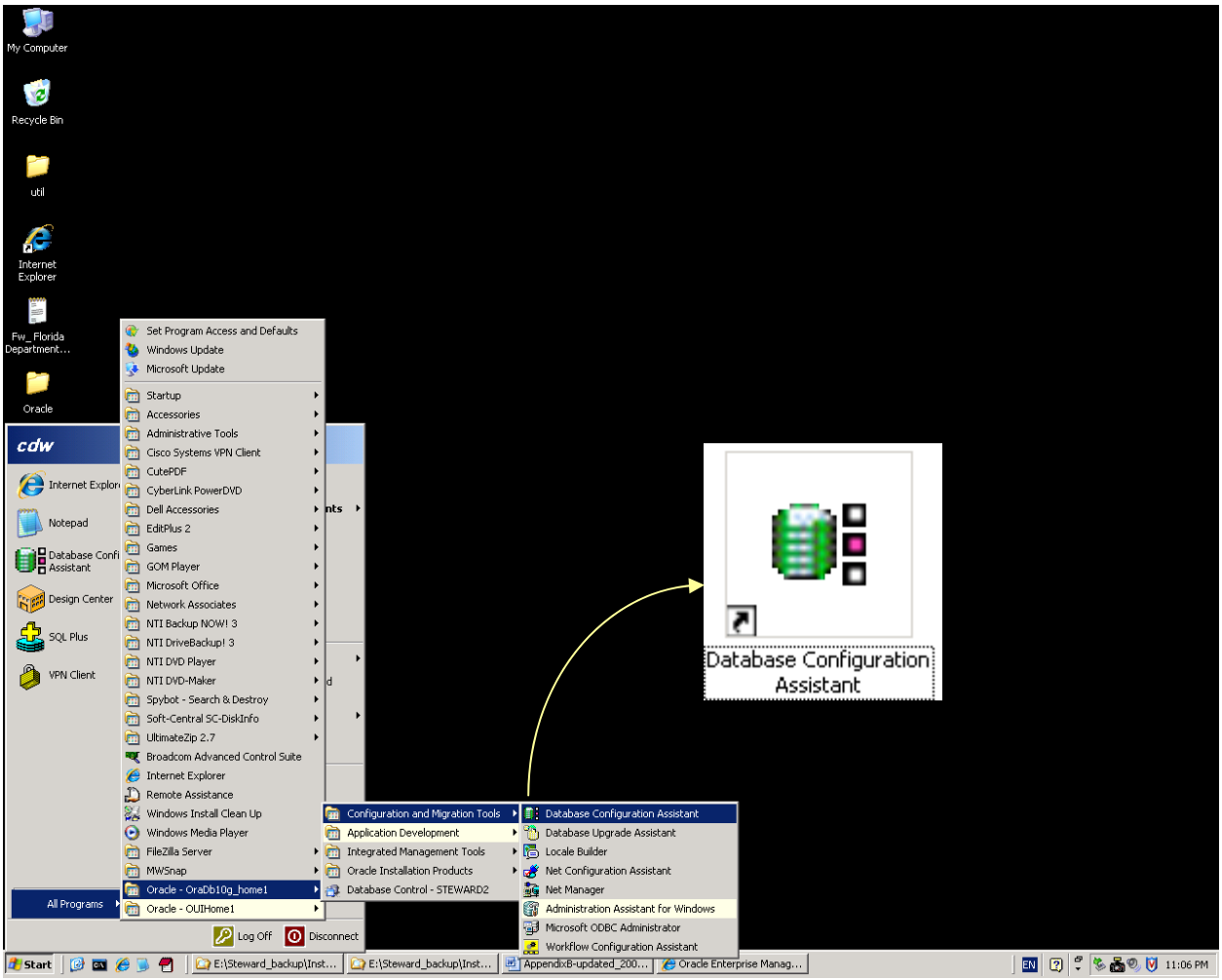


Figure 20 Oracle Database Configuration Assistant - run program and icon

2.4.1 Oracle Database Configuration Assistant -Welcome

As shown in Figure 21, the Database Configuration Assistant shows the welcome message. Click 'Next.'

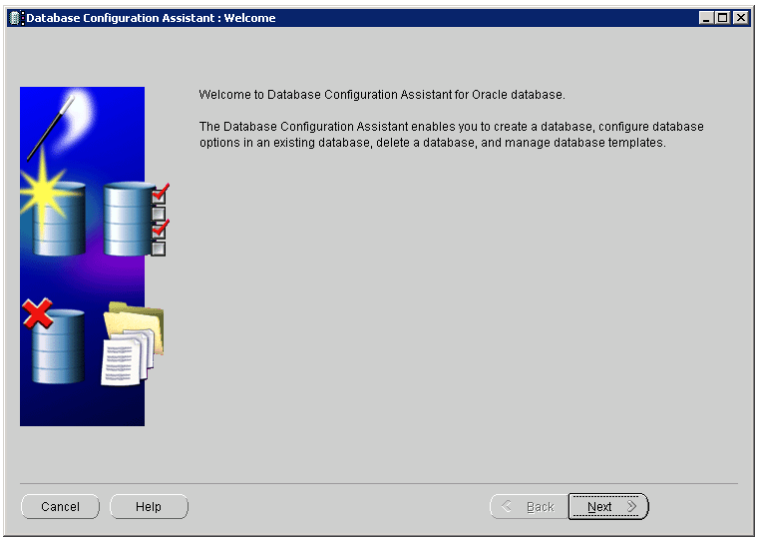


Figure 21 Oracle Database Configuration Assistant - Welcome

2.4.2 Oracle Database Configuration Assistant - Operations

As shown in Figure 22, select 'Create a Database' and click 'Next.'

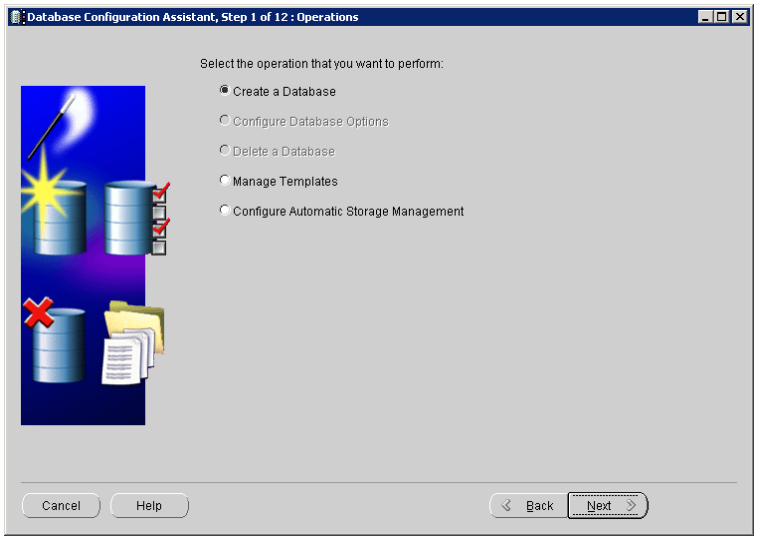


Figure 22 Oracle Database Configuration Assistant - operations

2.4.3 Oracle Database Configuration Assistant -Database Templates

As shown in Figure 23, select 'Data Warehouse' and click 'Next.'

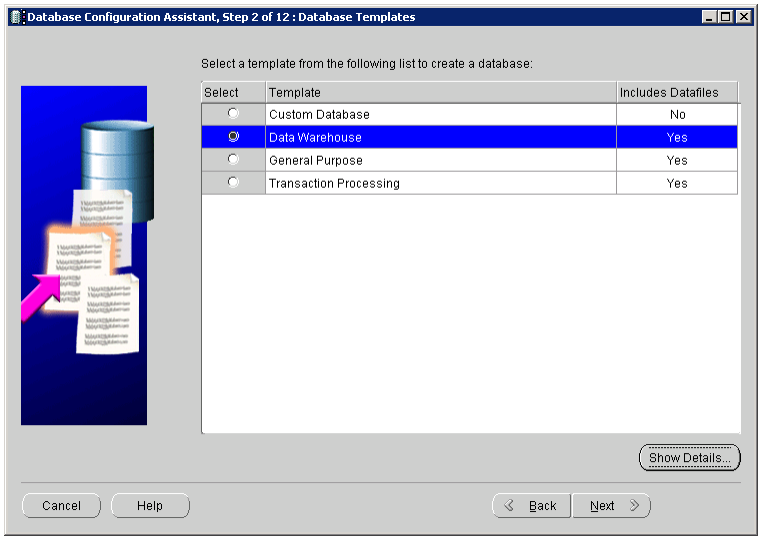


Figure 23 Oracle Database Configuration Assistant - database templates

2.4.4 Oracle Database Configuration Assistant - Database Identification

As shown in Figure 24, enter 'STEWARD' as a global database name. SID will be filled with 'STEWARD' automatically. Click 'Next.'

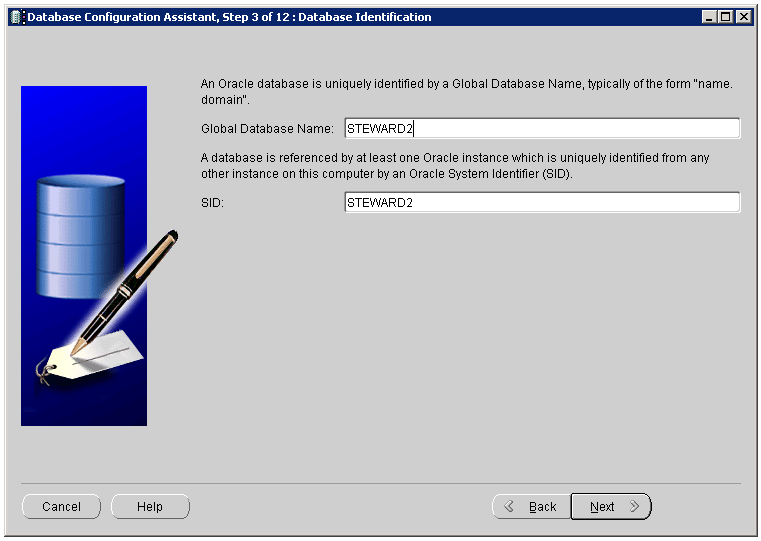


Figure 24 Oracle Database Configuration Assistant - database identification

2.4.5 Oracle Database Configuration Assistant -Management Options

As shown in Figure 25, nothing needs to be changed. Just click 'Next.'

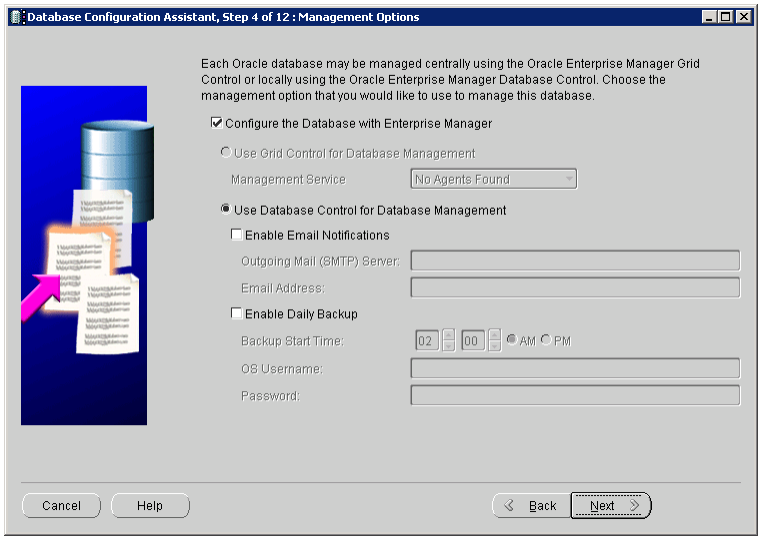


Figure 25 Oracle Database Configuration Assistant - management options

2.4.6 Oracle Database Configuration Assistant - Database Credentials

As shown in Figure 26, enter your password. 'trc513' is used in this example. Click 'Next.'

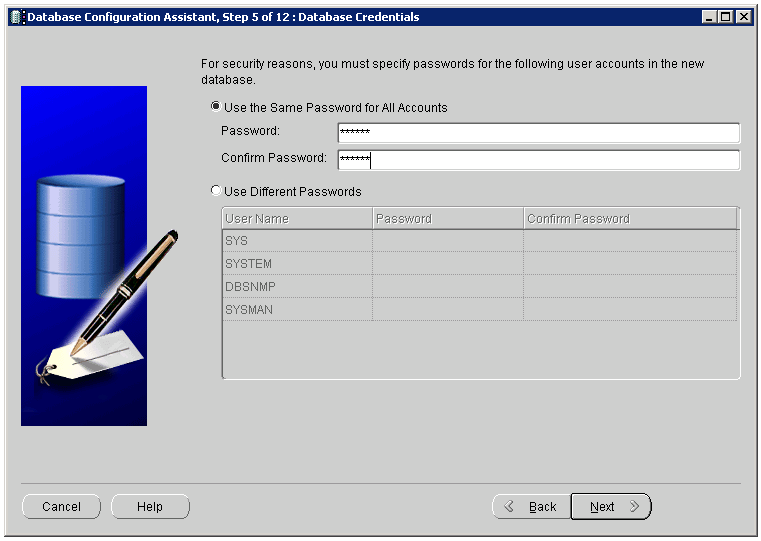


Figure 26 Oracle Database Configuration Assistant - database credentials

2.4.7 Oracle Database Configuration Assistant -Storage Options

As shown in Figure 27, select 'File System' as a default. Click 'Next.'

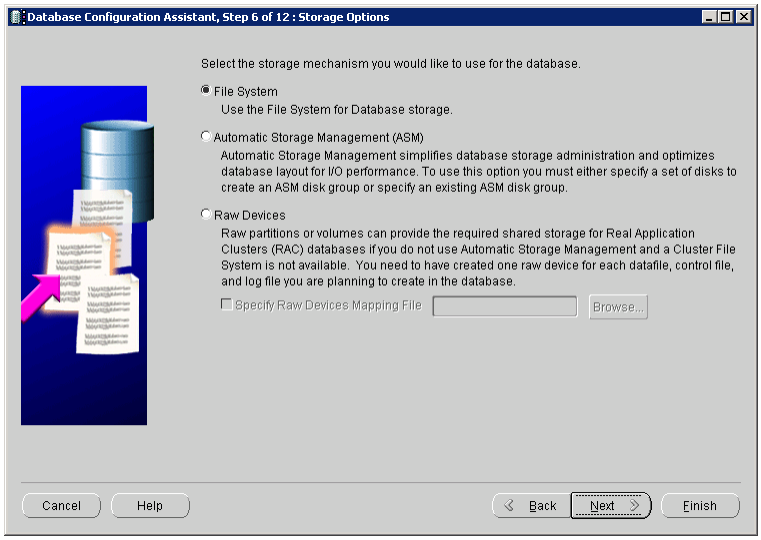


Figure 27 Oracle Database Configuration Assistant - storage options

2.4.8 Oracle Database Configuration Assistant - Database File Locations

In this step, you may select a default option 'Use Database File Location from Template.' Or if you want to select another location (or directory), as shown in Figure 28, you may select 'Use Common Location for All Database Files.' After selecting 'Use Common Location for All Database Files,' enter 'K:\ORADATA' into the Database Files Locations. Click 'Next.'

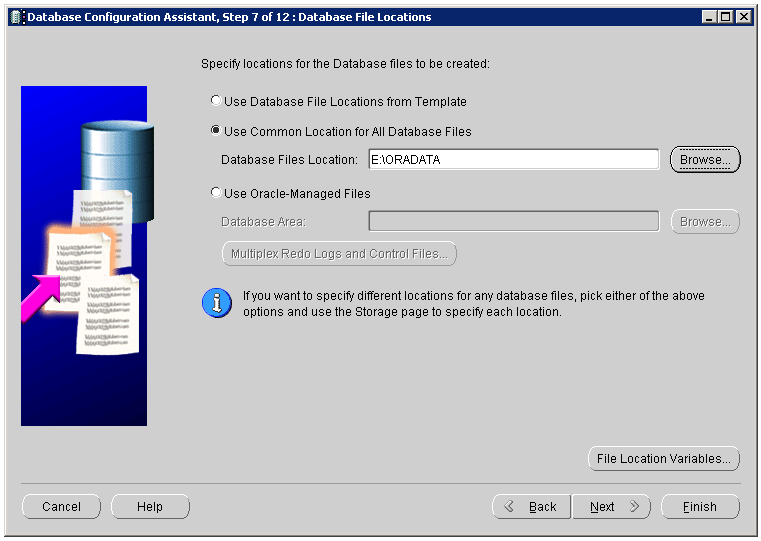


Figure 28 Oracle Database Configuration Assistant - database file locations

2.4.9 Oracle Database Configuration Assistant -Recovery Configuration

As shown in Figure 29, nothing needs to be changed. Just click 'Next.'

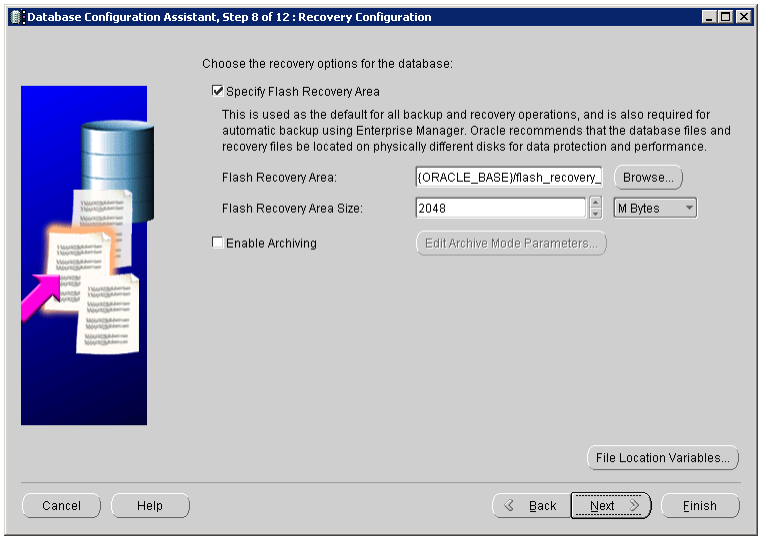


Figure 29 Oracle Database Configuration Assistant - recovery configuration

2.4.10 Oracle Database Configuration Assistant - Database Contents

As shown in Figure 30, nothing needs to be changed. Just click 'Next.'

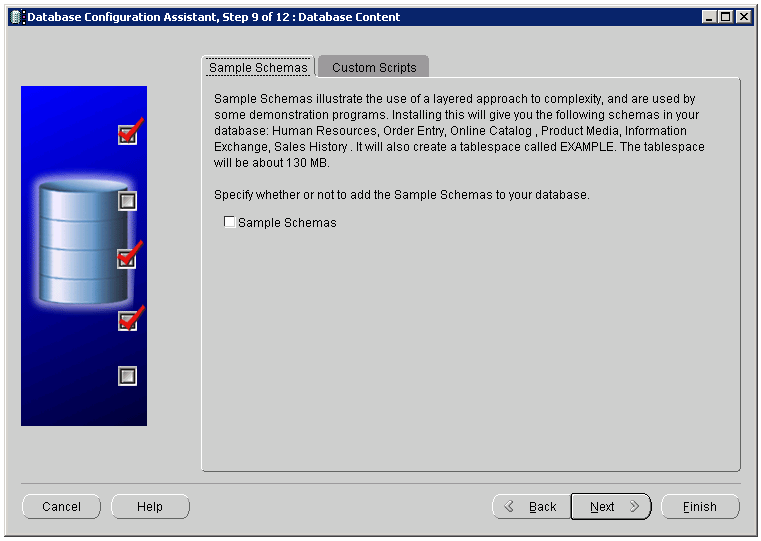


Figure 30 Oracle Database Configuration Assistant - database contents

2.4.11 Oracle Database Configuration Assistant - Initialization Parameters

As shown in Figure 31, nothing needs to be changed. Just click 'Next.'

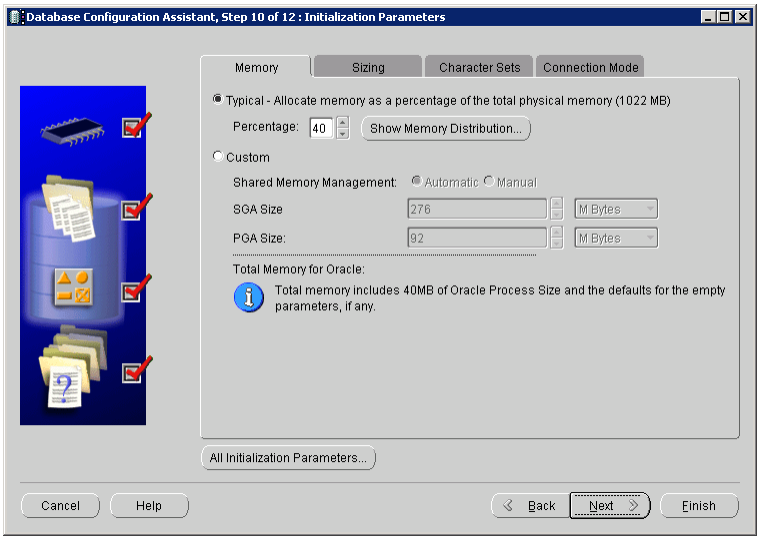


Figure 31 Oracle Database Configuration Assistant - initialization parameters

2.4.12 Oracle Database Configuration Assistant - Database Storage

As shown in Figure 32, The 'Datafiles' subfolder is set by the 'K:\ORADATA\{DB_NAME}\.' After checking it, click 'Next.'

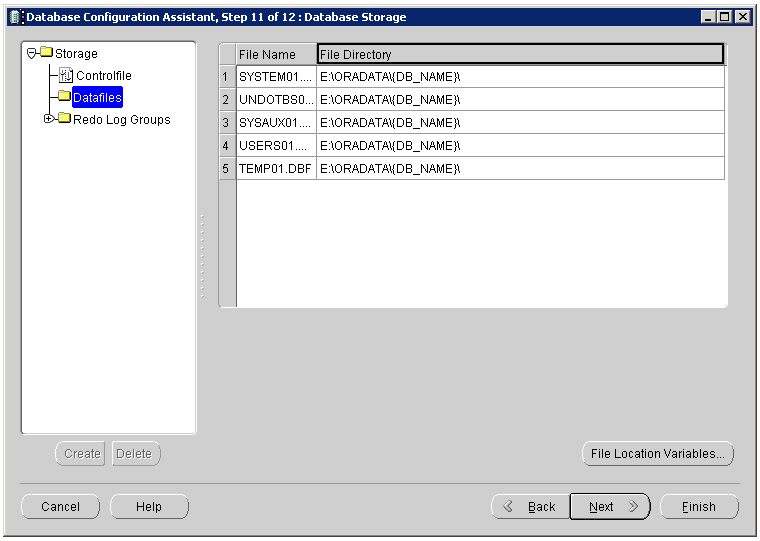


Figure 32 Oracle Database Configuration Assistant - database storage

2.4.13 Oracle Database Configuration Assistant -Creation Options

As shown in Figure 33, nothing needs to be changed. Just click 'Finish.'

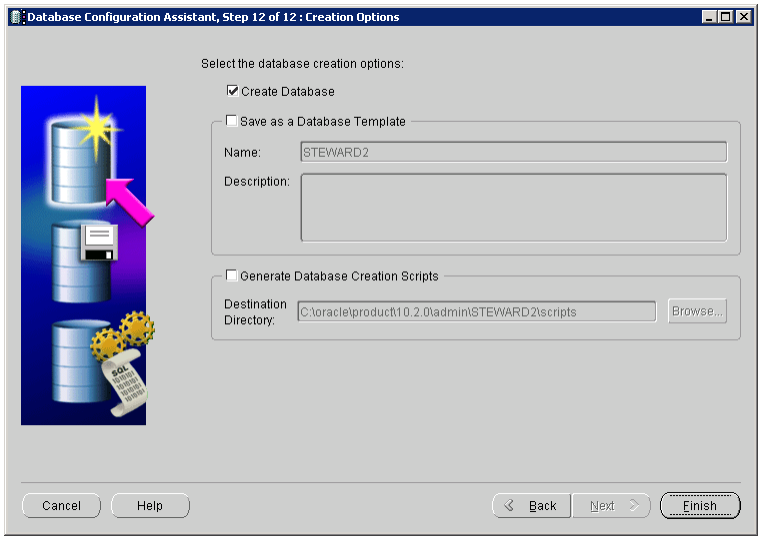


Figure 33 Oracle Database Configuration Assistant - creation options

2.4.14 Oracle Database Configuration Assistant - Confirmation

As shown in Figure 34, after checking the summary of details, click 'OK.'

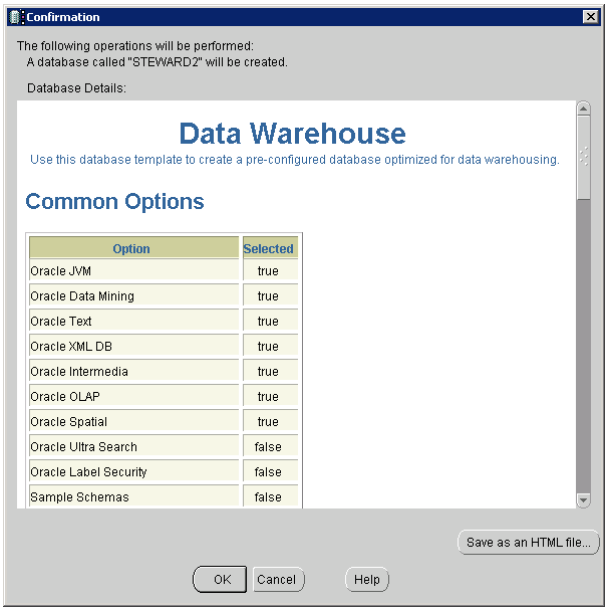


Figure 34 Oracle Database Configuration Assistant - confirmation

2.4.15 Oracle Database Configuration Assistant -Creating Database

As shown in Figure 35, the database is being created.

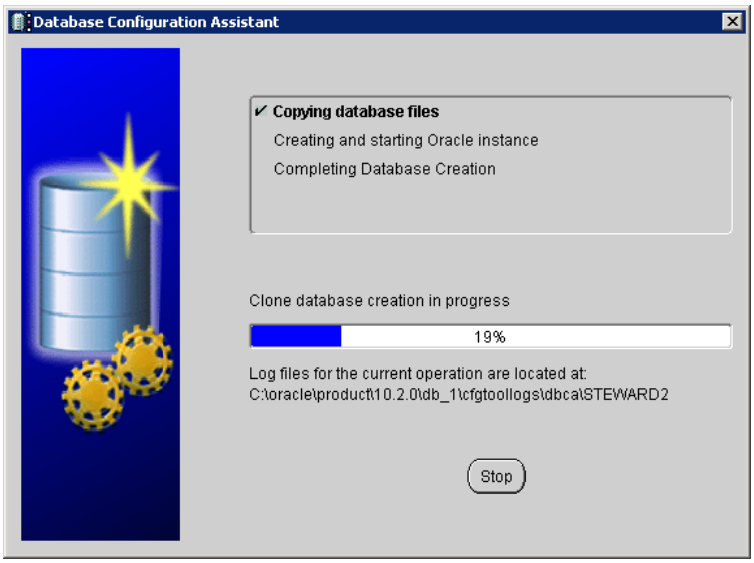


Figure 35 Oracle Database Configuration Assistant - creating the database

2.4.16 Oracle Database Configuration Assistant - End of Creating Database

As shown in Figure 36, the end of creating database is displayed.

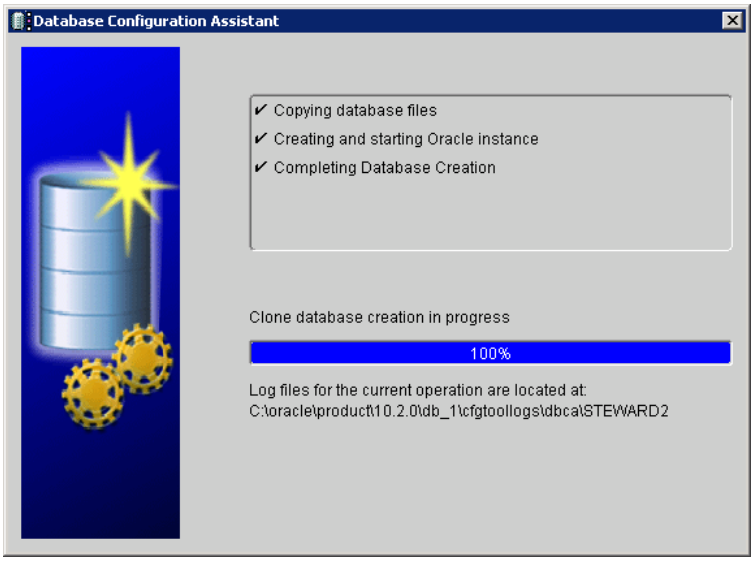


Figure 36 Oracle Database Configuration Assistant - end of creating database

2.4.17 Oracle Database Configuration Assistant -Password Management

As shown in Figure 37, the database configuration needs to be set.
The Database Control URL is 'http://cdw880.ce.ufl.edu:1158/em'.
When you go to the above URL, the username 'sys' and password 'trc513' will be used.

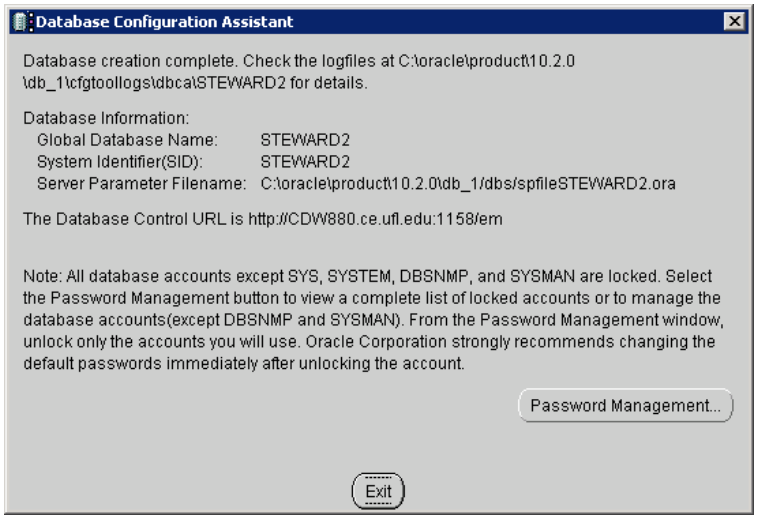


Figure 37 Oracle Database Configuration Assistant - password management

2.5 Oracle Net Configuration Assistant

Before setup the Oracle Warehouse Builder using the Design Center, the Oracle Net Configuration should be checked to ensure that it is working. Run the Oracle Net Configuration Assistant (Start → All Programs → Oracle-OraDb10g_home1 → Configuration and Migration Tools → Net Configuration Assistant) as shown in Figure 38.

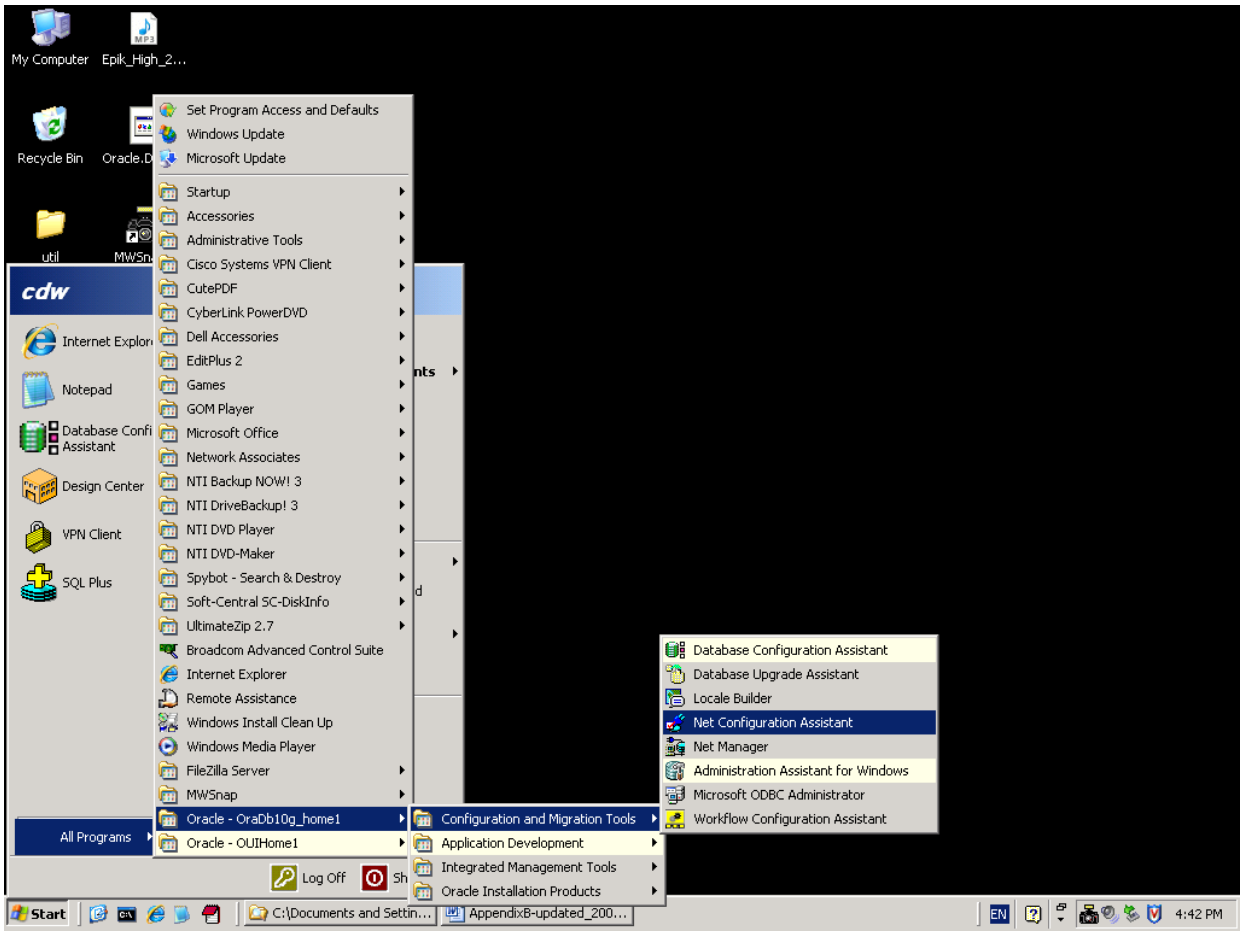


Figure 38 Oracle Net Configuration Assistant - run the 'Net Configuration Assistant'

2.5.1 Oracle Net Configuration Assistant - Listener Configuration

As shown in Figure 39, select the 'Listener Configuration' and click 'Next.' Although it could be already setup it should be checked.

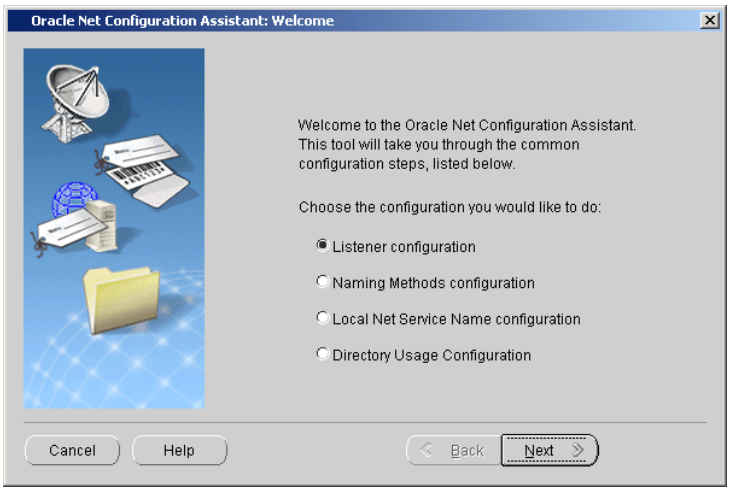


Figure 39 Oracle Net Configuration Assistant - listener configuration

As shown in Figure 40, the 'Add' option is the only one to be selected, which means that there is no running listener on this server. Click 'Next.'

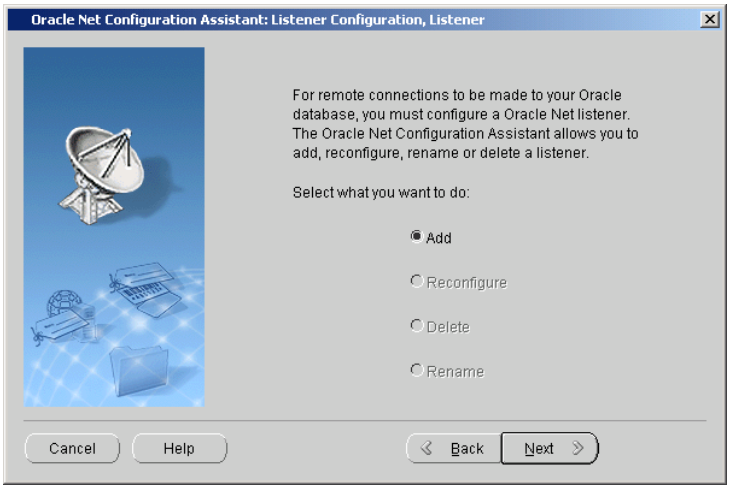


Figure 40 Oracle Net Configuration Assistant - listener configuration

As shown in Figure 41, set the listener name for remote connection to your Oracle database. You may enter the name of any listener you want but there is no problem if you use the default name 'STEWARD.' After setting the name, click 'Next.'

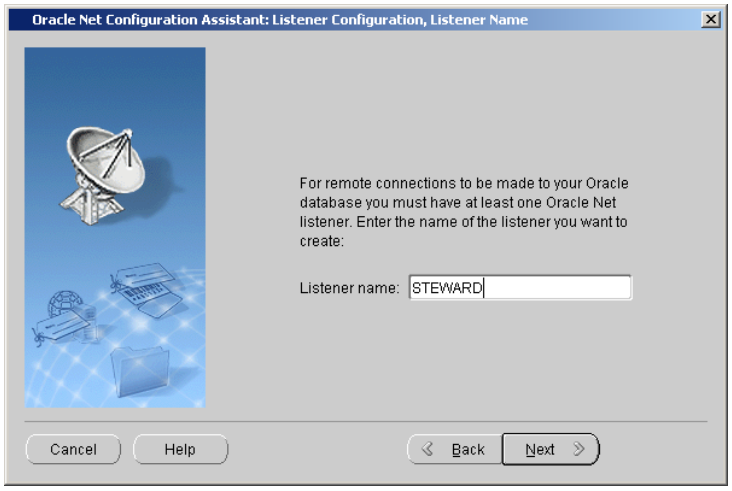


Figure 41 Oracle Net Configuration Assistant - listener name

As shown in Figure 42, select the protocols to accept connection. You may change or select the protocols if you want to but there will be no problem if you use the default protocol 'TCP.' After setting the protocols, click 'Next.'

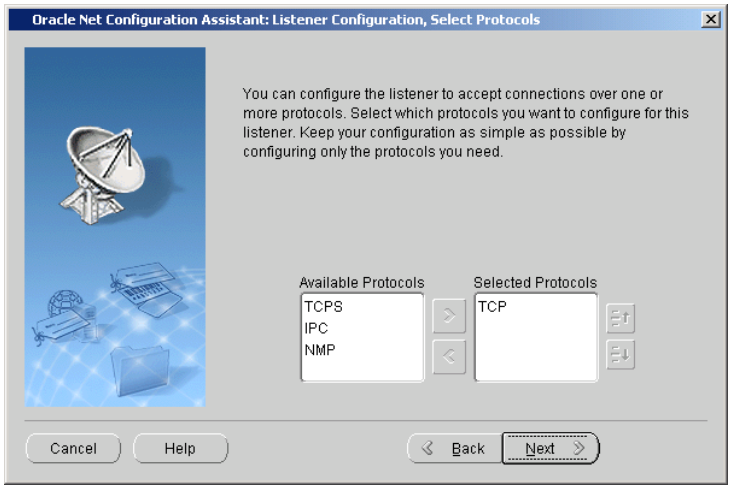


Figure 42 Oracle Net Configuration Assistant - Select Protocols

Under the assumption that the TCP protocol has been selected as shown in Figure 43, the next step is to select the port number for connection. As shown in Figure 43, you can choose the default port number '1521.' Or you can change the port number to whatever you want. After setting the port number, click 'Next.'

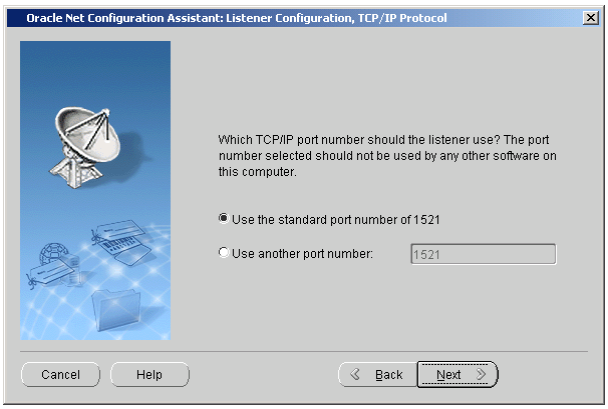


Figure 43 Oracle Net Configuration Assistant - TCP/IP protocols and port number

As shown in Figure 44, the InstallShield asks if there is another listener to be set up. In the current configuration, you don't need to add another listener. Select 'No' and click 'Next.'

Figure 44 Oracle Net Configuration Assistant - listener configuration

As shown in Figure 45, the listener configuration has been completed. Click 'Next' for the next step (Oracle Net Configuration Assistant – Local Net Service Name Configuration).

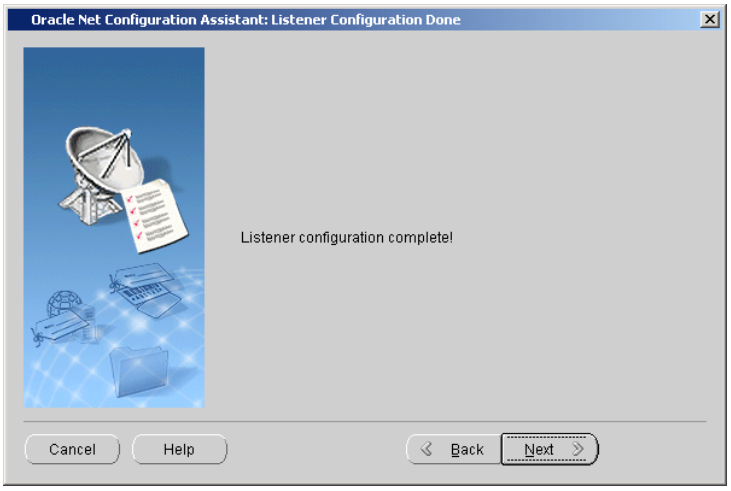


Figure 45 Oracle Net Configuration Assistant -listeners configuration done

2.5.2 Oracle Net Configuration Assistant -Local Net Service Name Configuration

As shown in Figure 46, the local net service name configuration should be processed after completing the setup of the listener configuration. Select the 'Local Net Service Name Configuration' menu and click 'Next.'

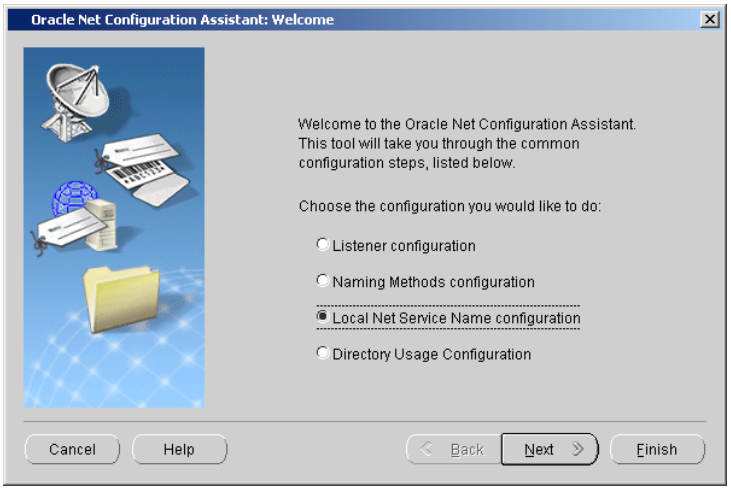


Figure 46 Oracle Net Configuration Assistant -local net service name configuration

As shown in Figure 47, select the 'Add' menu and click 'Next.' (It is assumed that there is no other net service name now.)

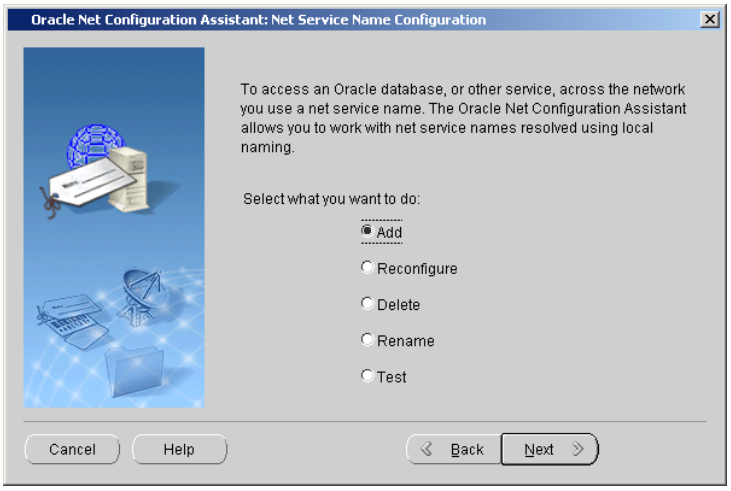


Figure 47 Oracle Net Configuration Assistant -local net service name configuration (Add)

As shown in Figure 48, enter 'STEWARD' as the service name and click 'Next.'

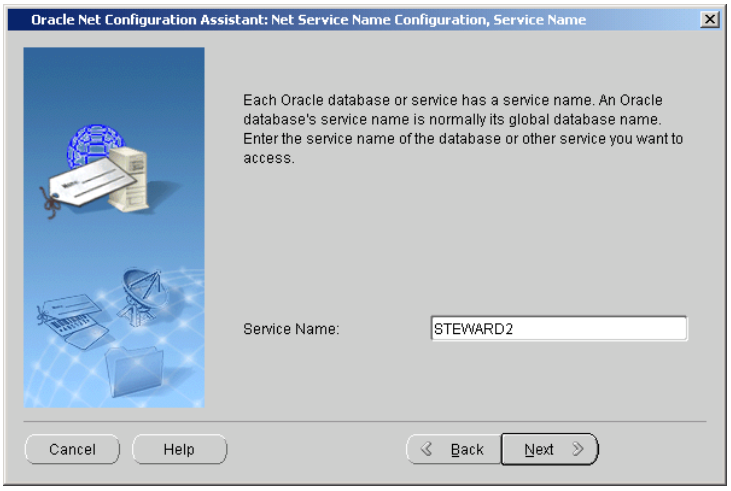


Figure 48 Oracle Net Configuration Assistant -local net service name configuration (Service Name)

As shown in Figure 49, select the protocols for communication. Select the 'TCP' and click 'Next.'

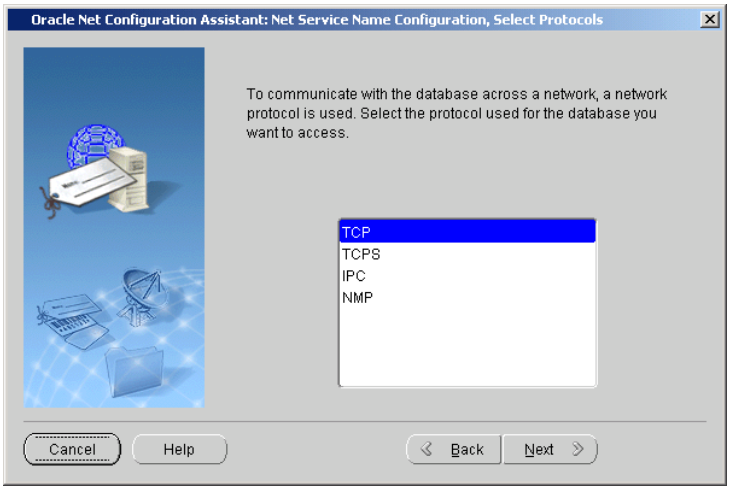


Figure 49 Oracle Net Configuration Assistant -local net service name configuration (Select Protocols)

As shown in Figure 50, enter the host name (e.g. 'cdwserver.ce.ufl.edu') and select the port number. Click 'Next.'

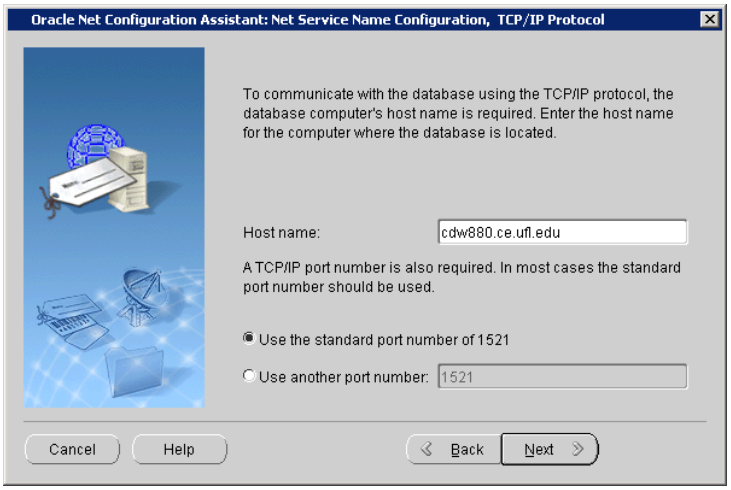


Figure 50 Oracle Net Configuration Assistant -local net service name configuration (TCP/IP Protocols)

As shown in Figure 51, select 'Yes, perform a test' to test the connection. Click 'Next.'

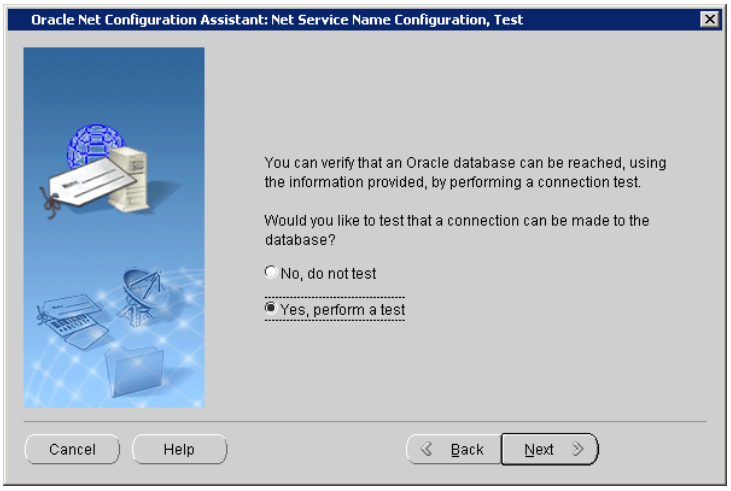


Figure 51 Oracle Net Configuration Assistant -local net service name configuration (Connection Test)

As shown in Figure 52, if the test result is failure, select 'Change Login' to test the connection.

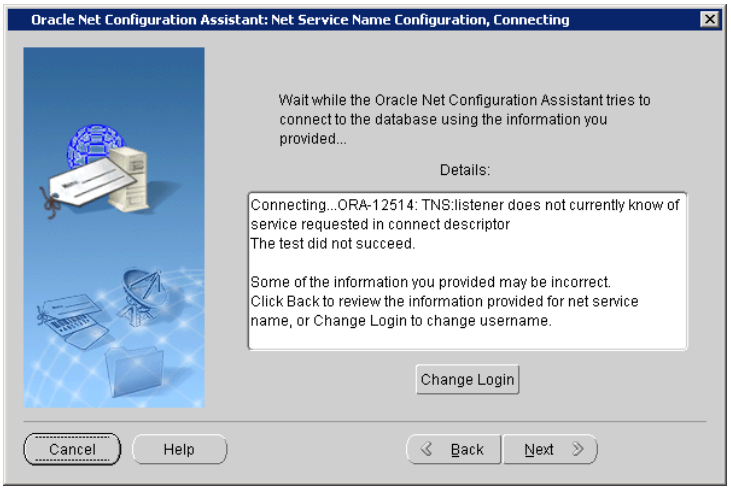


Figure 52 Oracle Net Configuration Assistant -local net service name configuration (Failure)

If the test is a failure, as shown in Figure 53, enter ‘sys(or system)’ as ‘Username’ and ‘trc513’ as the password. Then click ‘OK.’

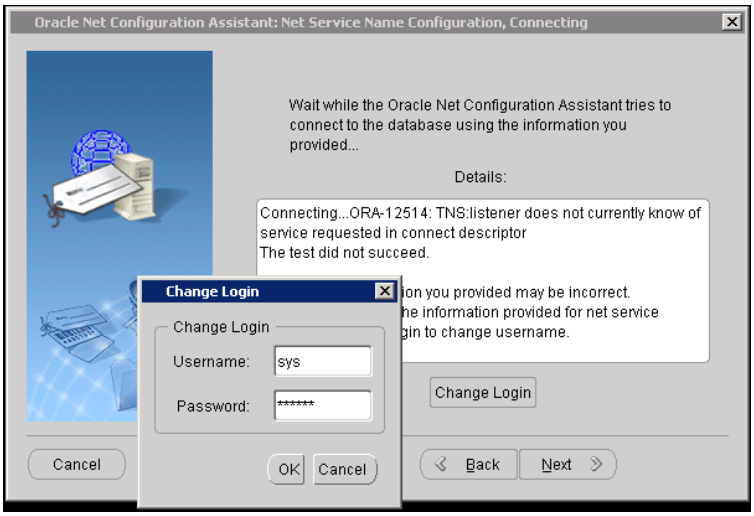


Figure 53 Oracle Net Configuration Assistant -change login

Please check the success of the connection, as shown in Figure 54 Then click ‘Next.’

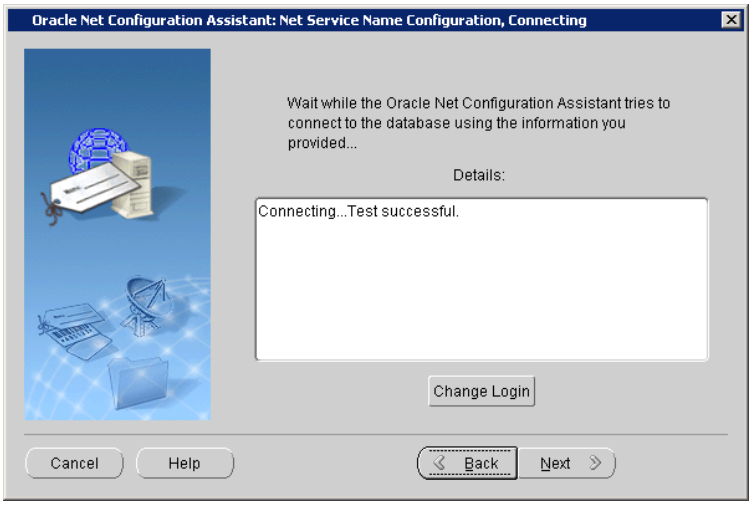


Figure 54 Oracle Net Configuration Assistant -connection succeeded

As shown in Figure 55, enter ‘STEWARD’ as ‘Net Service Name.’ Click ‘Next.’

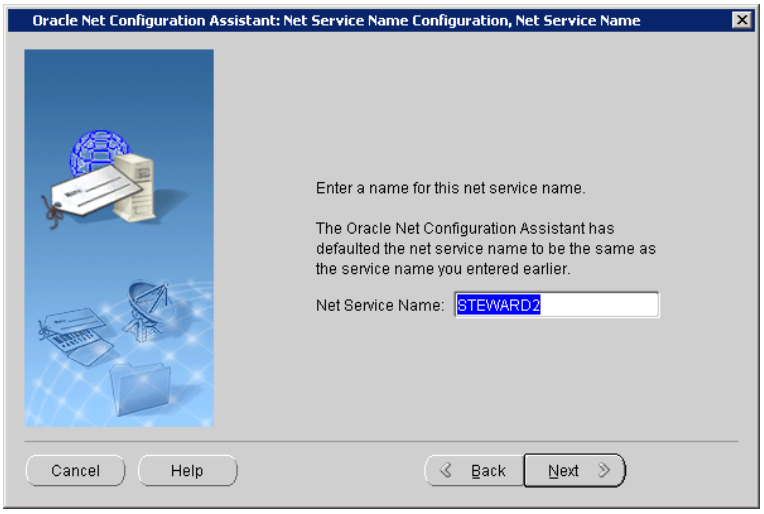


Figure 55 Oracle Net Configuration Assistant -net service name

As shown in Figure 56, select 'No' and click 'Next.'

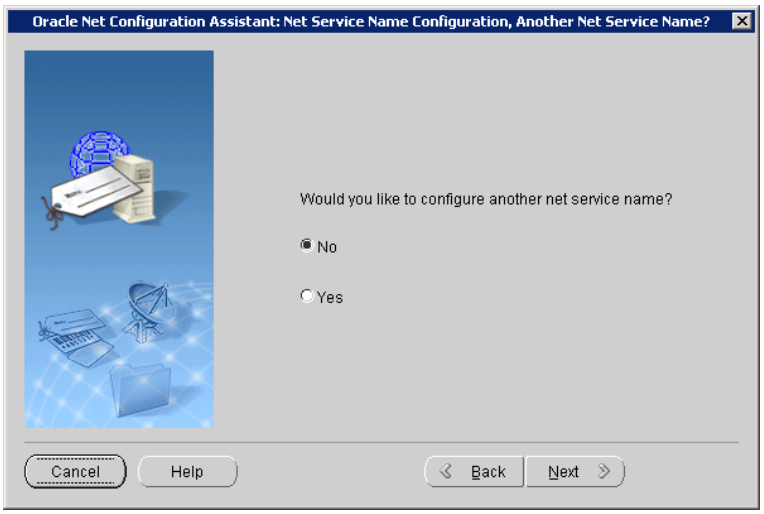


Figure 56 Oracle Net Configuration Assistant - another net service name?

As shown in Figure 57, click 'Next.' And then, just click 'Finish.'

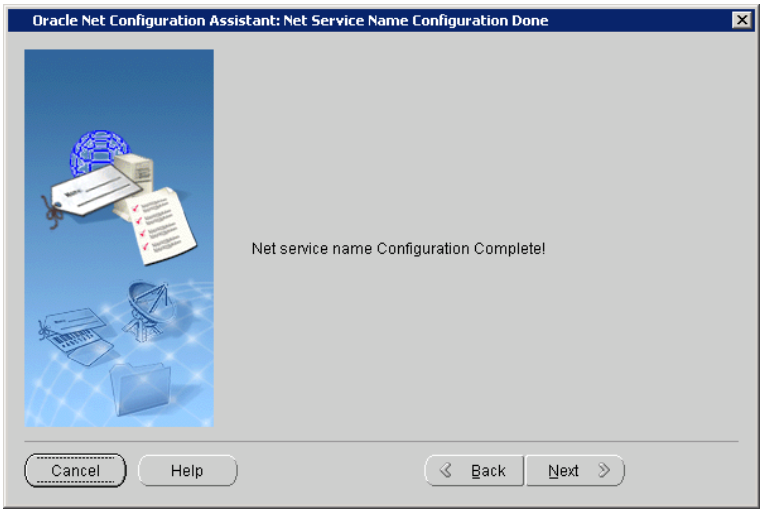


Figure 57 Oracle Net Configuration Assistant -net service name (configuration done)

2.5.3 System Variables Setting

In the Windows system, ORACLE_HOME needs to be defined as a system variable. If not, the following error message might be generated.

"ORA-12154: TNS: could not resolve the connect identifier specified."

From the Windows start menu, run the system properties windows as Figure 58 shows. Then set the system variables as follows.

Start→ Control Panel → System properties
System properties: Advanced → Environmental Variables → System variables

Variable: ORACLE_HOME
Value: C:\oracle\product\10.2.0\db_1

Then the system needs to be restarted.

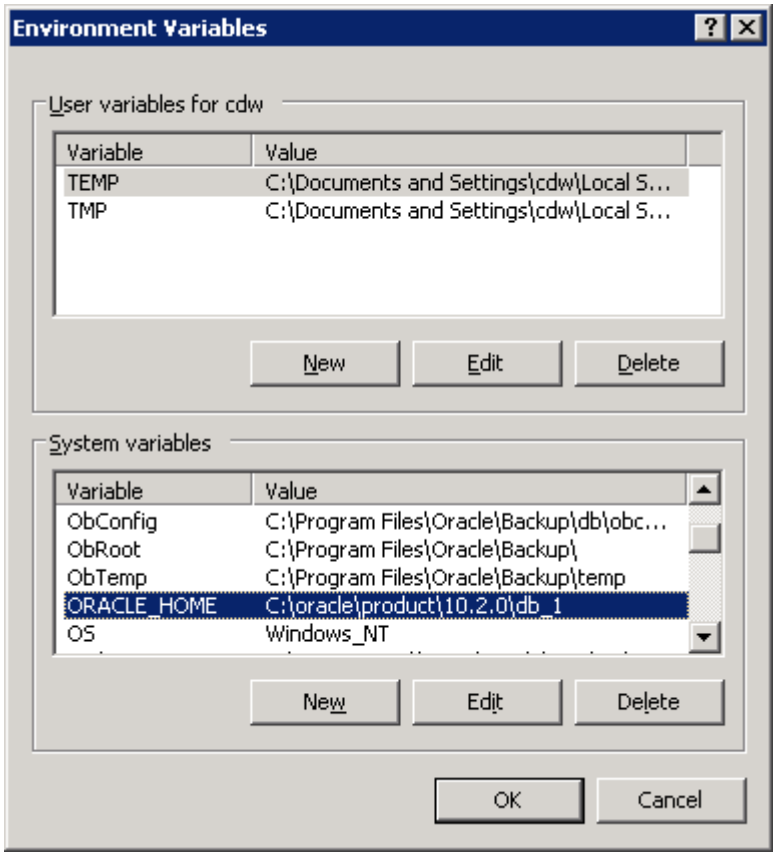


Figure 58 System variables setting

2.6 Oracle Enterprise Manager

2.6.1 Oracle Enterprise Manager - Login

As shown in Figure 59, go to 'http://cdwserver.ce.ufl.edu:1158/em'(or your server host URL) and login (User Name: sys, Password: trc513, Connect As: SYSDBA).

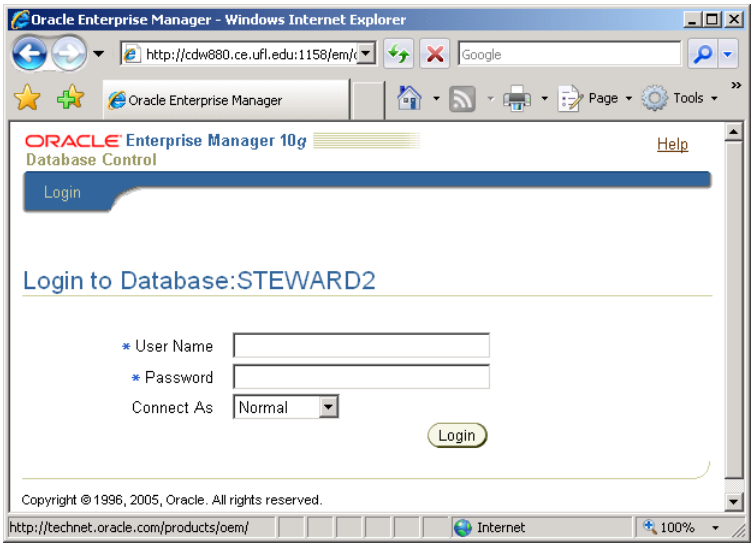


Figure 59 Oracle Enterprise Manager - login

2.6.2 Oracle Enterprise Manager - Database Control Main Page

After login, as shown in Figure 60, you can see the main page of the Oracle Enterprise Manager.

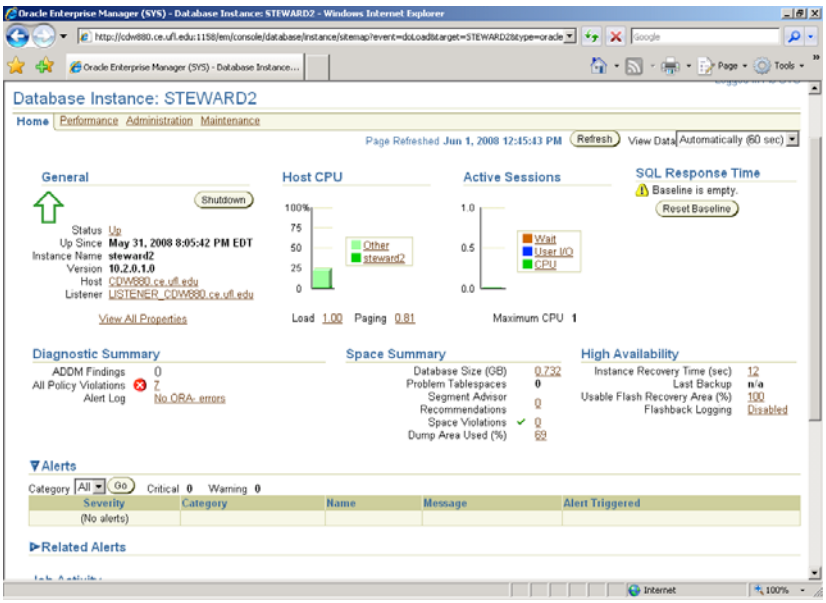


Figure 60 Oracle Enterprise Manager - main page

2.7 Oracle Warehouse Builder Setup

The Design Center in the Oracle Warehouse Builder needs to be configured during the installation. The Design Center is executable as shown in Figure 61.

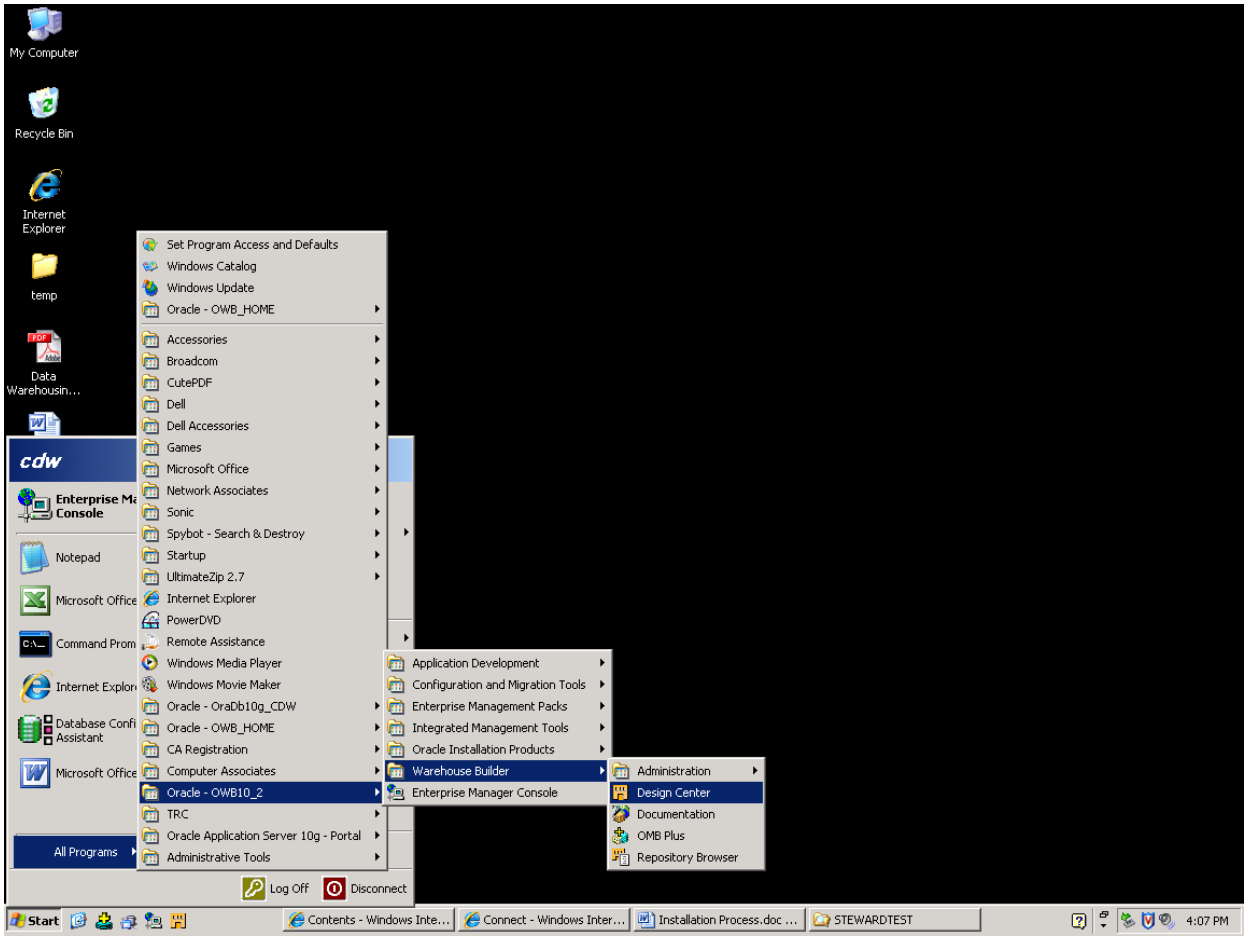


Figure 61 Oracle WarehouseBuilder configuration

2.7.1 Create the Design Center User and its Repository

Run the Design Center and create a "Warehouse Builder User" and "Warehouse Builder Repository." They will be used to archive information for the new data warehouse. As shown in Figure 62, click 'Get Started.' If the *Get Started* button is not visible, click 'Show Details.'

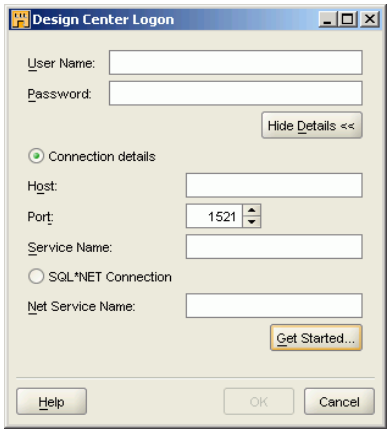


Figure 62 Oracle Warehouse Builder login window

As shown in Figure 63, choose 'Basic Install' in the *Install Type* window.

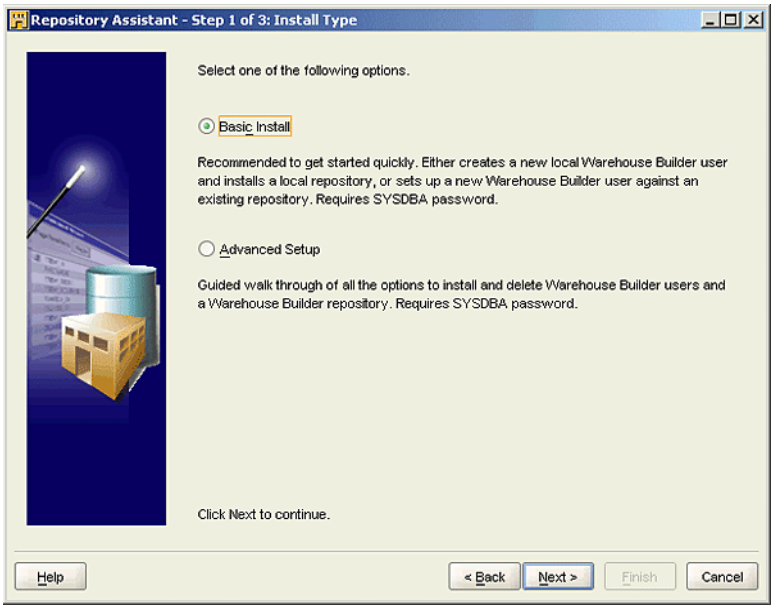


Figure 63 Oracle Warehouse Builder- installation type

As shown in Figure 64, create a user named steward_user and other items as follows in the *Repository User and Connection Information* window.

Repository User Name: **steward_user**
Repository User Password: **trc513**

SYSDBA User Name: **SYS**
SYSDBA Password: **trc513**

Host Name: **LOCALHOST**
Port Number: **1521**
Oracle Service Name: **steward2**

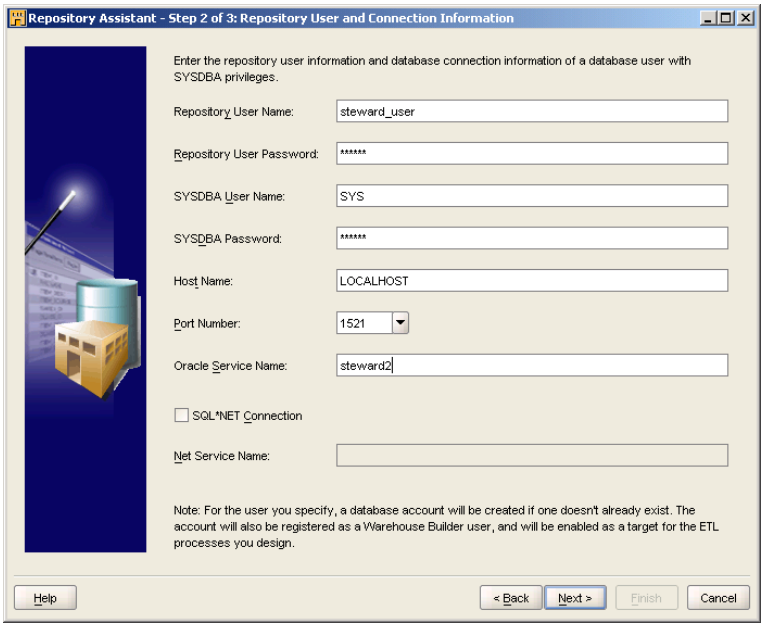


Figure 64 Oracle Warehouse Builder- repository user and connection information

As shown in Figure 65, re-enter the password for steward_user in the *Password Confirmation* window.

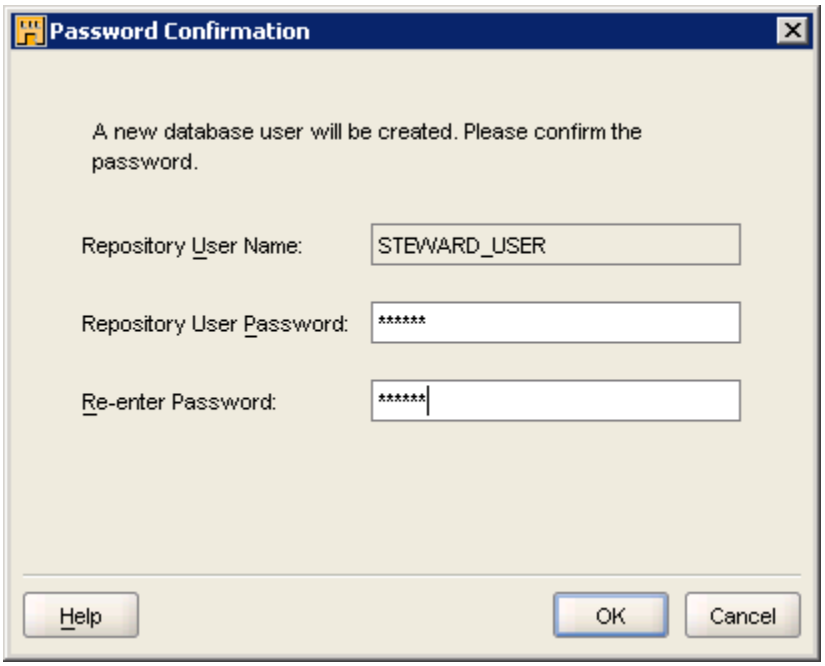


Figure 65 Oracle Warehouse Builder- password confirmation

As shown in Figure 66, type the owner name in the *Repository Owner Information* window. The repository owner is a highly privileged Warehouse Builder user with access to additional security features. Enter "steward_owner" as the username and password "trc513."

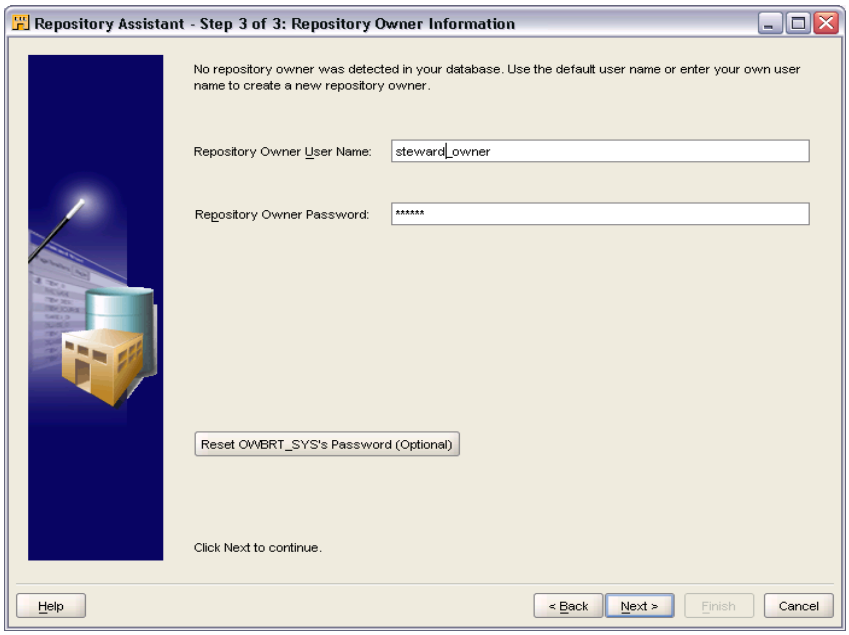


Figure 66 Oracle Warehouse Builder- repository owner information

The repository installation will take several minutes as shown in Figure 67.

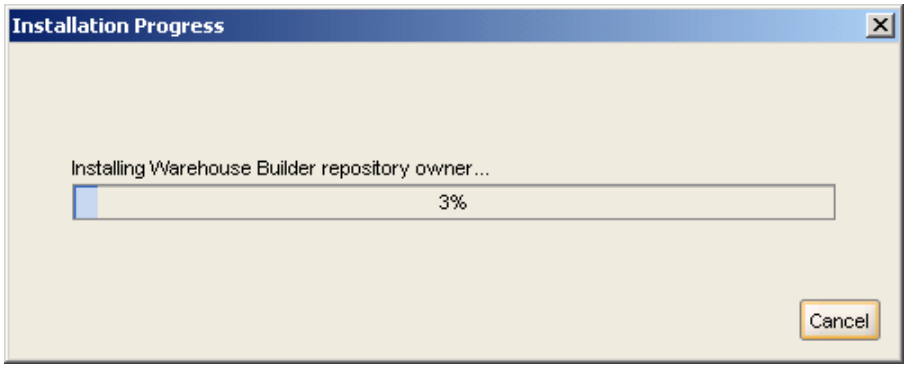


Figure 67 Oracle Warehouse Builder- installation progress

An *Installation Successful* message will appear as shown in Figure 68.

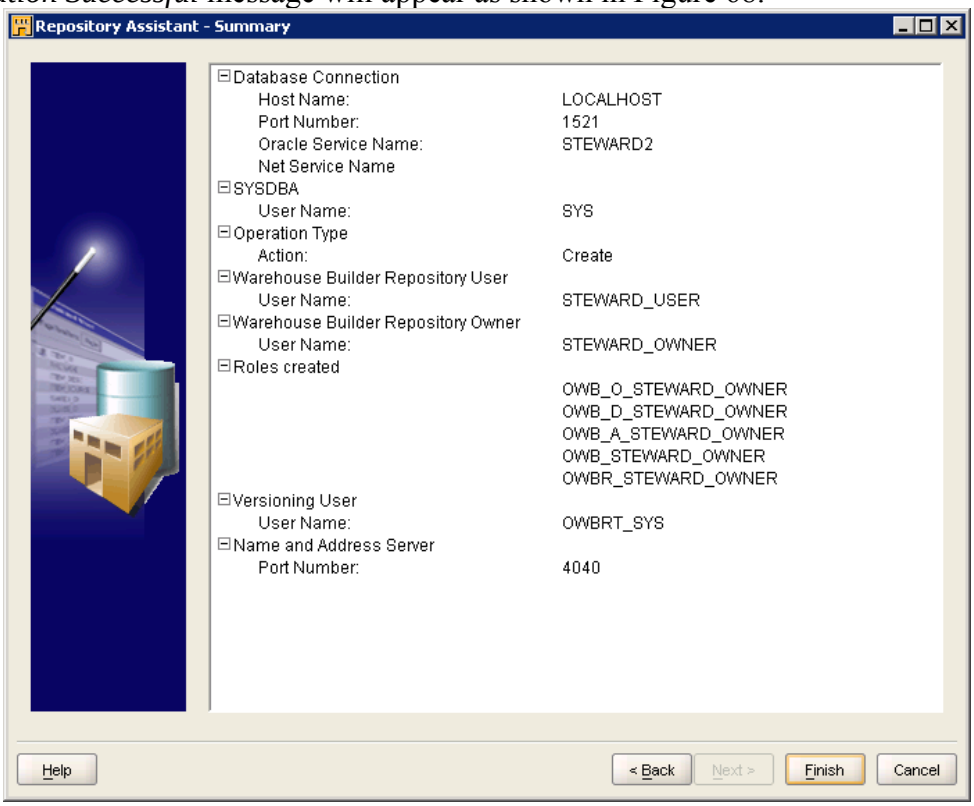


Figure 68 Oracle Warehouse Builder- installation summary

2.8 Oracle Workflow Configuration

2.8.1 Create the owf_mgr Workflow Schema using the Workflow Configuration Assistant.

Run the Oracle Workflow Configuration Assistant program as shown in Figure 69.

START > All Programs > {Oracle_Home} > Configuration and Migration tools > Workflow Configuration Assistant.

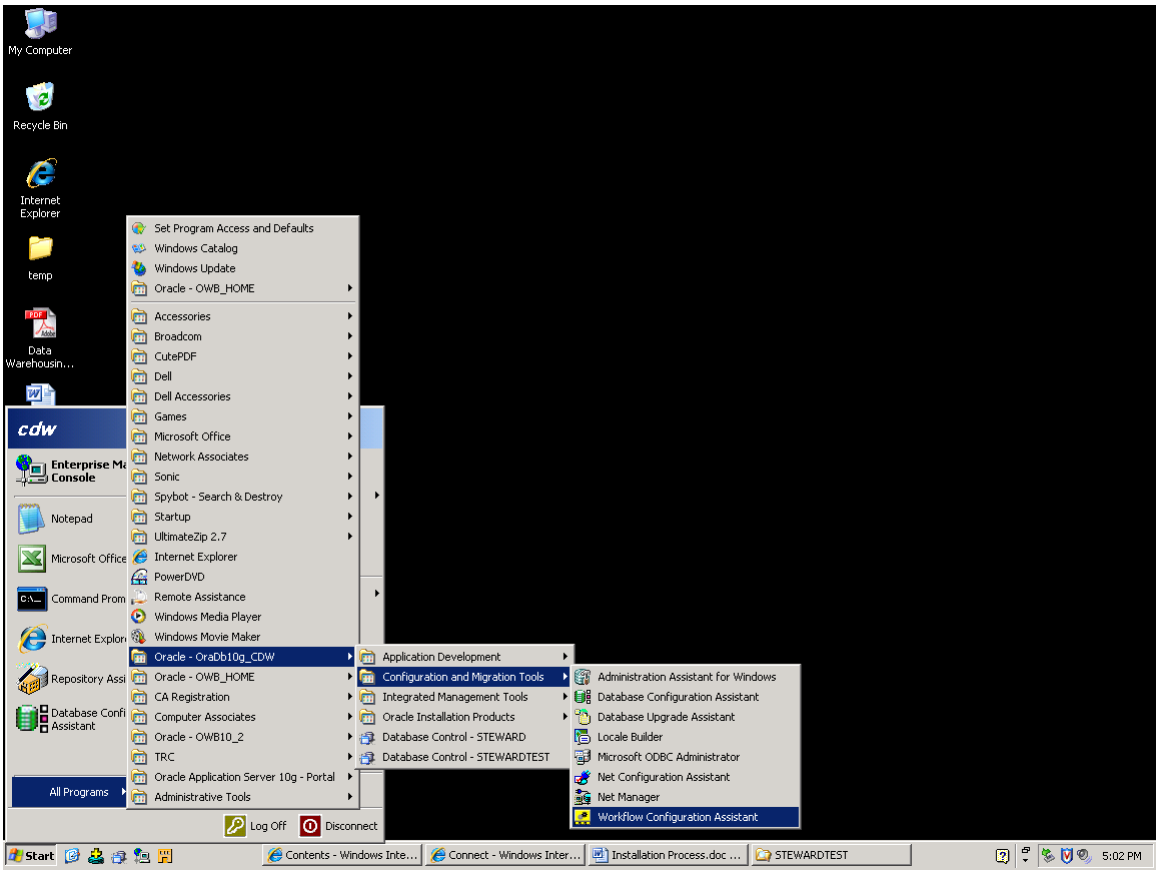


Figure 69 Oracle workflow configuration

2.8.2 Set the Workflow Configuration

Set the workflow account as follows. The Workflow Configuration Assistant window is shown in Figure 70.

- Workflow Account owf_mgr
- Workflow Password: **trc513**
- Sys Password: **trc513**
- TNS Connect descriptor: **localhost:1521:steward**

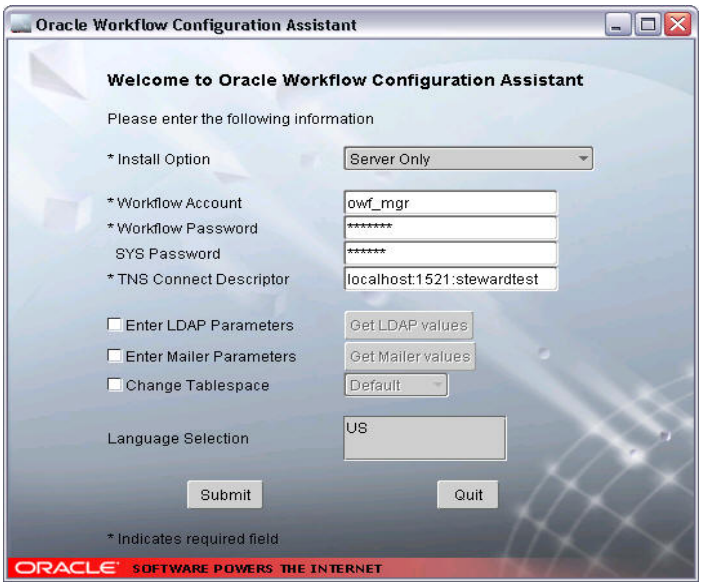


Figure 70 Oracle Workflow Configuration Assistant

2.8.3 Unlock the owf_mgr Account

After the Workflow Configuration Assistant has completed successfully, unlock the owf_mgr account. Log into SQL Plus as a SYSDBA(system/trc513/steward2). As shown in Figure 71, enter the following commands.

- alter user owf_mgr account unlock;
- grant execute any procedure to owf_mgr;
- grant owb_o_steward_owner to owf_mgr;

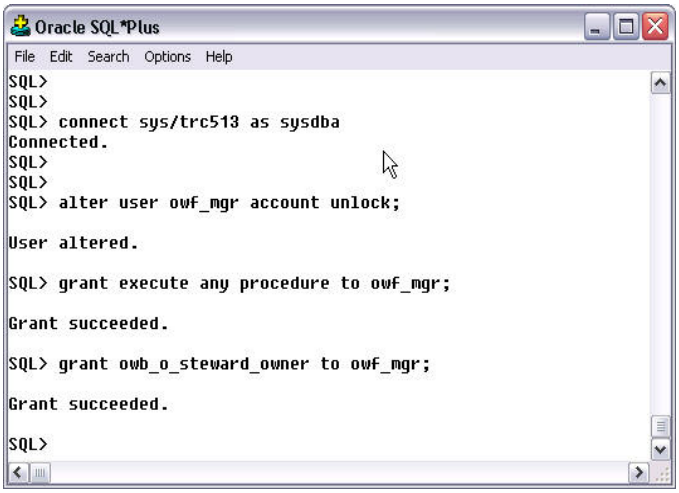


Figure 71 Oracle workflow configuration - unlock the owf_mgr

3 STEWARD Deployment

3.1 First Step - Login

Since all tasks are done by the 'steward_owner' authority, you must logon as 'steward_owner.' As shown Figure 72, login to the OWB Design Center with the following information.

- User Name: steward_owner
- Password: trc513
- Host: cdwserver.ce.ufl.edu
- Service Name: steward

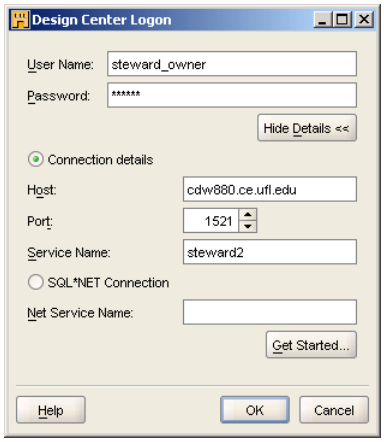


Figure 72 Oracle workflow configuration - unlock the owf_mgr

3.2 Prerequisites- Create a Target User

The target schema in the Oracle Warehouse Builder is a target to which the user can load the data and the data objects, such as cubes, dimensions, views, or mappings. This target schema physically stores target objects on deployment.

Target module consists of target schema and references a target schema by an assigned location. Every target module must be mapped to a target user schema. To create the STEWARD target schema users, perform the following steps.

Expand the Security node in the Global Explorer panel. Right-click the Users node and select New. Then select Create DB User to display the screen shown in Figure 73.

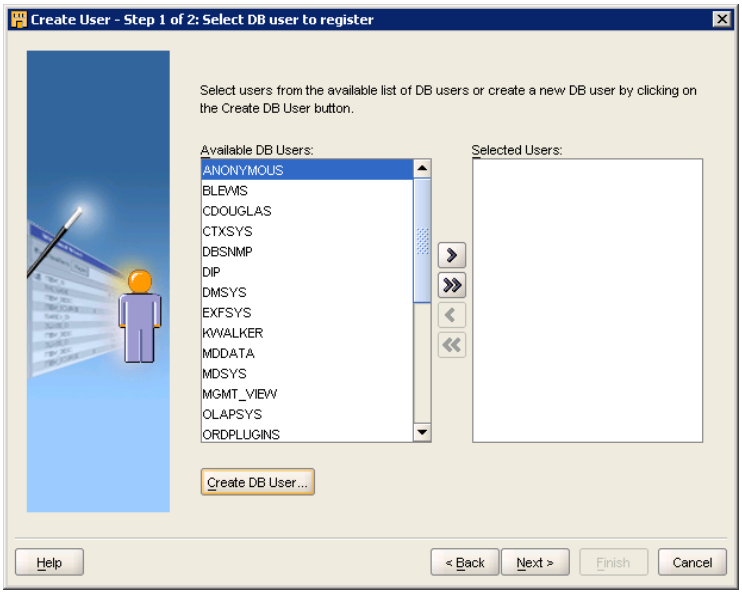


Figure 73 Oracle Warehouse Builder - create DB user

Create the DB user as follows: As shown Figure 74, enter the information required. (Name: gator Password: trc513)

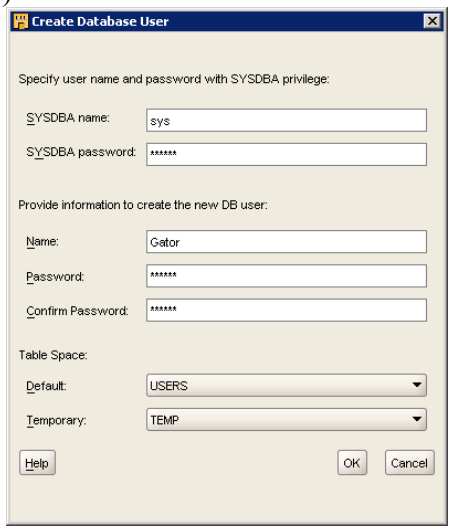


Figure 74 Oracle Warehouse Builder - create database user

After checking the status as shown Figure 75, click 'Next.'

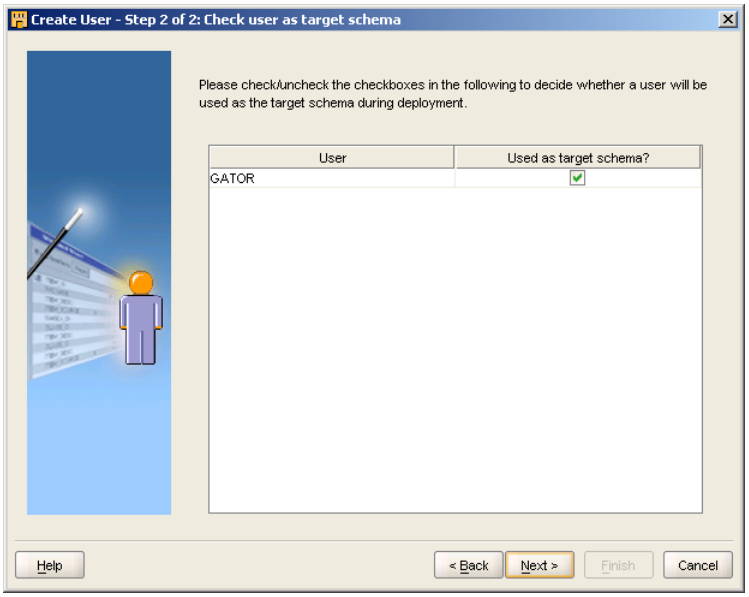


Figure 75 Oracle Warehouse Builder - check user as target schema

When the password is required as shown Figure 76, enter 'trc513.' Click 'OK.'

Figure 76 Oracle Warehouse Builder - target user password

As shown Figure 77, the “create user” procedure is done. Click ‘Finish.’

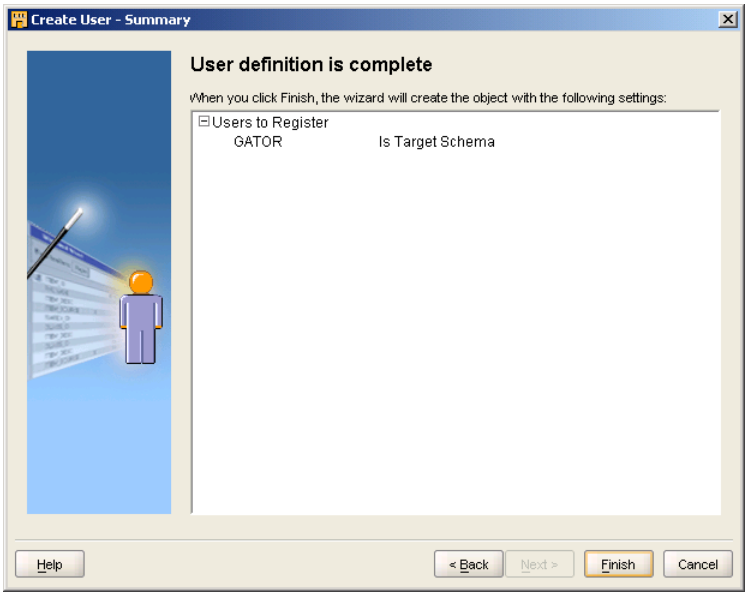


Figure 77 Oracle Warehouse Builder - check user summary

As shown Figure 78, the ‘Register User’ procedure is being performed over a few minutes.

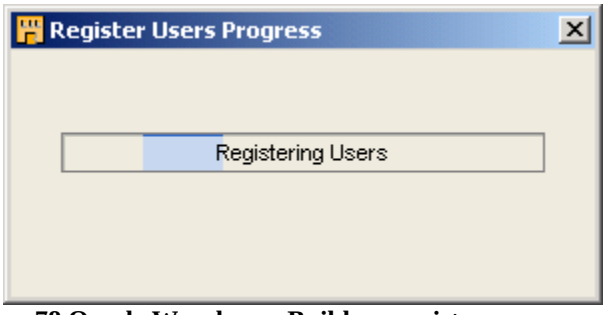


Figure 78 Oracle Warehouse Builder - register users progress

In the Connection Explorer panel, expand the Oracle Locations node as shown in Figure 79. Verify that a new location, "Steward_user _LOCATION." has been added. (Connection Explorer → Locations →Databases→Oracle)

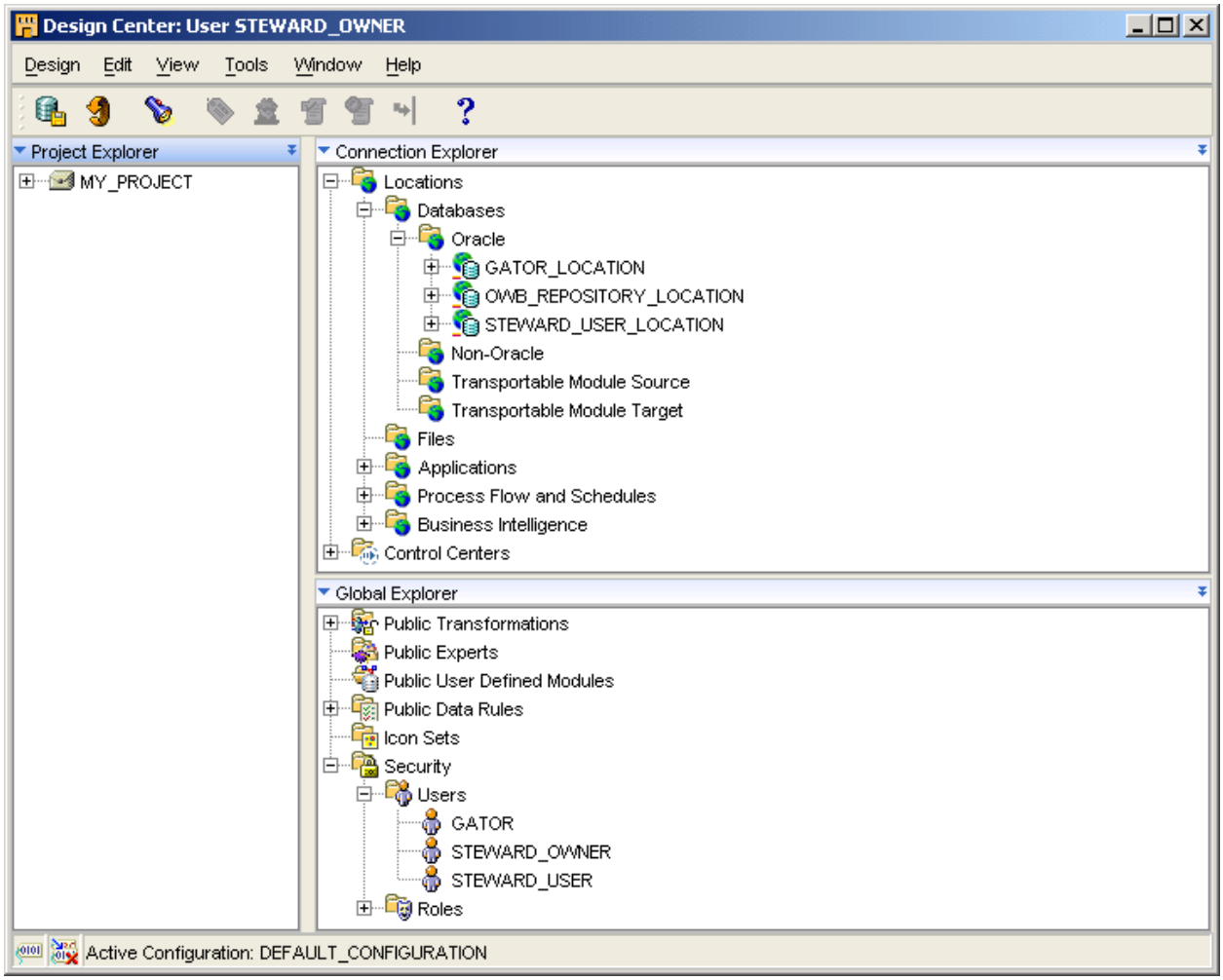


Figure 79 Oracle Warehouse Builder - Connection Explorer

Delete the 'GATOR_LOCATION' in the Oracle Locations before importing Metadata. Figure 80 shows the message for deleting Gator_Location

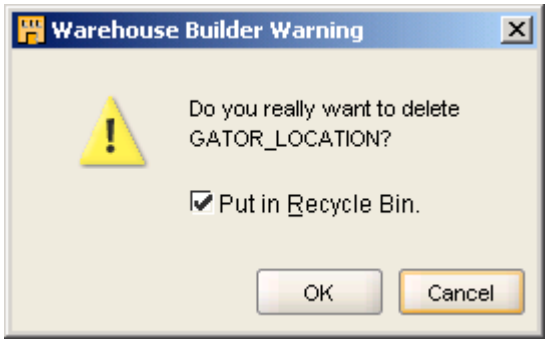


Figure 80 Oracle Warehouse Builder - delete Gator_Location

3.3 Prerequisites- Create New Files

Files in the Oracle Warehouse Builder are used for external file loading. If the new STEWARD system needs external file locations for file loading, they can be defined as *new* for the system. These files need to be defined before the STEWARD schemas are moved to new system.

To make the new system the same as the existing STEWARD system, perform the following steps.

Create the following folders in c:\Steward folder

- C:\Steward\Scripts
- C:\Steward\TSS_FacilityData
- C:\Steward\TSS_LANE_DATA
- C:\Steward\TSS_STATION_DATA
- C:\Steward\TSS_STATION_DATA\TSS_5Min_Data
- C:\Steward\TSS_STATION_DATA\TSS_15Min_Data
- C:\Steward\TSS_STATION_DATA\TSS_1HR_Data
- C:\Steward\TSS_LANE_DATA\TSS_5Min_Data
- C:\Steward\TSS_LANE_DATA\TSS_15Min_Data
- C:\Steward\TSS_LANE_DATA\TSS_1HR_Data

The following sections from 3.3.1 to 3.3.18 show an example of new files for TSS_15Min_Data. All of the following files need be created as *new*.

1) TIME_REF
2) TSS_15MIN_DATA
3) TSS_15MIN_LANE_DATA
4) TSS_1HR_DATA
5) TSS_1HR_LANE_DATA
6) TSS_5MIN_DATA
7) TSS_5MIN_LANE_DATA
8) TSS_ETL_REPORTS
9) TSS_FACILITY

3.3.1 Oracle Warehouse Builder - Create New File

As Figure 81 shows, select *new* from the Files menu.

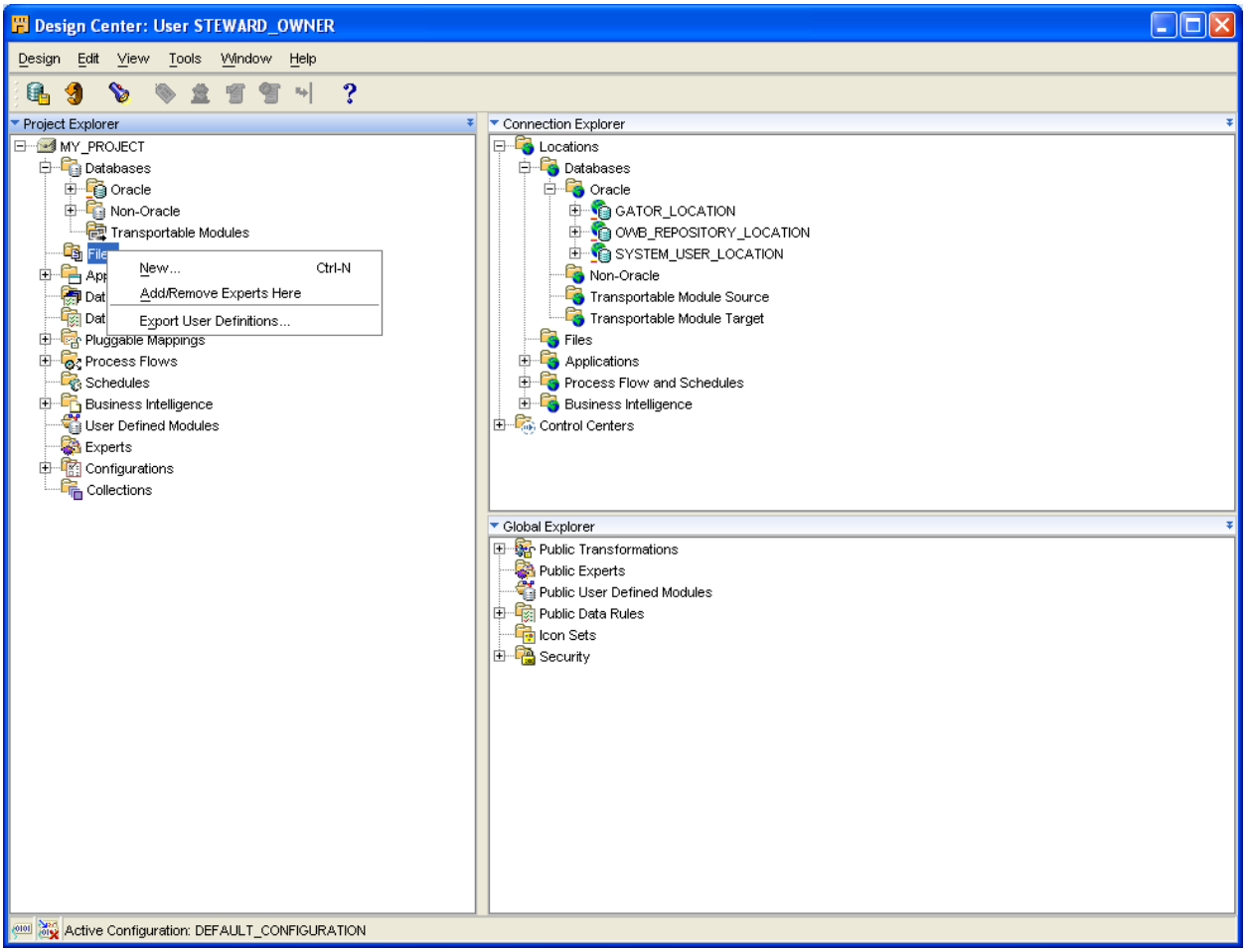


Figure 81 Oracle Warehouse Builder - create new file

3.3.2 Oracle Warehouse Builder - Create Module Wizard

As Figure 82 shows, the *create module wizard* will be executed.

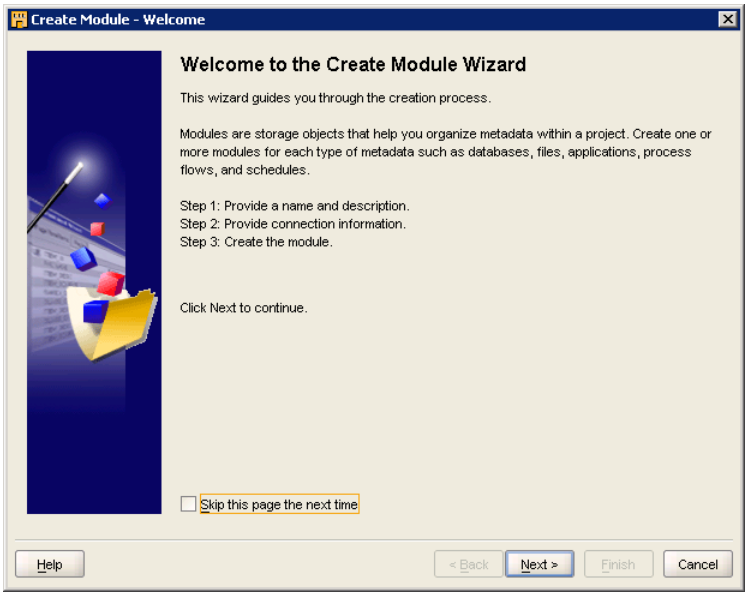


Figure 82 Oracle Warehouse Builder - Create Module Wizard

3.3.3 Oracle Warehouse Builder - Name and Description

As Figure 83 shows, add the name and description of the new module.

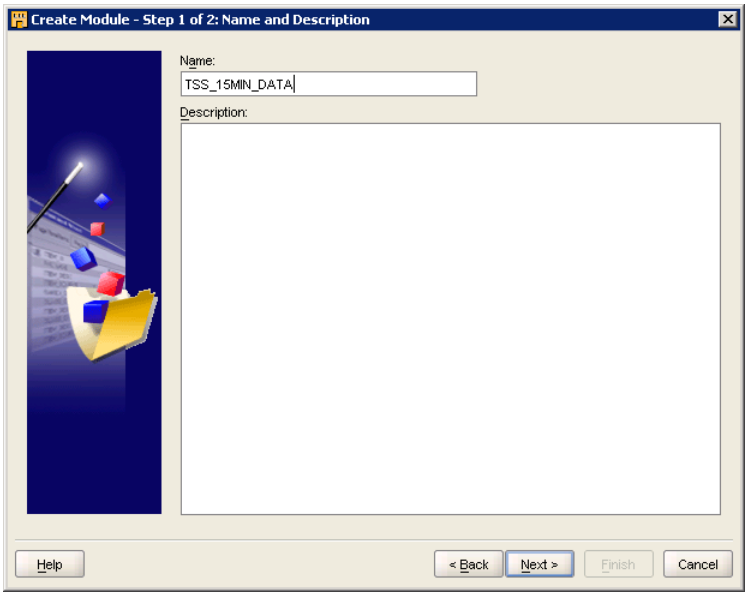


Figure 83 Oracle Warehouse Builder - name and description

3.3.4 Oracle Warehouse Builder - Select Location

As Figure 84 shows, select the location for new file and click the edit button.

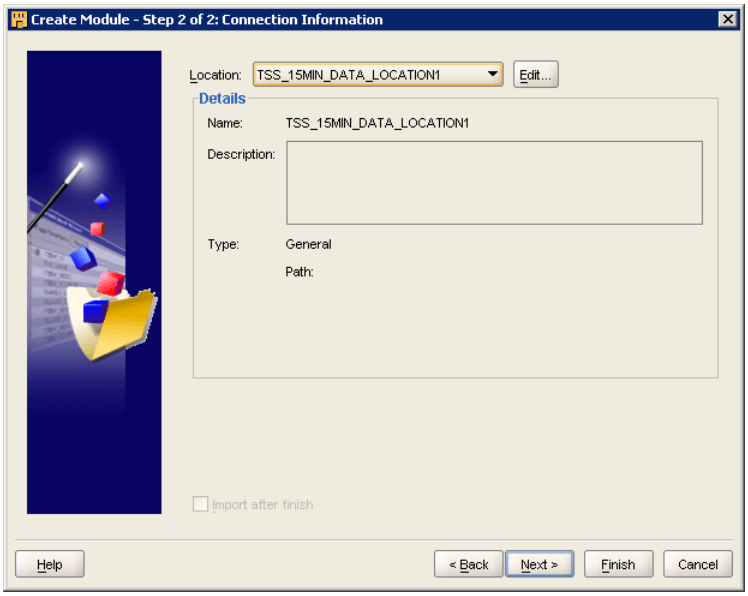


Figure 84 Oracle Warehouse Builder - connection information

3.3.5 Oracle Warehouse Builder - Select the Location Path

As Figure 85 shows, click *browse* and select the right location path.

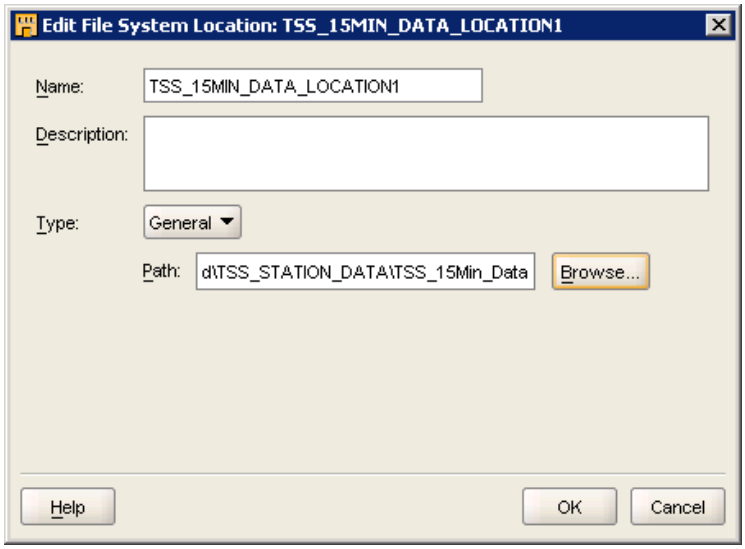


Figure 85 Oracle Warehouse Builder - select location path

3.3.6 Oracle Warehouse Builder - Finish Connection Information

As Figure 86 shows, the connection information is updated. Click 'next' for the next step.

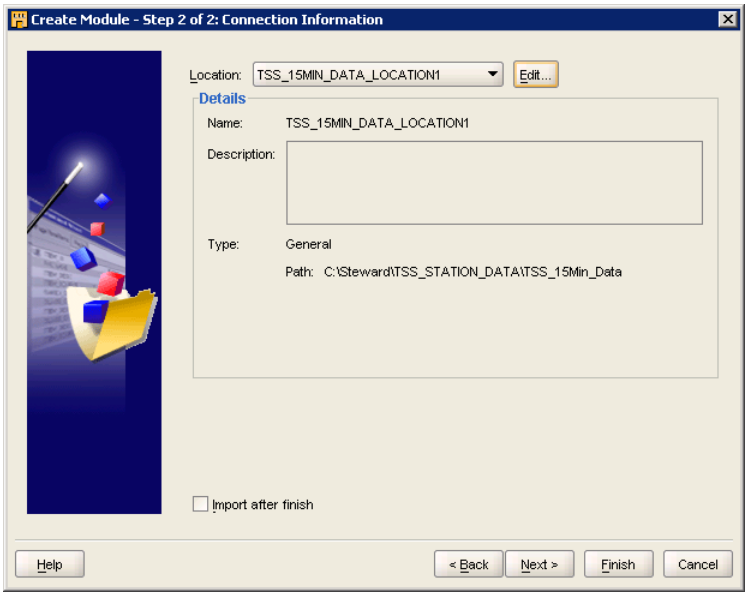


Figure 86 Oracle Warehouse Builder - finish connection information

3.3.7 Oracle Warehouse Builder - Module is Created

As Figure 87 shows, the module creation is finished.

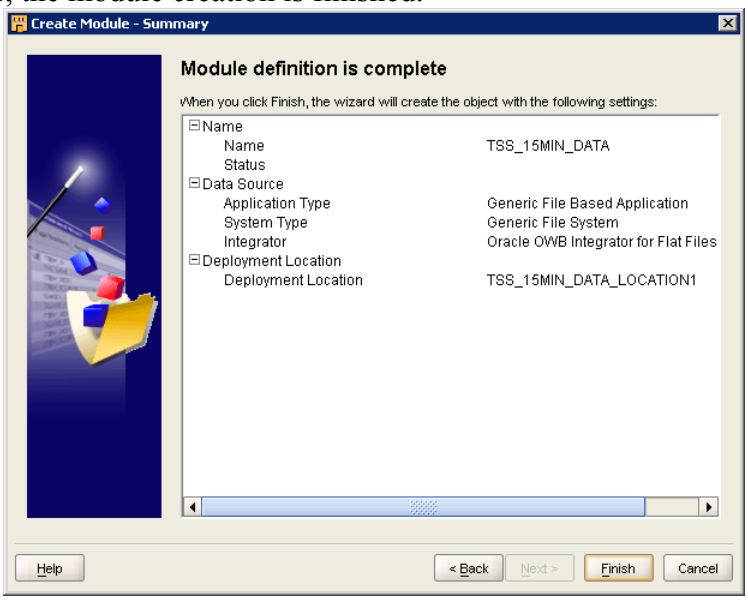


Figure 87 Oracle Warehouse Builder - finishing module creation

3.3.8 Oracle Warehouse Builder – Importing from the New Module

As Figure 88 shows, run *import* from the newly created module.

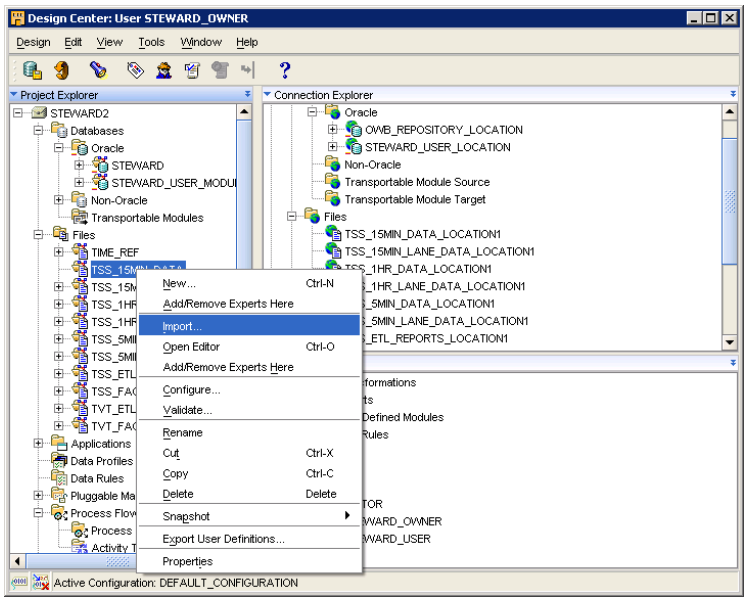


Figure 88 Oracle Warehouse Builder - importing files

3.3.9 Oracle Warehouse Builder - Import Metadata Wizard

As Figure 89 shows, the Import Metadata Wizard will be executed.

Figure 89 Oracle Warehouse Builder - Import Metadata Wizard

3.3.10 Oracle Warehouse Builder - Filter Information

As Figure 90 shows, select *All Data Files* for the filter information.

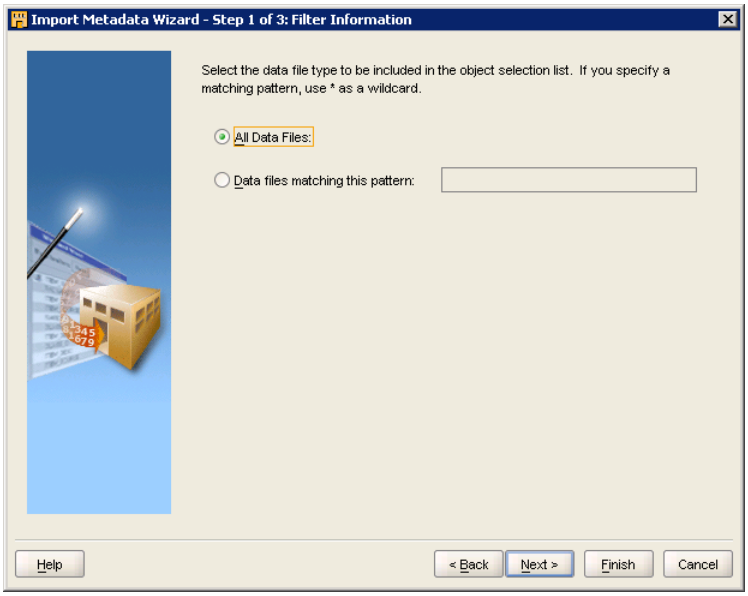


Figure 90 Oracle Warehouse Builder - filter information

3.3.11 Oracle Warehouse Builder - Object Selection

As Figure 91 shows, select the right object for the module.

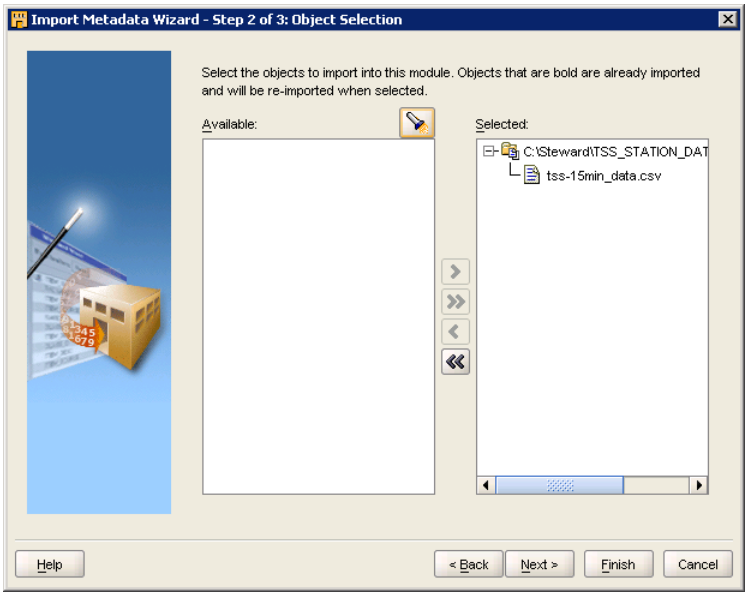


Figure 91 Oracle Warehouse Builder - select the right object for the module

3.3.12 Oracle Warehouse Builder - Run the Sample

As Figure 92 shows, after the object selection, click the sample for the object.

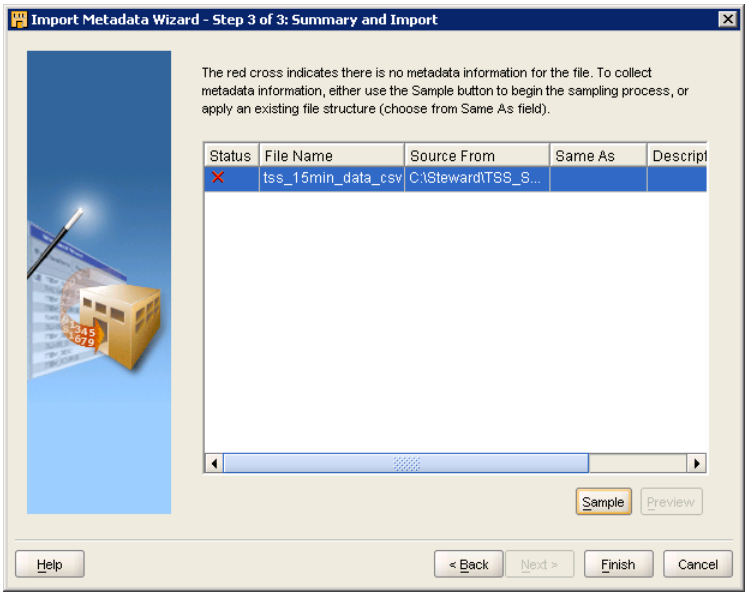


Figure 92 Oracle Warehouse Builder - run the sample

3.3.13Oracle Warehouse Builder - Flat File Sample Wizard

As Figure 93 shows, run the Flat File Sample Wizard

Figure 93 Oracle Warehouse Builder - Flat File Sample Wizard

3.3.14 Oracle Warehouse Builder - Check the File Name

As Figure 94 shows, Sep 1 is to check the file name.

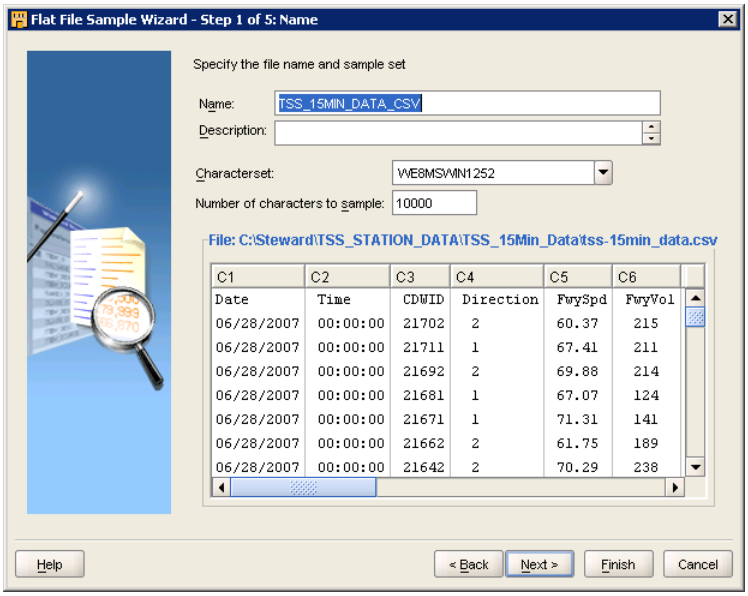


Figure 94 Oracle Warehouse Builder - check the file name

3.3.15 Oracle Warehouse Builder - Record Organization, File Format and File Layout

Figure 95, Figure 96 and Figure 97 show Step 2 (record organization), Step 3 (file format) and Step 4 (file layout), respectively. Default settings are used.

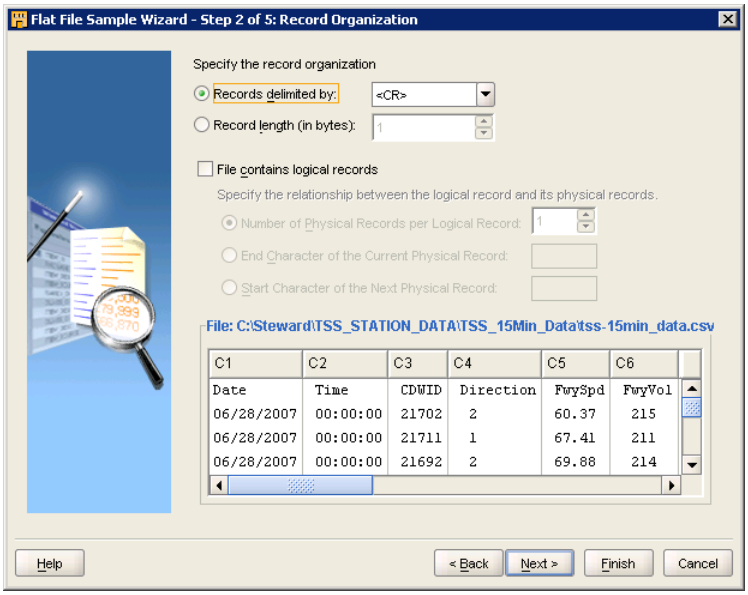


Figure 95 Oracle Warehouse Builder - report organization

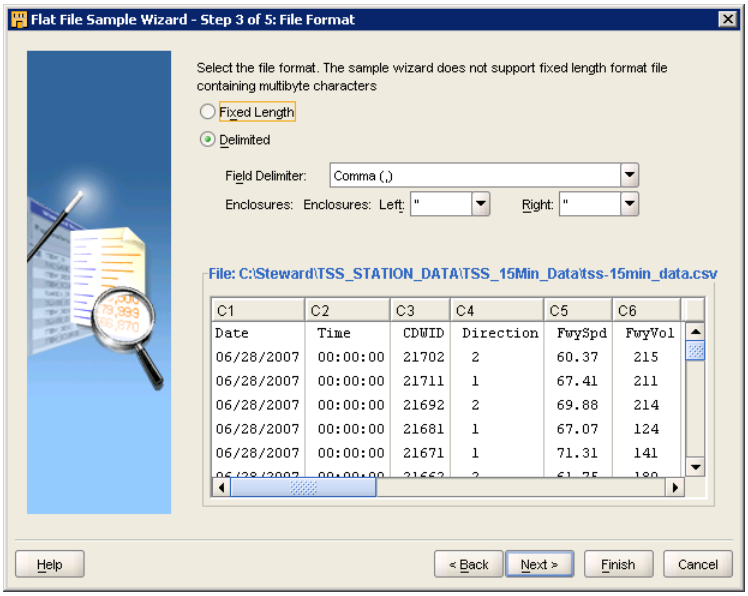


Figure 96 Oracle Warehouse Builder - file format

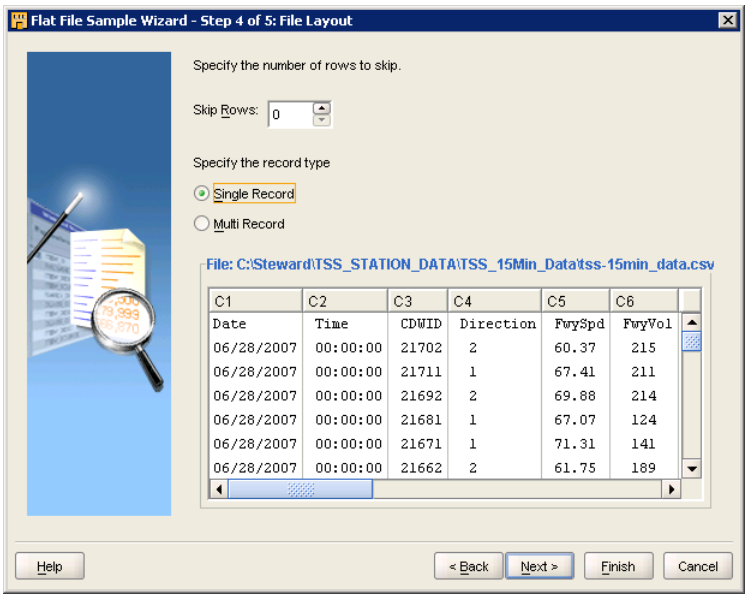


Figure 97 Oracle Warehouse Builder - file layout

3.3.16 Oracle Warehouse Builder - Field Properties

As Figure 98 shows, Step 5 sets the field properties. The field name, type and length need to be updated as defined in Figure 99. Also, choose “Use the first record as the field names.”

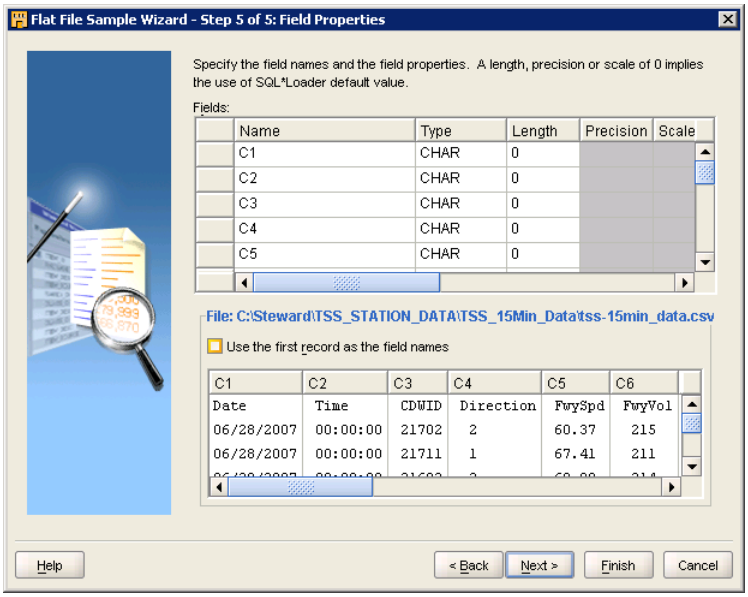


Figure 98 Oracle Warehouse Builder - field properties

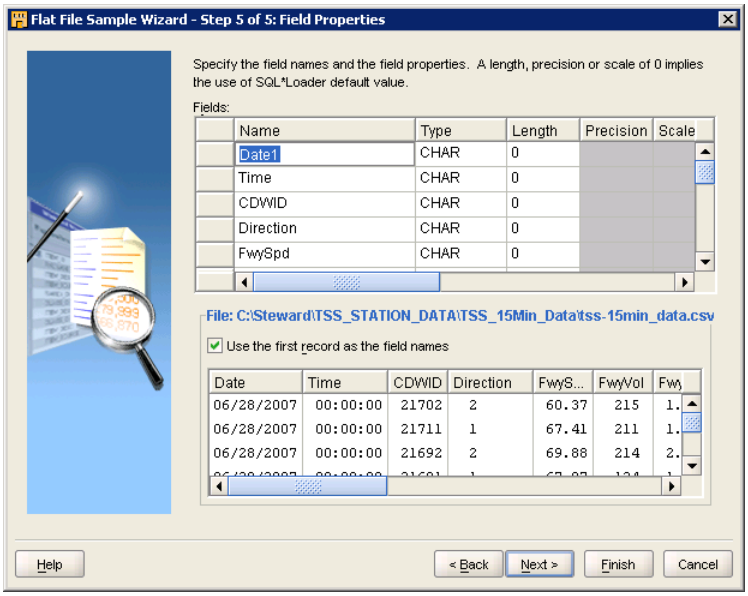


Figure 99 Oracle Warehouse Builder - updated field properties

3.3.17 Oracle Warehouse Builder - Summary

As Figure 100 shows, the flat file definition is done with summary windows.

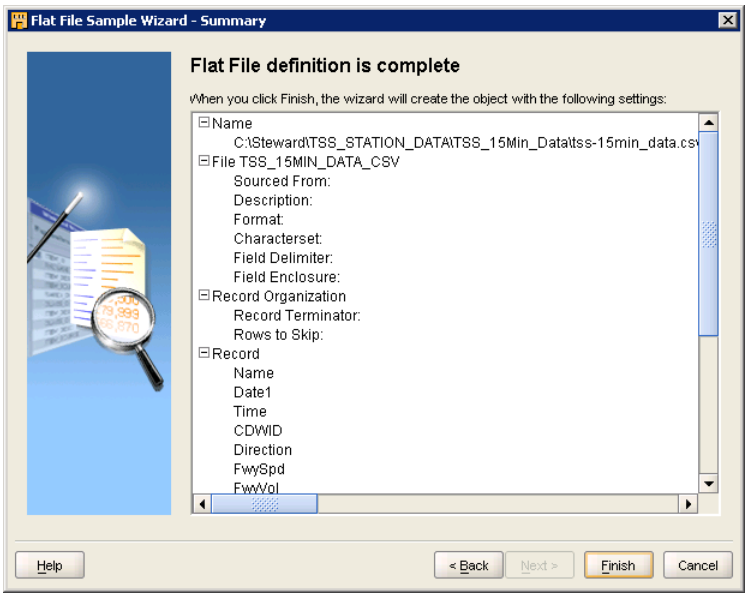


Figure 100 Oracle Warehouse Builder – flat file summary

3.3.18 Oracle Warehouse Builder - Import Data Wizard Summary

As Figure 101 shows, the Import Data Wizard is done with summary windows

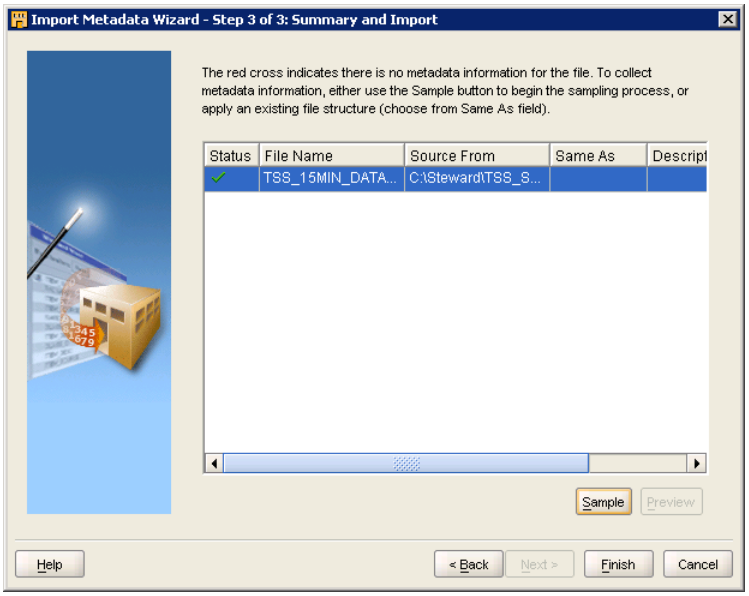


Figure 101 Oracle Warehouse Builder - Import Data Wizard summary

3.4 Importing Metadata

CASE tools such as Oracle Designer and Computer Associates ERwin generally provide a metadata repository to store definitions. In the Oracle Warehouse Builder 10g, a CASE tool for the design of business intelligence systems, also provides a metadata store.

Metadata in the Oracle Warehouse Builder includes all the information about the data. It can be exported and imported using the Metadata Loader (MDL) utility. For STEWARD migration, all of the STEWARD project metadata needs be exported from current STEWARD system and imported into the new STEWARD system.

3.4.1 Prerequisite- Export Warehouse Builder Metadata (Step 1)

Before importing metadata to a new location, the metadata needs to be exported from the current STEWARD location at the TRC. Figure 102 shows how to start the *Export Warehouse Builder Metadata* process from the Design Center.

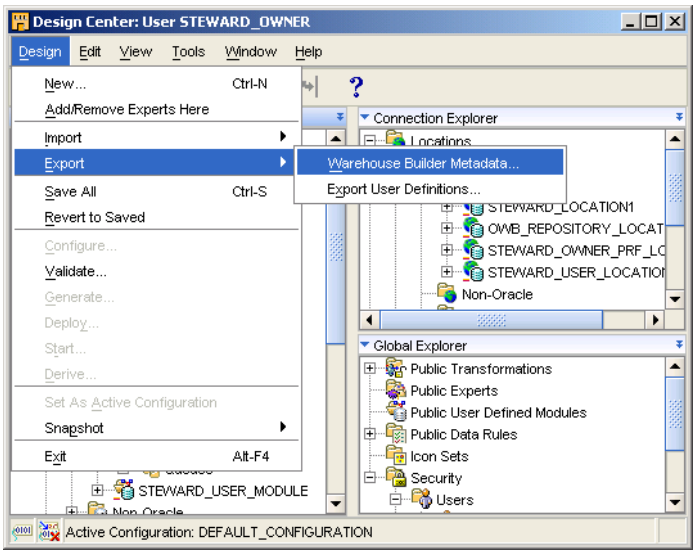


Figure 102 Oracle Warehouse Builder -export warehouse builder metadata

3.4.2 Prerequisite- Export Warehouse Builder Metadata (Step 2)

Figure 103 shows the metadata export window. In this window, check the checkbox 'Export all object dependencies.' And click the 'Export' button.

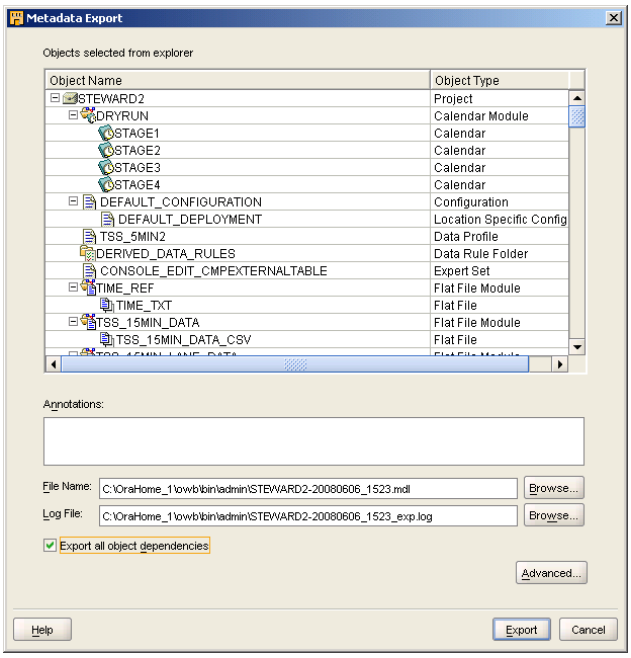


Figure 103 Oracle Warehouse Builder - metadata export window

3.4.3 Importing Metadata - Import Warehouse Builder Metadata (Step 1)

From the Design Center menu, select [Design > Import > Warehouse Builder Metadata]. Select the File Name and Log File that will be imported using the *Browse* button. Then click the *Import* button as shown in Figure 104.

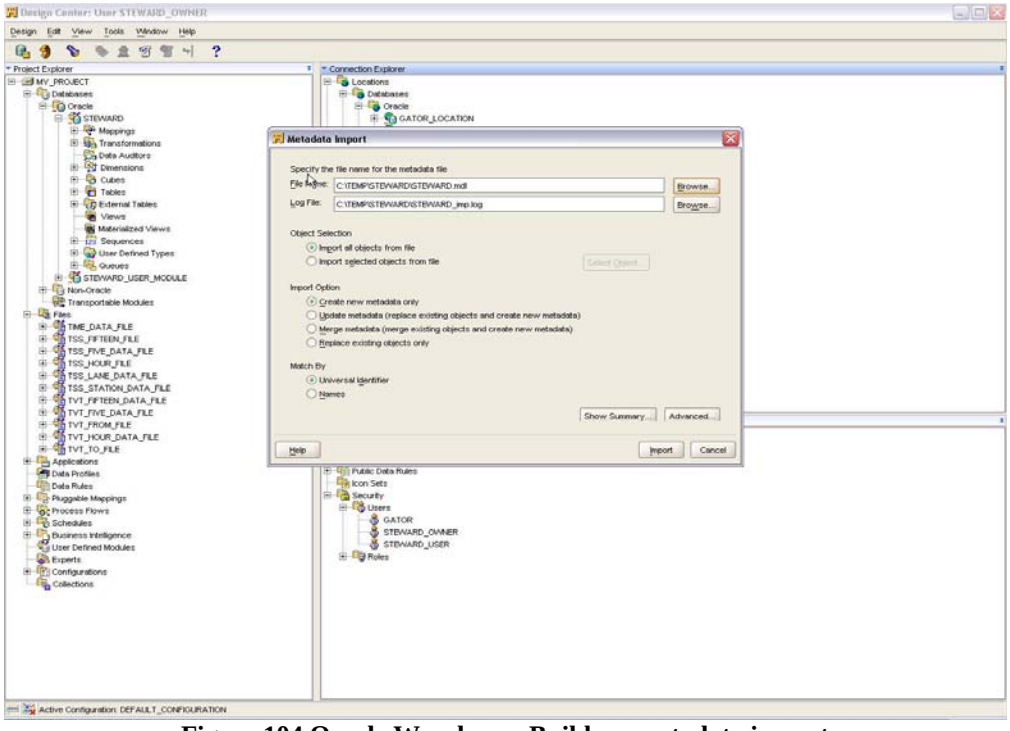


Figure 104 Oracle Warehouse Builder - metadata import

3.4.4 Importing Metadata - Import Warehouse Builder Metadata (Step 2)

The metadata needs to be registered using the *Control Center* tool. To register metadata, right click the following three locations: "GATOR_LOCATION," "LOAD_SCHEDULE_LOCATION," and "STEWARD_PF_MODLUE_LOCATION" and open the editor in the *Connection Explorer* as shown in Figure 105. Click the *Test* button.

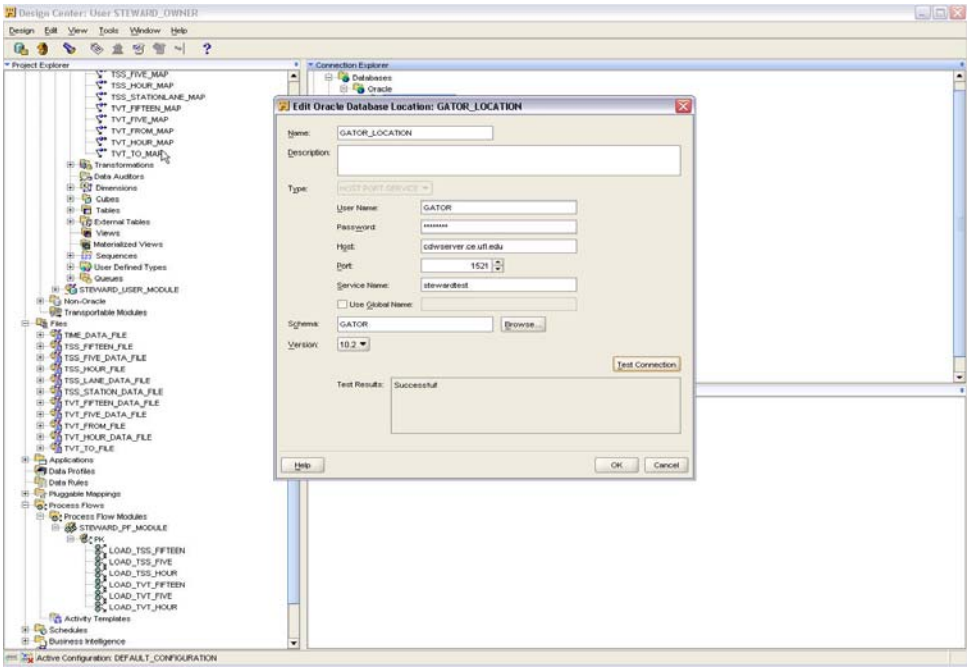


Figure 105 Oracle Warehouse Builder - edit database location

3.4.5 Importing Metadata - Import Warehouse Builder Metadata (Step 3)

Click *Tools* in the Design Center menu and click *Control Center Manager*.

In the Control Center, the previous three locations will be shown on the left side. They should be registered when saved as shown in Figure 106. When the external tables are deployed, the file locations can be registered automatically.

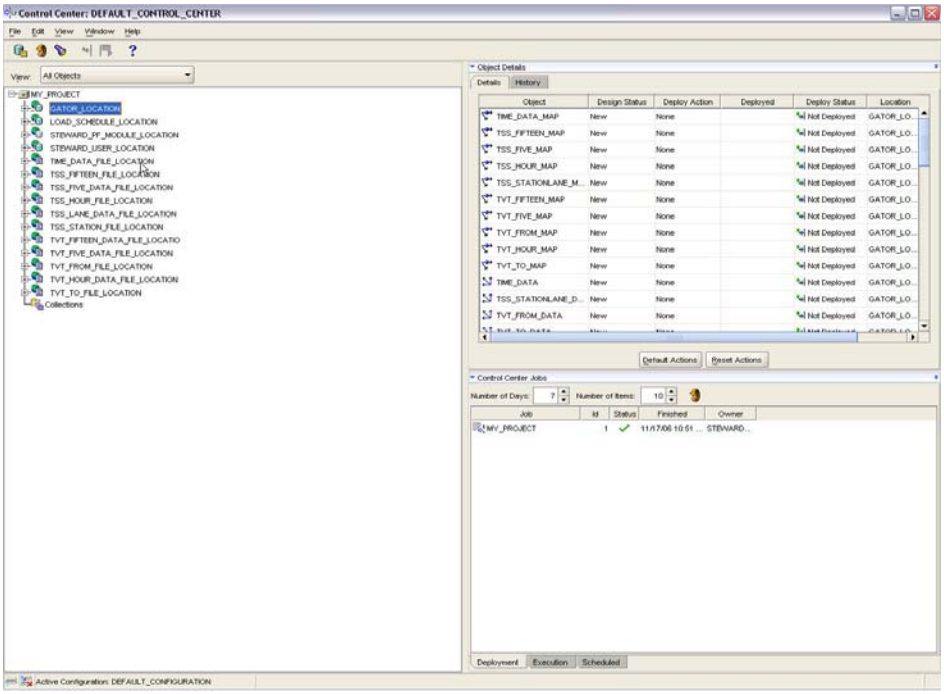


Figure 106 Oracle Warehouse Builder - file locations in the control center

3.5 Registration of the Control Center Manager

This step registers the imported module to the newly created depository. All locations need to be registered as follows.

As shown in Figure 107, select the *Oracle* tab from Connection Explorer and select all the *.LOCATION one by one. Select Open Editor menu and enter the user name, password, host, and service name as Figure 107 shows. Press Test Connection.

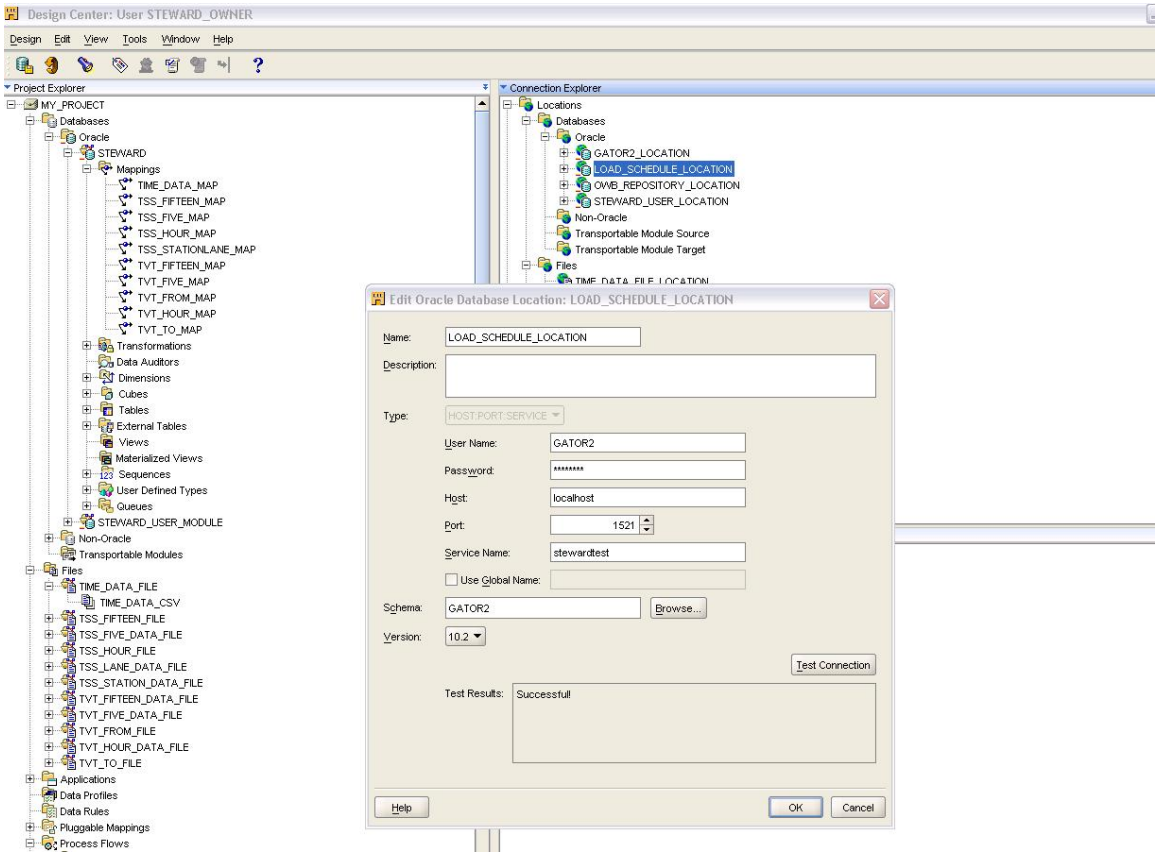


Figure 107 Oracle Warehouse Builder - edit *Connection Explorer*

As shown in Figure 106, Open the Control Center Manager and press the Refresh Button. You can now see all new LOCATION items on the left side.

As shown in Figure 108, select the File tab in the Connection Explorer and register all the locations as before.

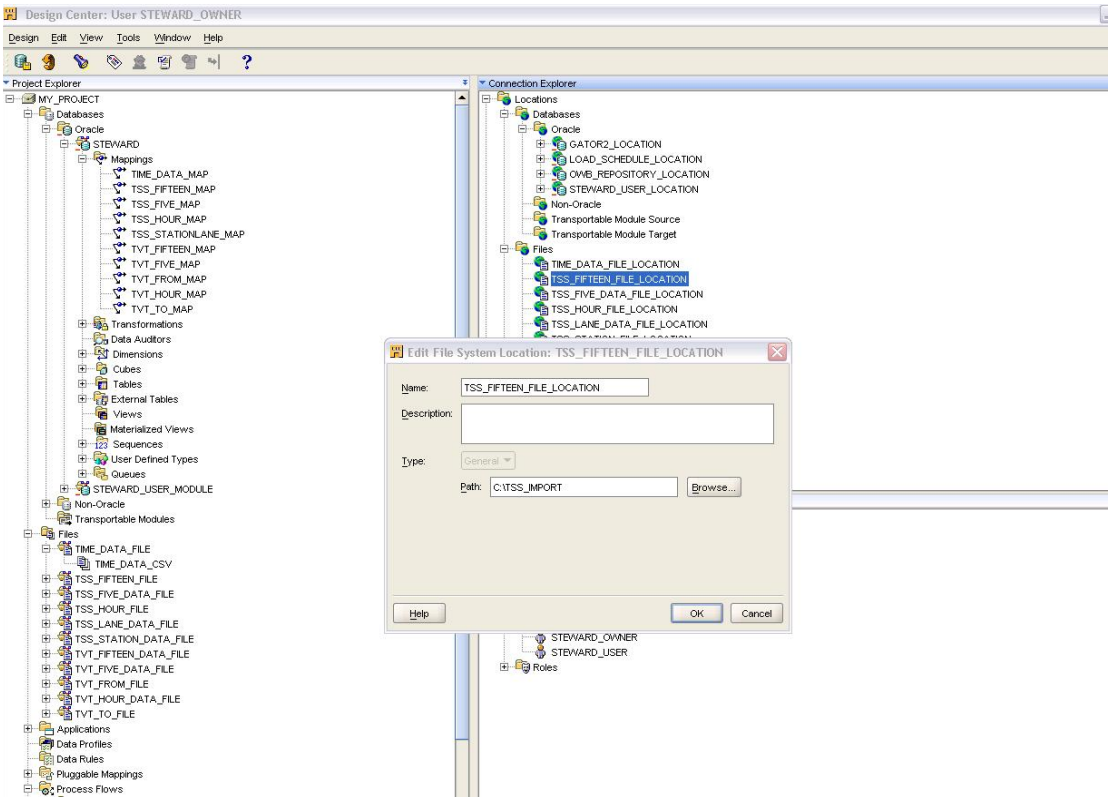


Figure 108 Oracle Warehouse Builder - edit *file system location*

As shown in Figure 106, open the Control Center Manager and refresh the left pane. All of the new LOCATION items will be on the left side.

As shown in Figure 109, select *Process Flow* and the *Schedule* tab in Connection Explorer and select *Oracle Workflow*. Select all *.LOCATION and register them as before.

Verify the new registration from the Control Center Manager on the left pane.

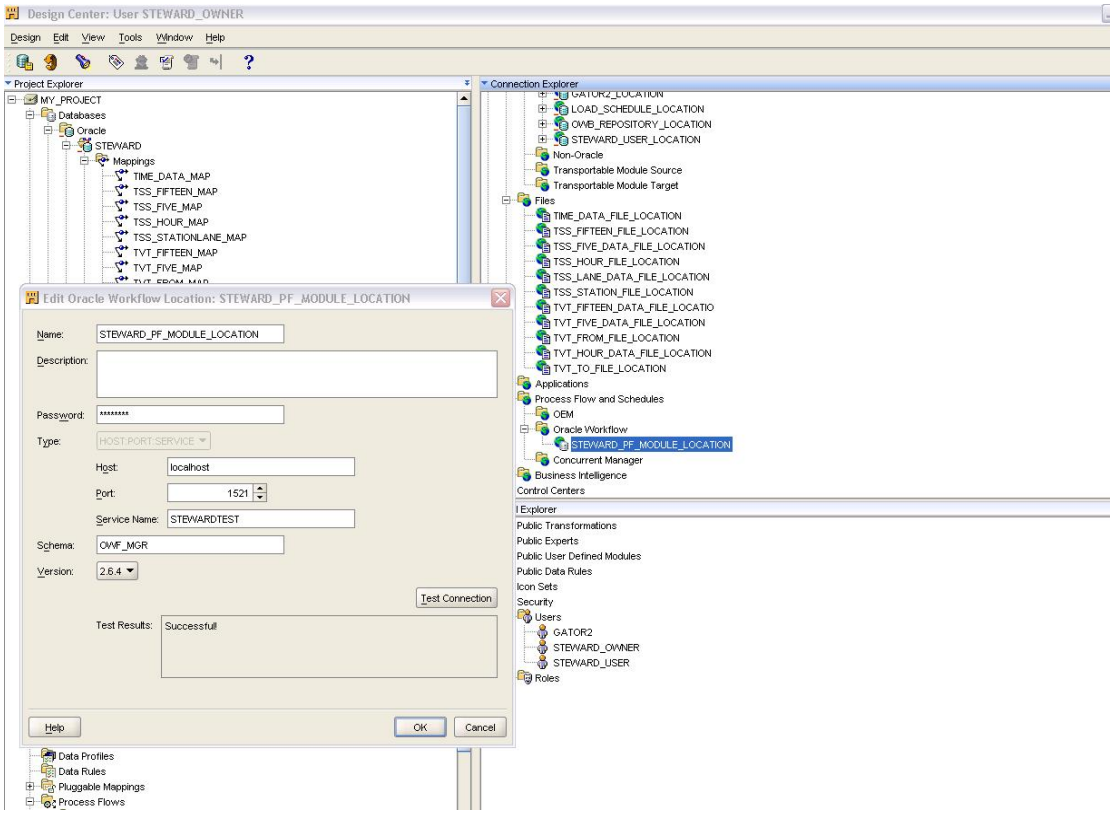


Figure 109 Oracle Warehouse Builder - edit process flow and schedule location

The updated Control Center is shown in Figure 110.

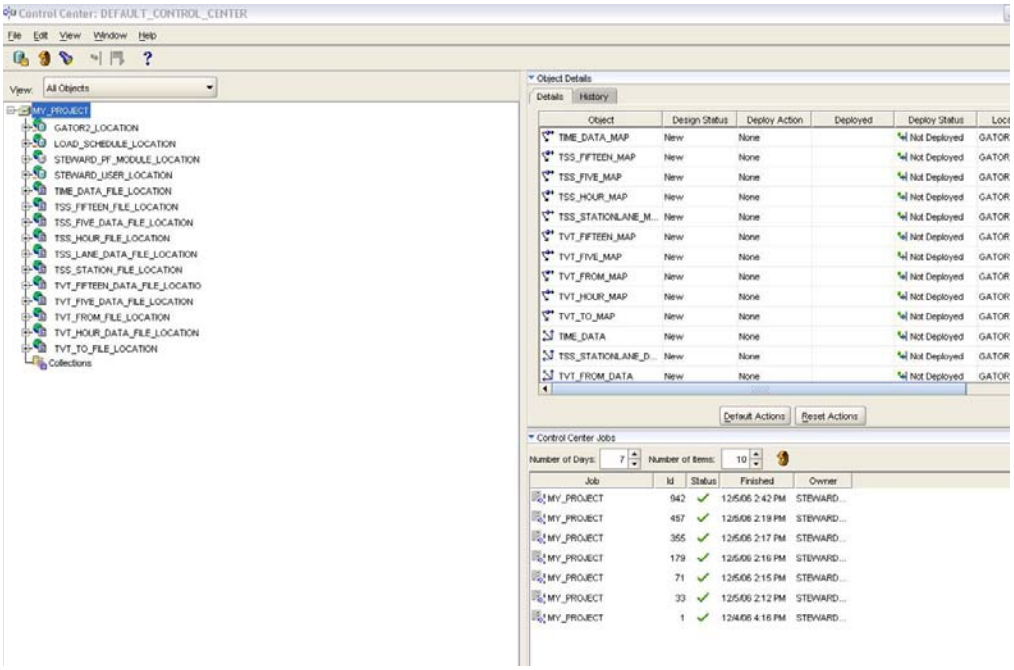


Figure 110 Oracle Warehouse Builder - updated Control Center

3.6 Data Deployment Process

Newly imported schemes need to be deployed in the control center to have the physical object within the database.

3.6.1 External Table Deployment

As shown in Figure 111, select *External Tables* and click *Default Actions* in the Control Center. Then, click the *Deploy* button to deploy the DB.

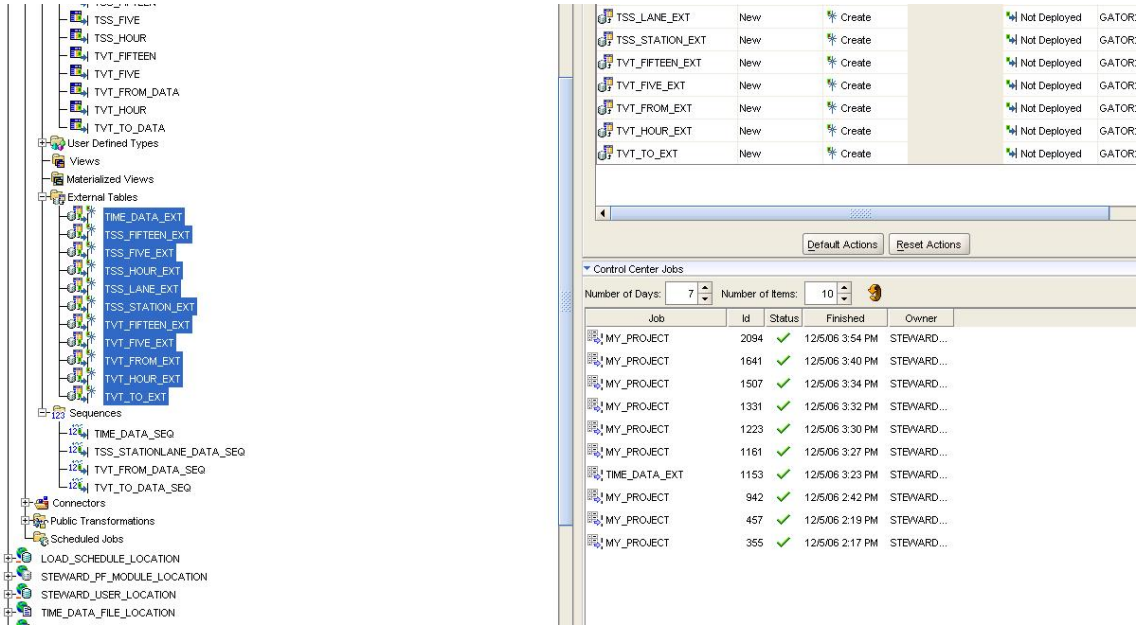


Figure 111 Oracle Warehouse Builder - external table deployment

3.6.2 Dimension Table Deployment

As shown in Figure 112, select all dimension tables and all sequences and click the *Default Actions* button. Then, click the *Deploy* button to deploy the DB.

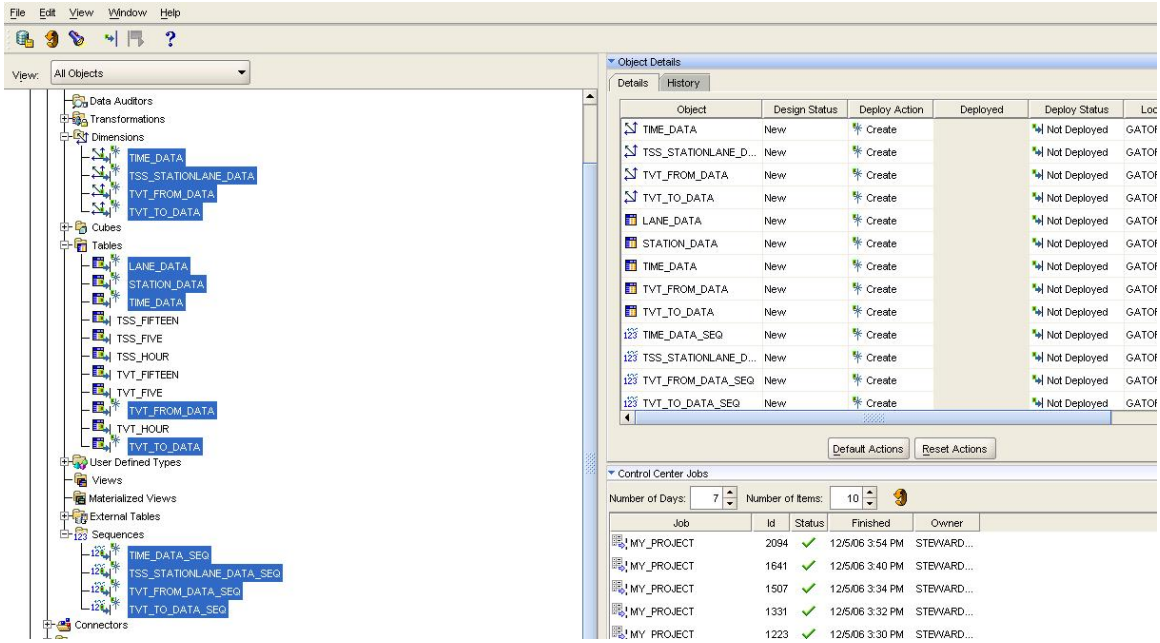


Figure 112 Oracle Warehouse Builder - Dimension Table Deployment

3.6.3 Fact Table Deployment

As shown in Figure 113, select all cubes and tables for the Fact Table and click the *Default Actions* button. Then, click the *Deploy* button to deploy the DB.

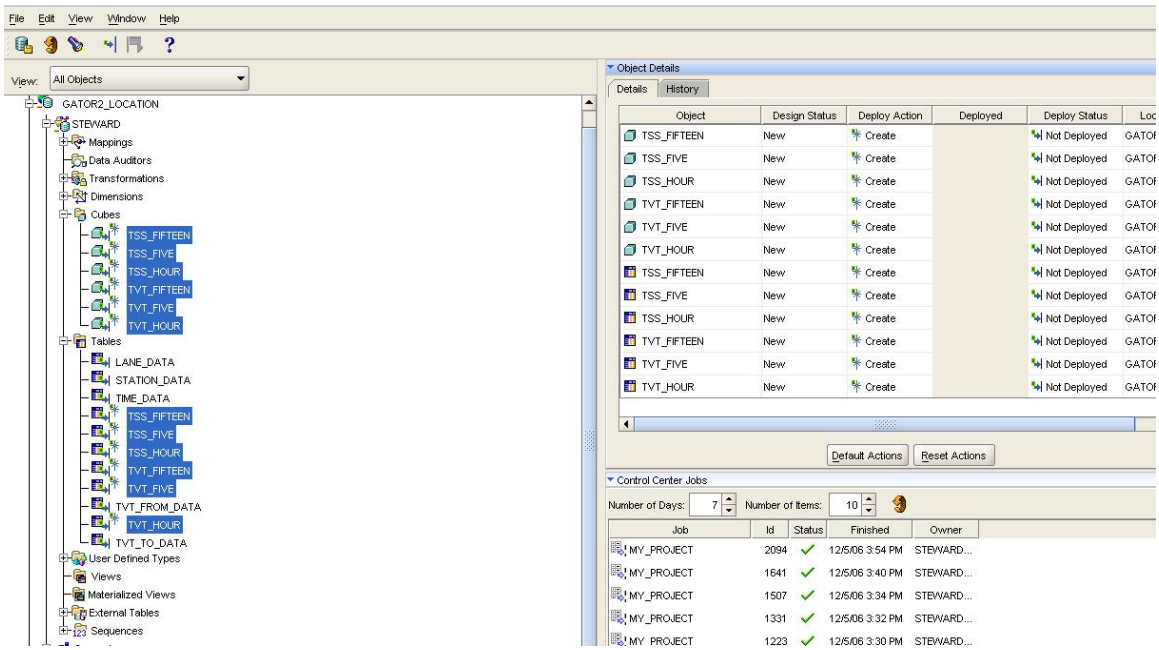


Figure 113 Oracle Warehouse Builder - fact table deployment

3.6.4 Mappings Deployment

As shown in Figure 114, select all mappings and click then *Default Actions* button. Then click the *Deploy* button to deploy the DB.

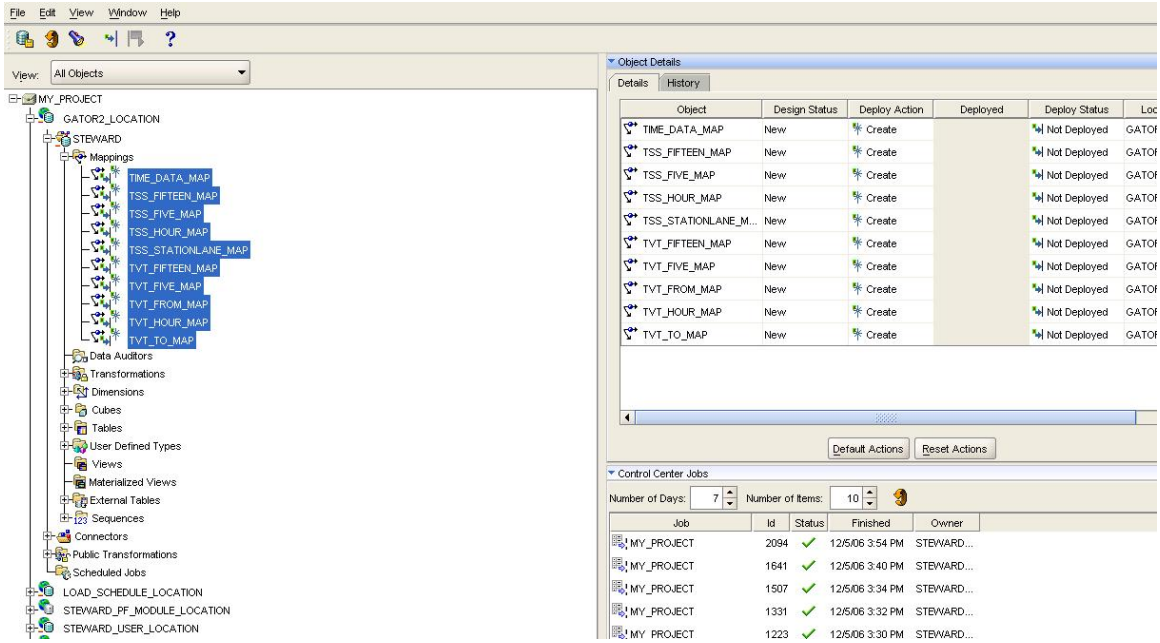


Figure 114 Oracle Warehouse Builder - mappings deployment

3.7 Data loading process

Target schema can be loaded with real data using the data loading process. Only one file defined in the external file mapping is loaded into table and can be used for the daily operation.

3.7.1 Dimension Table Data loading

The dimension table doesn't change at all. As shown in Figure 115, select mappings for each dimension table one by one. Then, click the *Start* button to begin loading data.

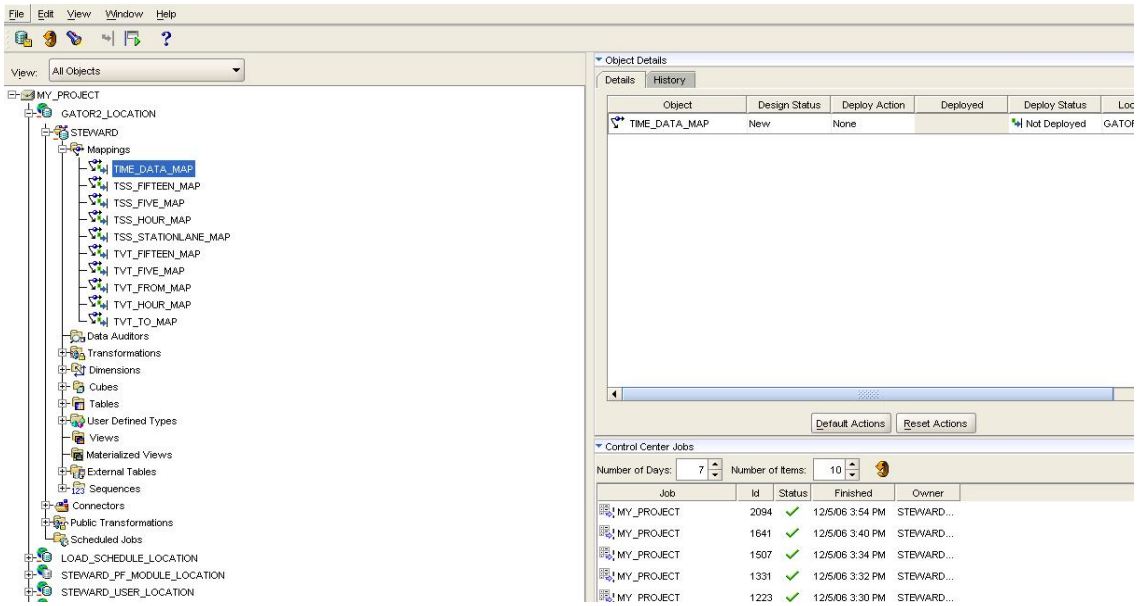


Figure 115 Oracle Warehouse Builder - loading dimension table data

3.7.2 Loading Fact Table Data

As shown in Figure 116, select *STEWARD_PF_MODULE_LOCATION* and deploy it first. Then select each item one by one and click the *Start* button. This will load one day's data. Remember to click *Start* button only one time.

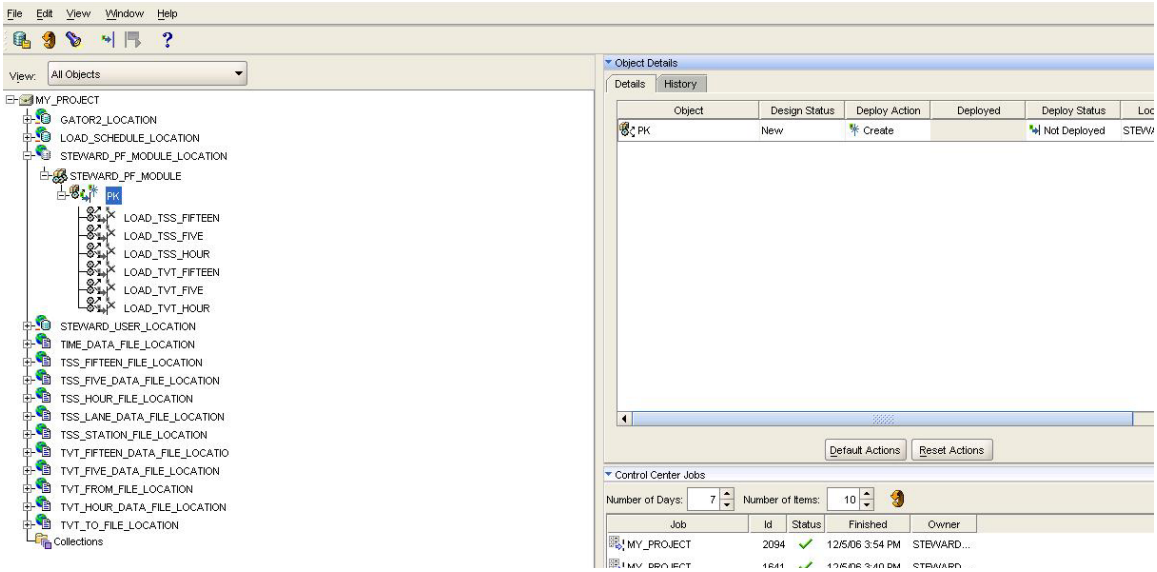


Figure 116 Oracle Warehouse Builder - loading fact table data

To load the next day's data, change the *.csv file in the source destination. Then, click the *Start* button. You can check the table contents using SQL*PLUS.

4 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. Access to the Net Configuration Assistant is illustrated in Figure 117.

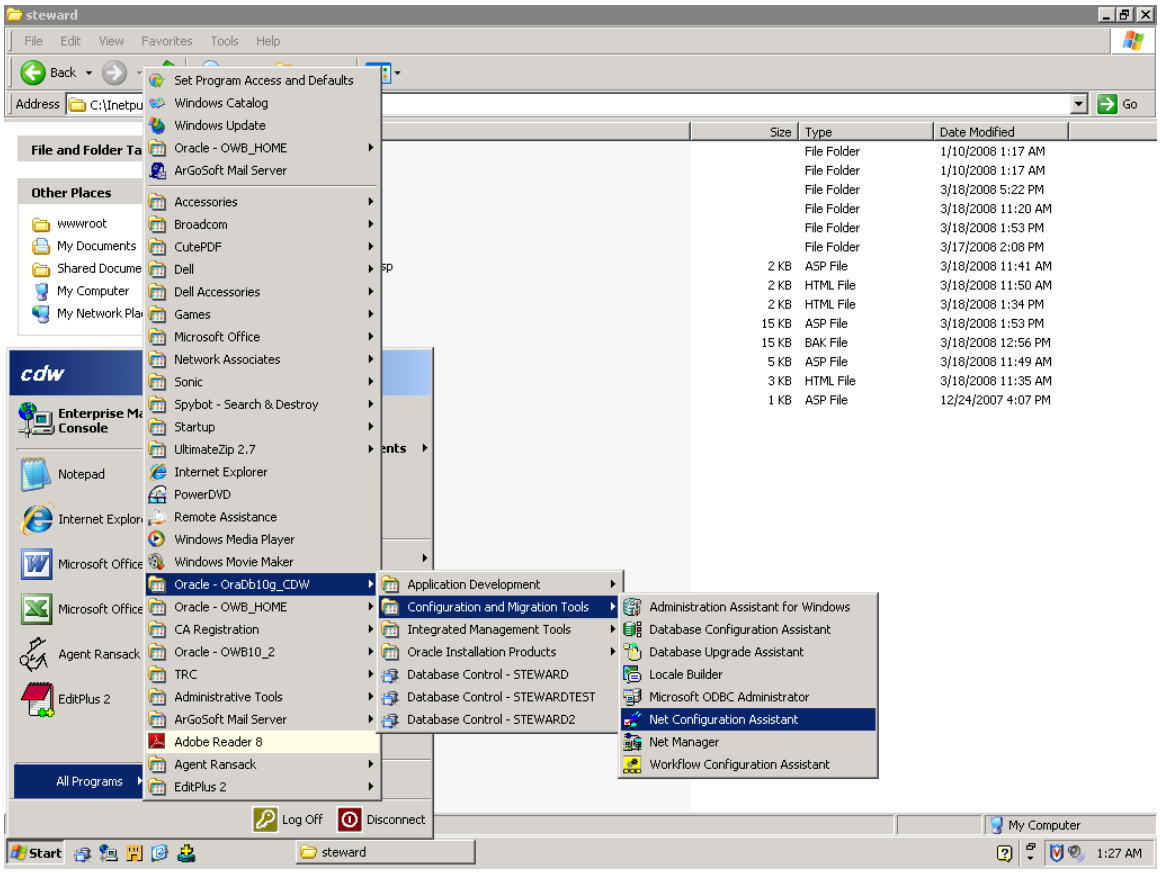


Figure 117 STEWARD web installation - Web Configuration Assistant

4.1 Net Configuration Assistant Steps

The following steps must be performed using the Oracle Net Configuration Assistant.

Step1: As shown in Figure 118, select local net service name configuration.

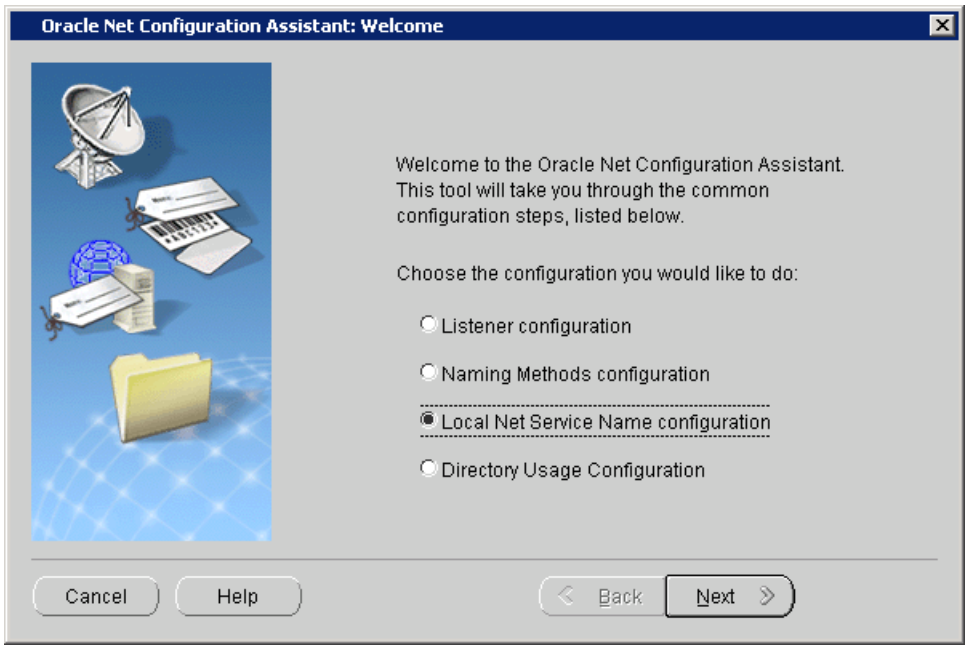


Figure 118 STEWARD web installation - Oracle net configuration (Step 1)

Step2: As shown in Figure 119, select service name as 'Steward2'.

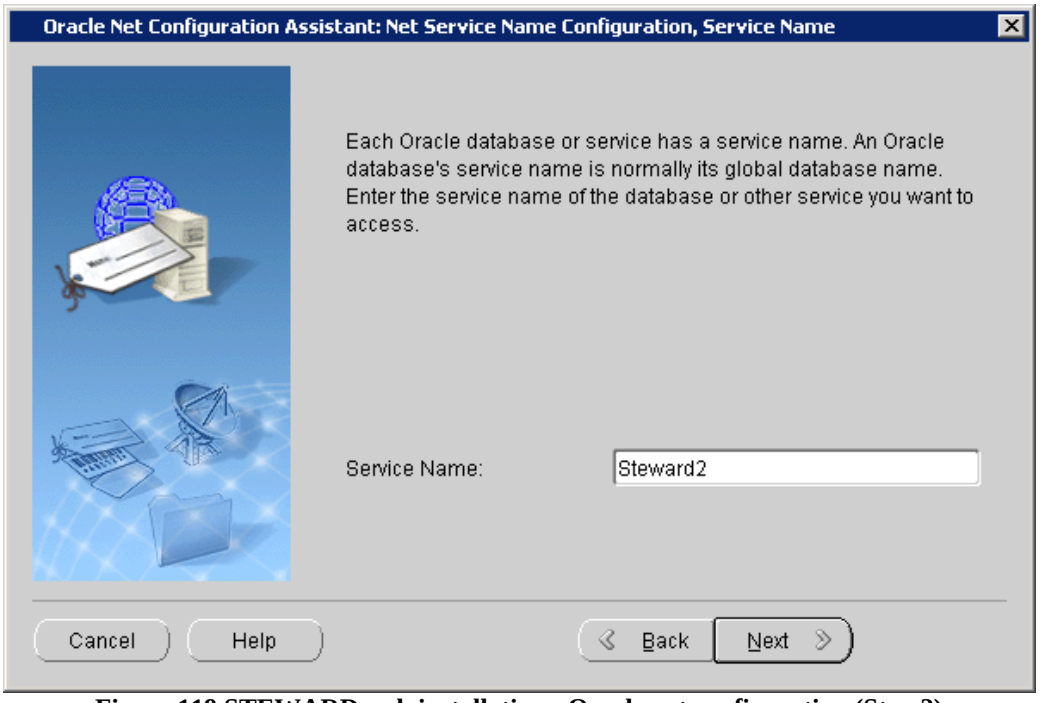


Figure 119 STEWARD web installation - Oracle net configuration (Step 2)

Step3: As shown in Figure 120, select the communication protocol as TCP.

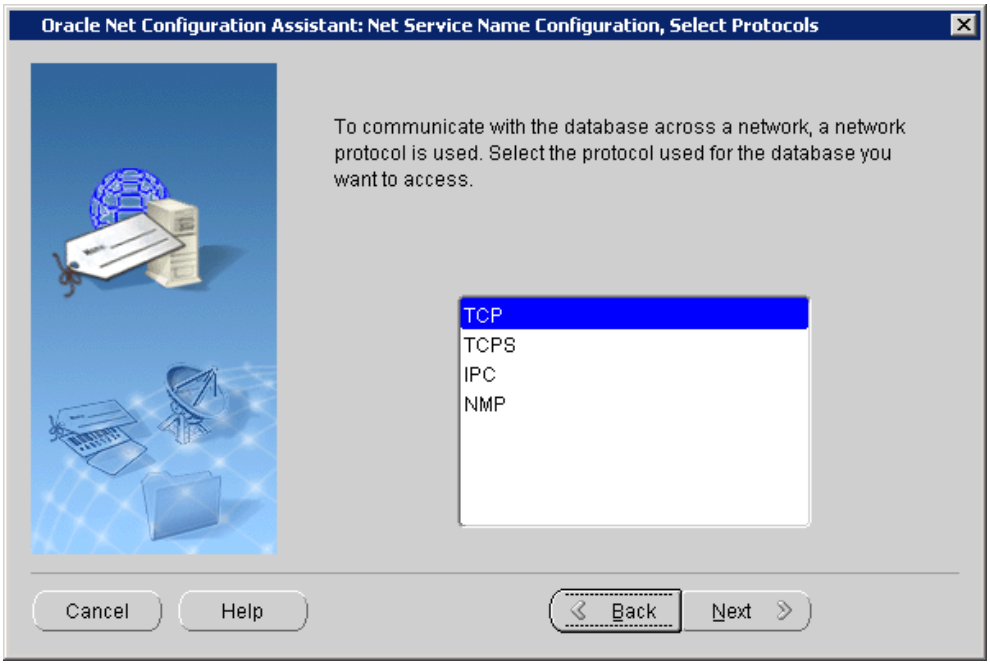


Figure 120 STEWARD web installation - Oracle net configuration(Step 3)

Step4: As shown in Figure 121, edit the host name as 'cdwserver.ce.ufl.edu.'

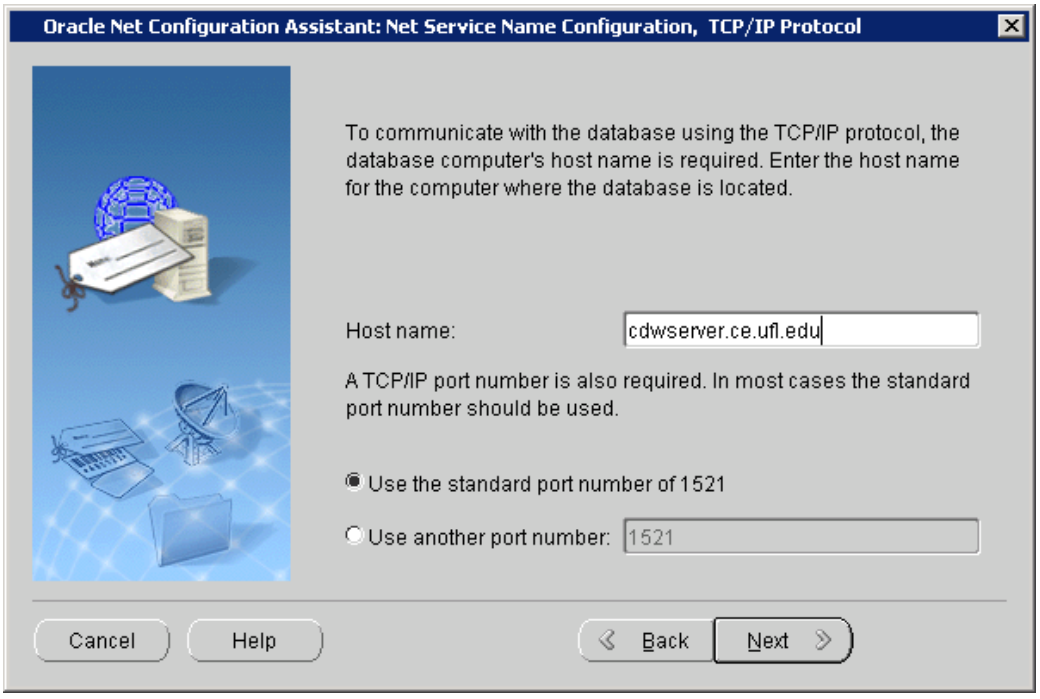


Figure 121 STEWARD web installation - Oracle net configuration(Step 4)

Step5: As shown in Figure 122, perform the test for the new Oracle net configuration

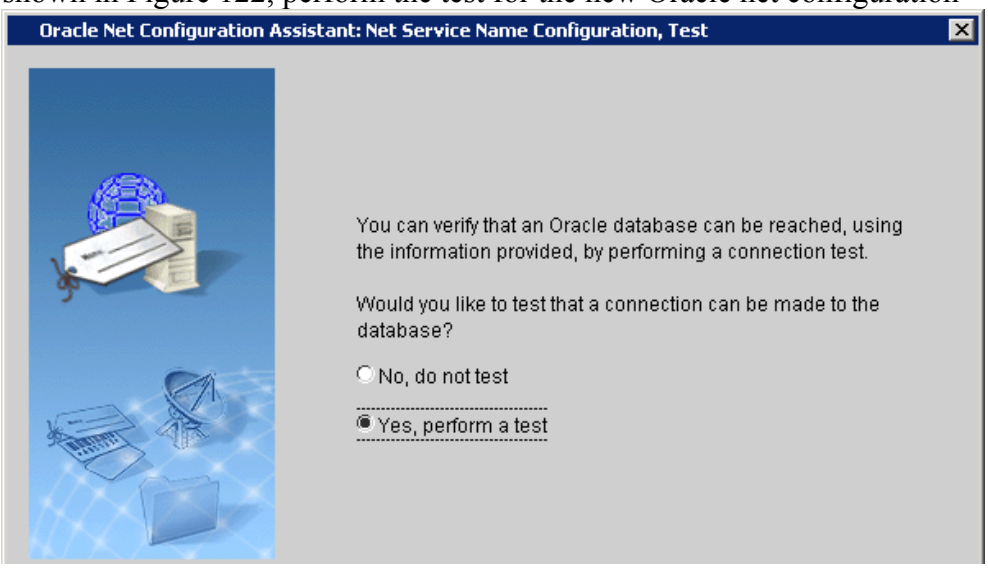


Figure 122 STEWARD web installation - Oracle net configuration(Step 5)

Step 6: If the test fails as Figure 123 for invalid login ID or password, change the login ID and password by clicking *change login*.

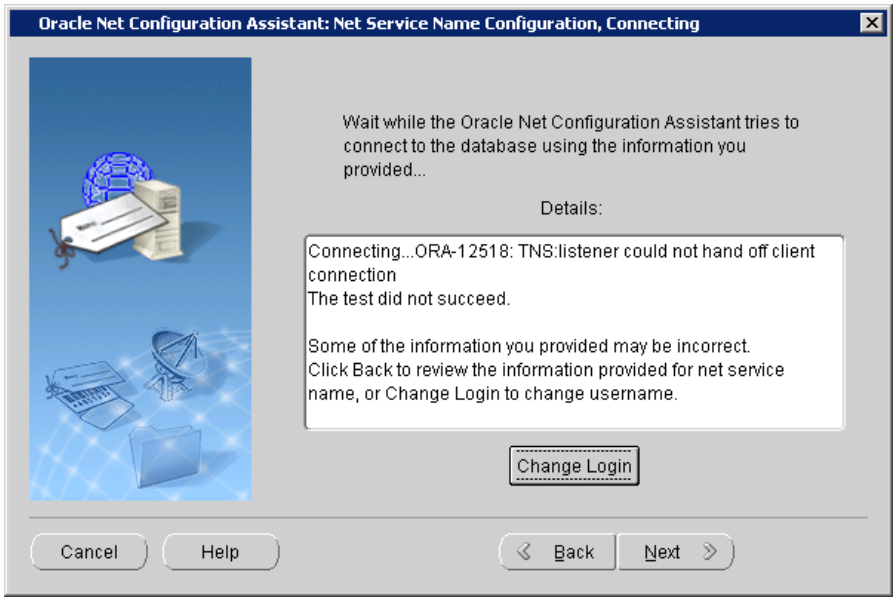


Figure 123 STEWARD web installation - Oracle net configuration(Step 6)

Step 7: As shown in Figure124, set the user name/password as “gator/trc513”

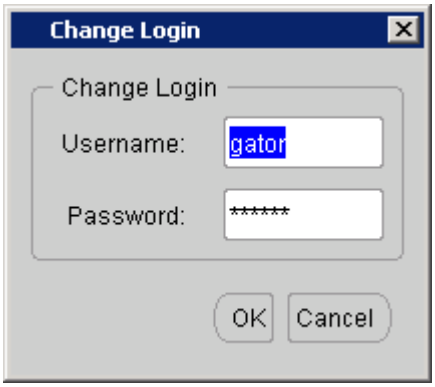


Figure 124 STEWARD web installation - Oracle net configuration(Step 7)

Step 8: As shown in Figure 125, the remote login should be successful.

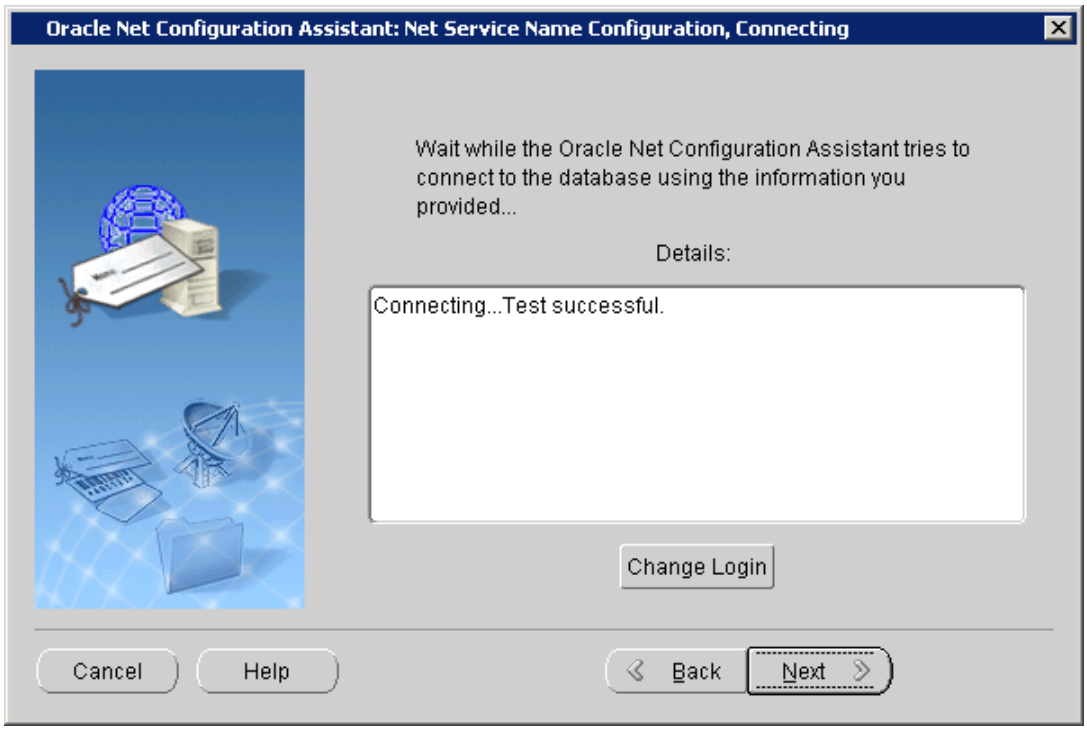


Figure 125 STEWARD web installation - Oracle net configuration(Step 8)

Asp codes and Java Scripts don’t need any specific installation procedures. The program works as the source codes are copied to the directory (steward) as the web root directory. In case of Microsoft Internet Information Services version 5.1, the default root directory is located at C:\Inetpub\wwwroot and the web program need be copied to the STEWARD directory.

4.2 STEWARD Web Program Installation

There are no special installers or routines for the STEWARD web program. The program itself needs be extracted to the specific directory (C:\Inetpub\wwwroot: Root directory for Microsoft Internet Information Services). If the database connection is successful, the STEWARD Web program will respond to the remote Microsoft Internet Explorer.

4.3 System Configuration

After the STEWARD web program is installed on web server, the system configuration needs to be updated. There are two main configuration operations. One is to enable the web server on the system firewall. The other is to add the sharing permission for public users to access the traffic data directory

4.3.1 Firewall Setting

The STEWARD web server is installed on Windows XP Service Pack 2 with the default firewall. In this case, the local area connection needs be updated to enable the web service. The procedures are as follows.

1. Open the Windows firewall icon from windows control panel as shown in Figure 126.

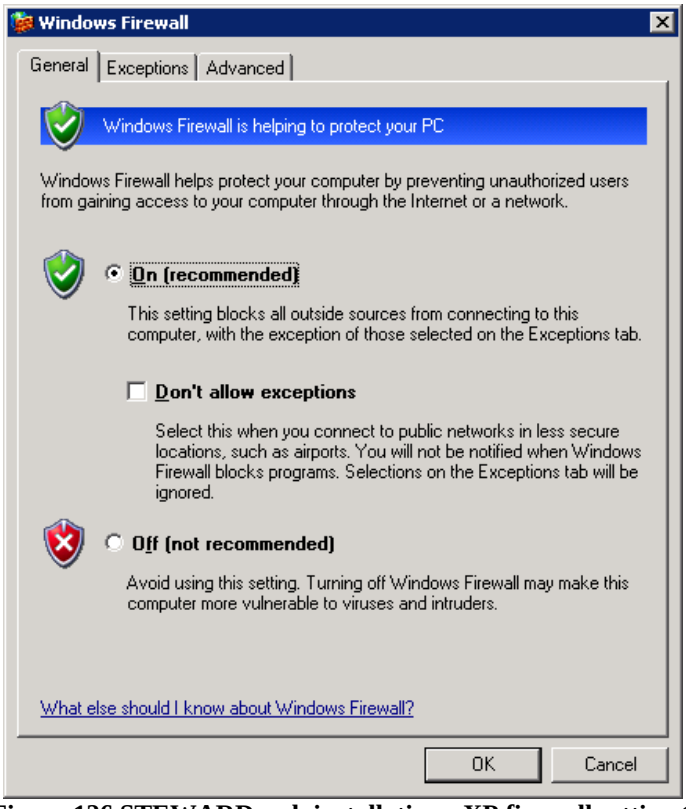


Figure 126 STEWARD web installation - XP firewall setting 1

2. Select the ***Advanced*** tab as shown in Figure 127.

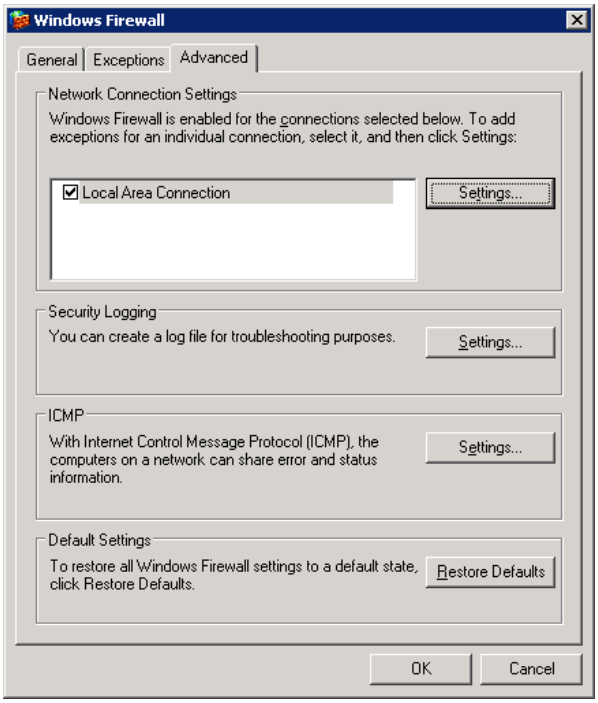


Figure 127 STEWARD web installation - XP firewall setting 2

3. Settings for the Local Area Connection
Check *Web Server* (HTTP) as shown in Figure 128.

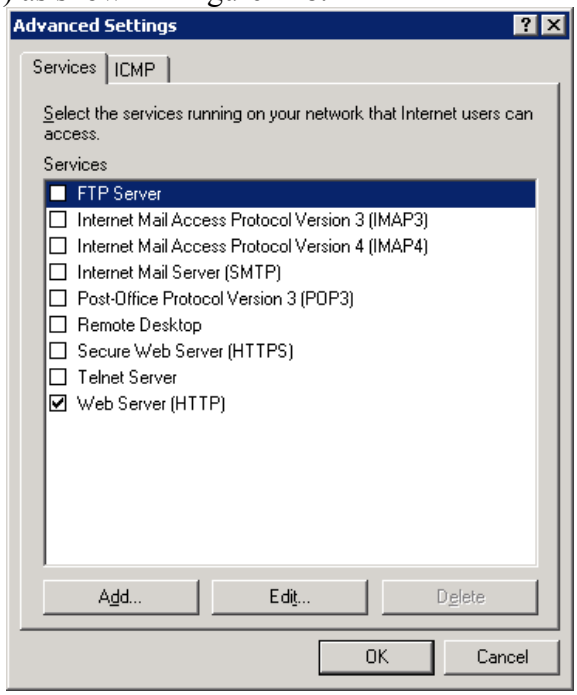


Figure 128 STEWARD web installation - XP firewall setting 3

4.3.2 Permission for File Sharing

The files folder is located in the STEWARD root directory (wwwroot\Steward\files). This directory needs to update the web sharing options for the web user to open or download data file. The procedure is as follows.

Open the *Properties* screen for the *Files* directory as shown in Figure 129.

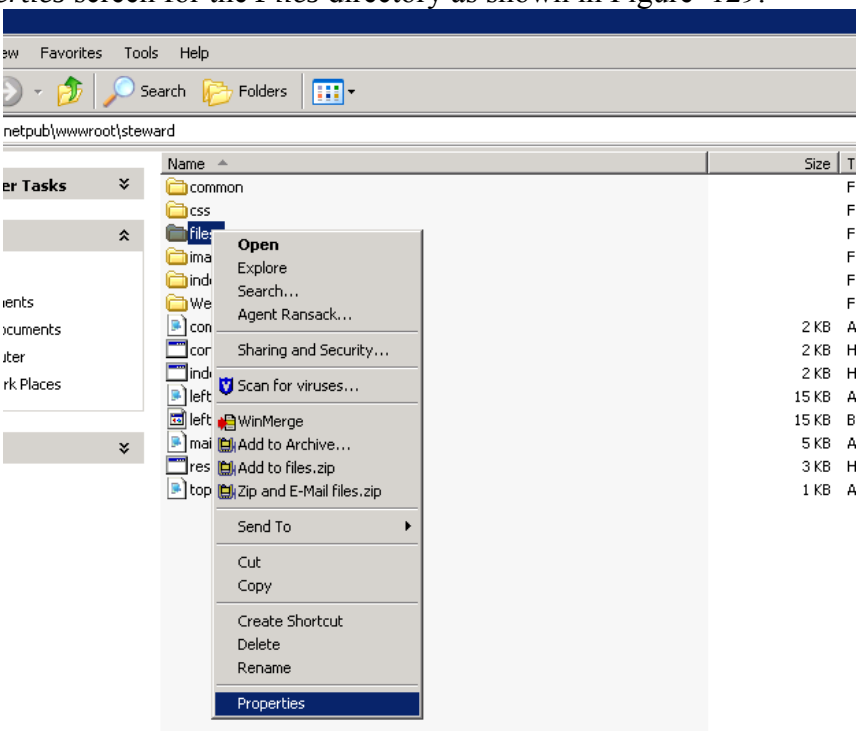


Figure 129 STEWARD web installation - file sharing permission 1

Select the web sharing tab and change the sharing option to allow sharing of this folder as shown in Figure 130.

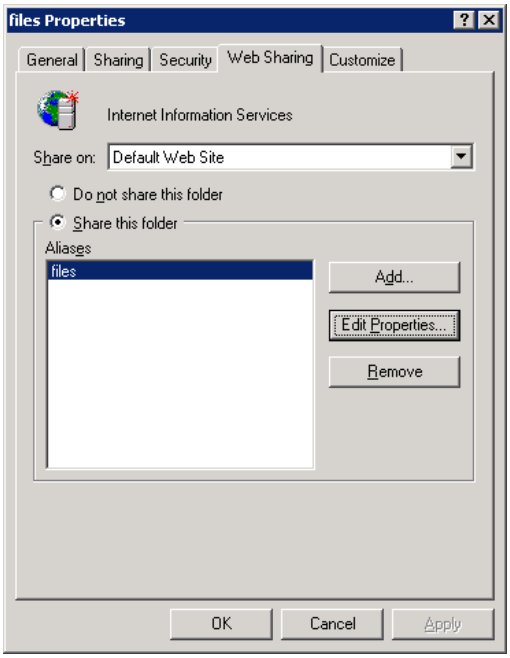


Figure 130 STEWARD web installation - File sharing permission 2

Edit properties as shown in Figure 131.

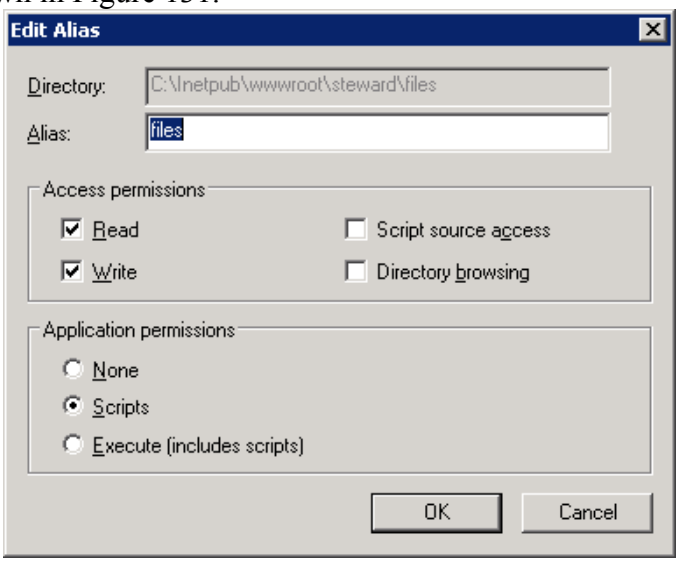


Figure 131 STEWARD web installation - File sharing permission 3

This will complete the set up of the STEWARD Internet access

4.4 Web Program Configuration

After the web program is ported into another web server, TRC-dedicated features need to be updated.

4.4.1 Web Server Address

The server address within the web source needs to be updated. Currently, the server address is defined as

server_url = http://cdwserver.ce.ufl.edu/

in the \Steward\common\asp\server_url.asp. This needs to be updated as necessary.

4.4.2 Google Map API Key

Google maps is used in STEWARD map pages and they need an API key from the Google web site. One Google map API key is assigned on each web server. STEWARD at TRC has the following Google map API key in the map pages:

```
<script
src="http://maps.google.com/maps?file=api&amp;v=2&amp;key=ABQIAAAAgV0e4ari6PxwJ
U39zhTFgRSmDFlu1ewgPW9_Vza0ZTxNM8JDAxRLEpwQSW4FnFN5h39b4jvThnSlMw"
  type="text/javascript"></script>
```

After migration, new key needs be replaced for the following map pages.

\Steward\maps\TSS_interactive_map.html
\Steward\maps\dX_interactive_map.html

4.4.3 STEWARD DB Login Information

DB access information is defined at the following files.

\Steward\common\asp\db_open.asp
\Steward\common\asp\db_close.asp

5 STEWARD Management

5.1 STEWARD Daily Operation

STEWARD receives archive data from SunGuide systems in District 2,4,5,6 and 7 everyday. All data are processed and loaded into STEWARD for users to access the various reports.

Figure 132 shows an overview of the ETL process. STEWARD consists of two servers – an FTP server and a Database/web server. The FTP server receives the data from each district; backs up the files, and processes them. The Database/web server receives the processed data from the FTP server, loads them into the STEWARD database, and creates the user reports. It also supports the web user interface.

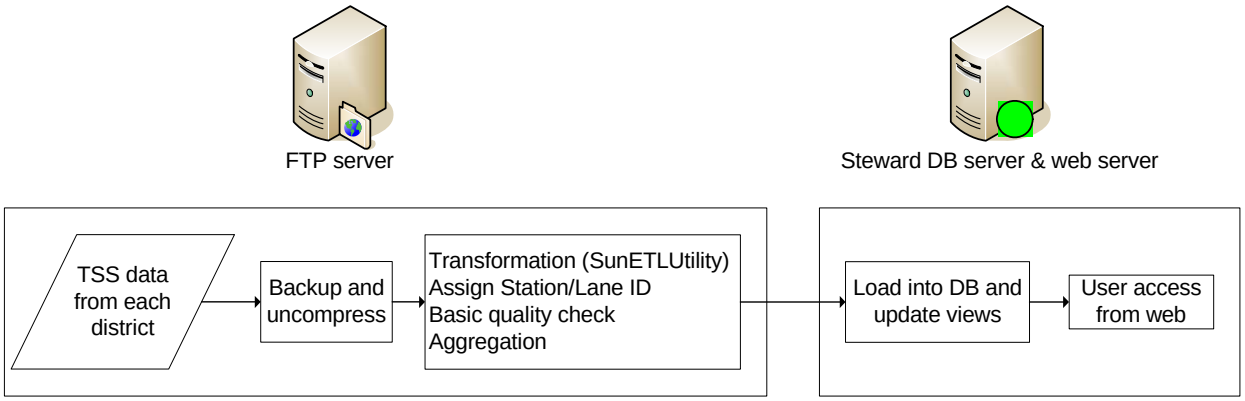


Figure 132 STEWARD system architecture

There are at least two computers are used in STEWARD. Extraction and transformation is done in one computer and loading is done in the other computer. Extraction and transformation is done using the Windows Scheduler and loading is done using the Oracle Scheduler.

5.1.1 Data Transfer from District SunGuide Systems

STEWARD receives TSS data from district SunGuide systems. Each district uses the SunGuide data archive subsystem to create archive files everyday at 11:59 PM. The archive files have predefined file names and formats. The file name format is *'TSS-MMDDYYYY—1.CSV.'* The contents are in a text file with a comma separator. The file names do not include the district ID and the STEWARD operator needs to identify each data file with caution to avoid confusion among districts. These files are archived with a zip utility and sent to the STEWARD FTP site.

- STEWARD FTP address: cdw880.ce.ufl.edu
- STEWARD FTP ID: cdw_user_d*x* (*x* denotes the district ID)

These files are pushed into the STEWARD FTP site on the next day. The following list shows the transfer time for each District.

- District 2 TSS data are posted at 3:30 AM next day
- District 4 data are posted at 2:00 AM next day.
- District 5 data are posted at 12:00 PM next day.

- District 6 data are posted at 2:00 AM next day.
- District 7 TSS data are posted at 3:00 AM next day

These files are archived into STEWARD FTP server at the following folders.

- District 2: C:\Steward_FTP\District2
- District 4: C:\Steward_FTP\District4
- District 5: C:\Steward_FTP\District5
- District 6: C:\Steward_FTP\District6
- District 7: C:\Steward_FTP\District7

5.1.2 Data Backup and Transformation in the STEWARD FTP Server

Scripts have been developed to support data backup and transformation using VBScripts. These scripts are designed to be run with Windows scheduler at the predefined time.

- District 2: C:\Steward_FTP\Steward_ETL_D2.vbs
- District 4: C:\Steward_FTP\Steward_ETL_D4.vbs
- District 5: C:\Steward_FTP\Steward_ETL_D5.vbs
- District 6: C:\Steward_FTP\Steward_ETL_D6.vbs
- District 7: C:\Steward_FTP\Steward_ETL_D7.vbs

These scripts were developed for each District and have the same content. The detailed contents are as follows.

- Backup the archived files into two backup folders.
  - C:\Steward_FTP\District2\Backup\2009
  - U:\DISTRICT_DATA\District2\RawData
- Unzip the archive files and archive the raw data file.
  - U:\DISTRICT_DATA\District2\RawData\TSS\2009
- Transform the input raw data file into STEWARD input format using SunETLUtility. This program is described in detail in Appendix 4a. During the transformation, one input TSS data file would create the 10 new output files and update one status file. For example, one sample input data file, *TSS-03032009--1.dat* from District 2 would be transformed into the following output files
  - ConvertedData\D2_TSS-03032009-0.csv
  - ConvertedData\D2_TSS-03032009-DailyReport.Log
  - GroupData\D2_TSS-03032009-1S.csv
  - GroupData\D2_TSS-03032009-1L.csv
  - GroupData\D2_TSS-03032009-5S.csv
  - GroupData\D2_TSS-03032009-5L.csv
  - GroupData\D2_TSS-03032009-15S.csv
  - GroupData\D2_TSS-03032009-15L.csv
  - GroupData\D2_TSS-03032009-60S.csv
  - GroupData\D2_TSS-03032009-60L.csv
  - TSSConversionHistory-D2.csv (updated)

- Backup the transformed 10 output files into the following backup folders for each district
  - U:\DISTRICT_DATA\District2\ProcessedData\ConvertedData\2009
  - U:\DISTRICT_DATA\District2\ProcessedData\GroupData\2009
- Send an email to STEWARD Admin with the attachment of each daily report log output file.

Three kinds of data backups exist for the incoming data: Raw data backup, unzipped data backup and processed data backup. Only processed data files are used for the STEWARD data loading. Drive U: is used as the network drive to share the data between the FTP server and DB/Web server.

5.1.3 Data Loading into the STEWARD Database

Processed data from the ETL utility are ready to be loaded into the STEWARD database. In the DB/Web server, the Oracle Warehouse Builder Design Center and Oracle Enterprise Manager are mainly used to manage STEWARD processes and schedules.

The DB server retrieves the processed data from the processed-data backup folder and loads them into the STEWARD database. The details of this process is described in detail in Appendix 1 of the Phase 2 final report.

As TSS data are ready to be loaded into STEWARD, the administrator can load these data into using the Oracle Warehouse Builder Design Center. The basic loading procedure is defined as the mappings for each table. Figure 133 shows the mappings defined in STEWARD

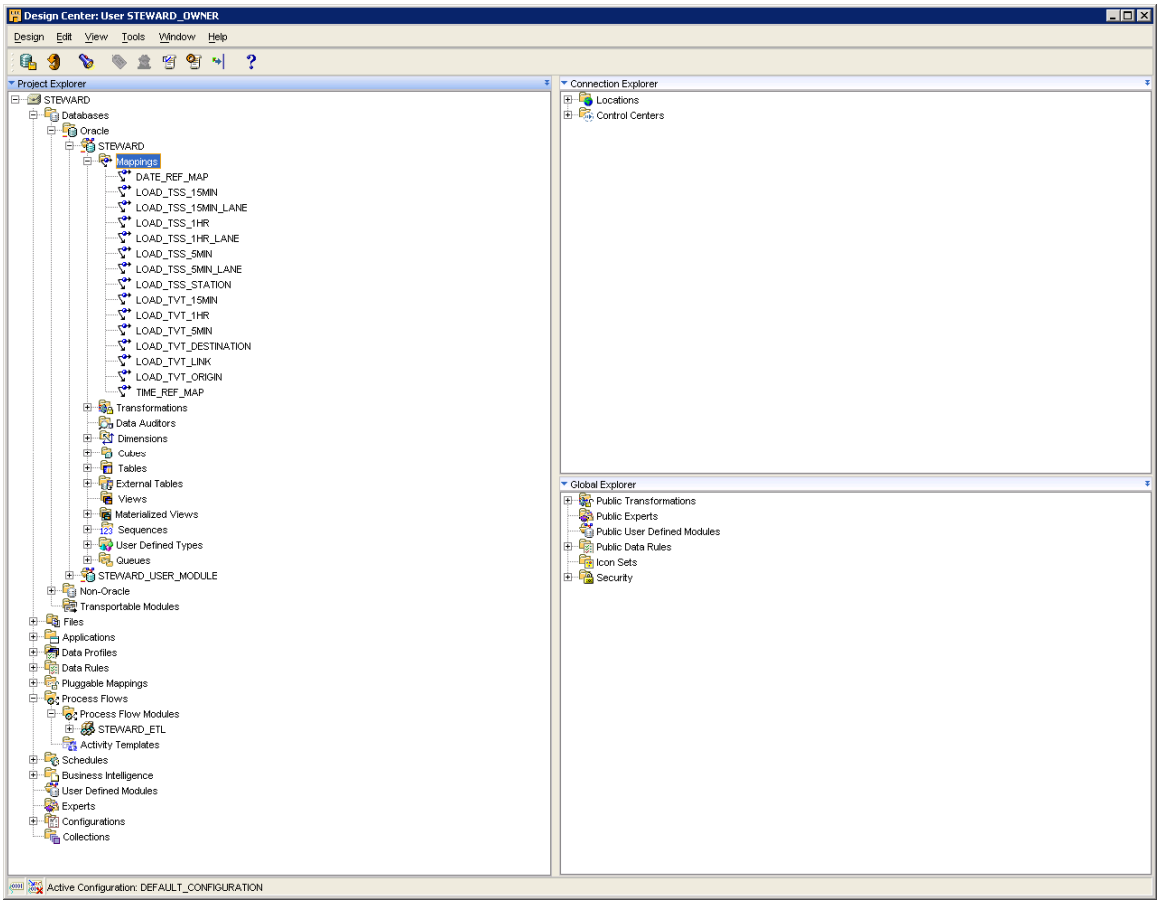


Figure 133 Mapping modules for STEWARD

The following mappings are used to load single TSS files for each aggregation level.

- LOAD_TSS_5MIN
- LOAD_TSS_15MIN
- LOAD_TSS_1HR
- LOAD_TSS_5MIN_LANE
- LOAD_TSS_15MIN_LANE
- LOAD_TSS_1HR_LANE

These mappings are used as the objects in Oracle Warehouse Builder process. With this object, a process can be created and run to load the multiple files.

- PRELOAD process is used to move the processed files into predefined STEWARD folders for loading.
- LOADING_TSS process is used to run three processes: LOADING_TSS_5MIN, LOADING_TSS_15MIN, LOADING_TSS_1HR
- LOADING_TSS_LANE process is used to run three processes: LOADING_TSS_5MIN_LANE, LOADING_TSS_15MIN_LANE, LOADING_TSS_1HR_LANE

The following processes run the predefined mapping until all the files in the folders are loaded.

- LOADING_TSS_5MIN, LOADING_TSS_15MIN, LOADING_TSS_1HR processes are used to run the LOAD_TSS_5MIN mapping to load all the TSS 5min files in the folders
- LOADING_TSS_5MIN_LANE, LOADING_TSS_15MIN_LANE, LOADING_TSS_1HR_LANE processes are used to run the LOAD_TSS_5MIN mapping to load all the TSS 5min files in the folders.

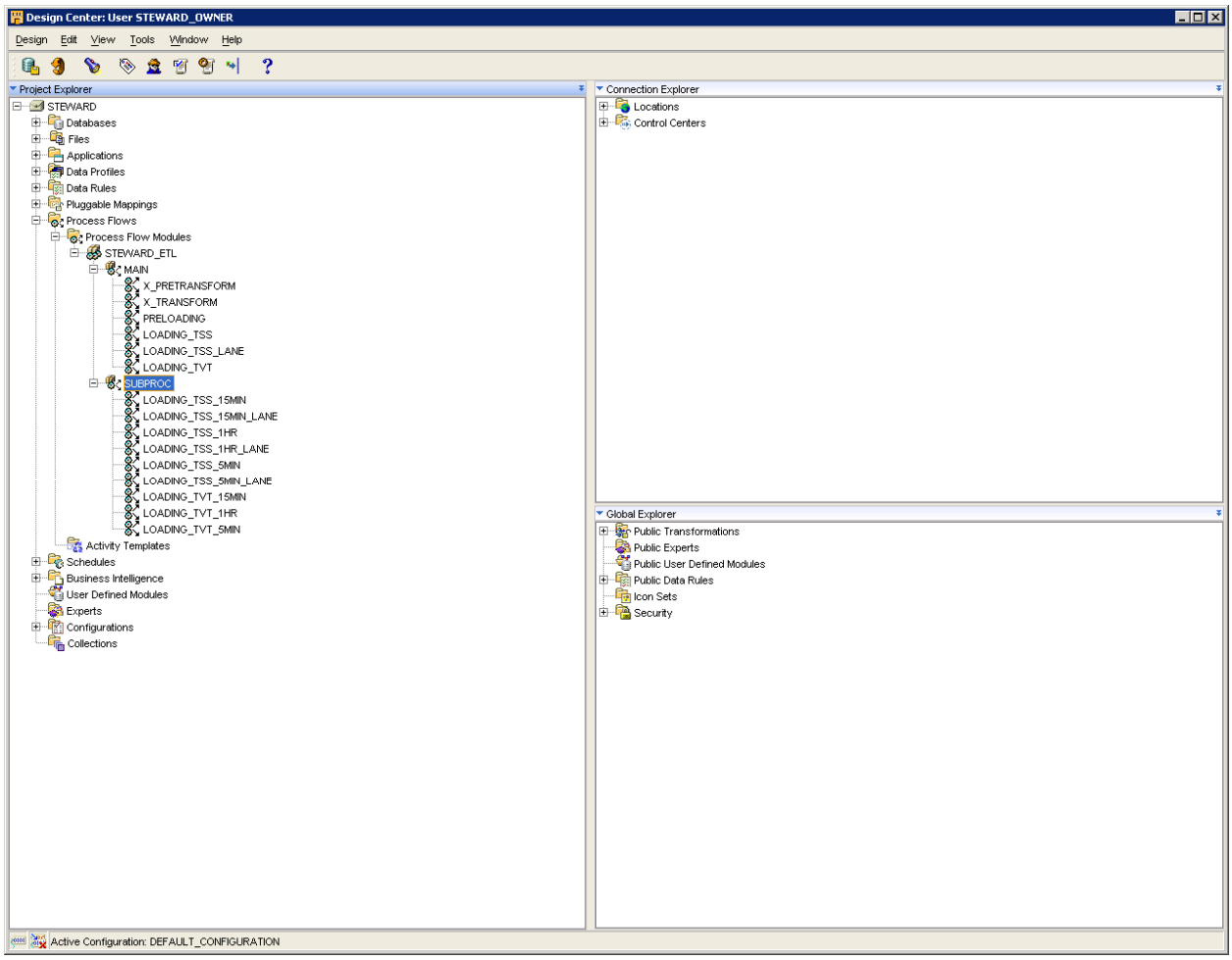


Figure 134 Process flow modules for STEWARD

Figure 134 shows the Process Flow modules defined in STEWARD. These processes can be executed at the predefined time using Oracle schedules or manually using Oracle Warehouse Builder Control Center.

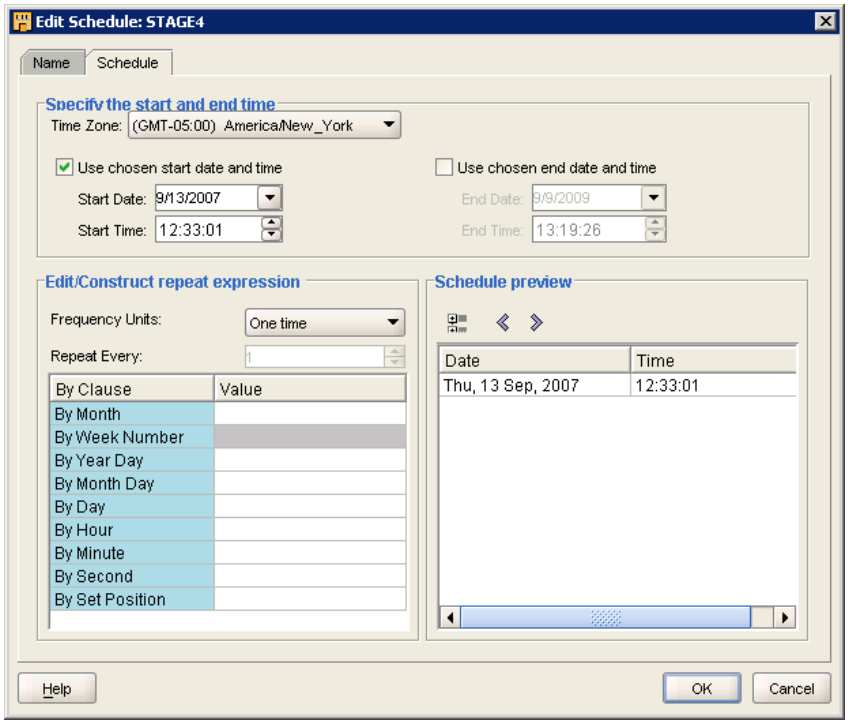


Figure 135 Oracle schedule sample for STEWARD

Figure 135 shows an Oracle Schedule sample. The Oracle Process Flow Version 10.2g has a known issue that the scheduled process flow runs only once even though the user sets it multiple times (Bug ID 6690762). Because of this, STEWARD data loading uses the Oracle Warehouse Builder Control Center.

Figure 136 shows a screen capture of Oracle Warehouse Builder Control Center.

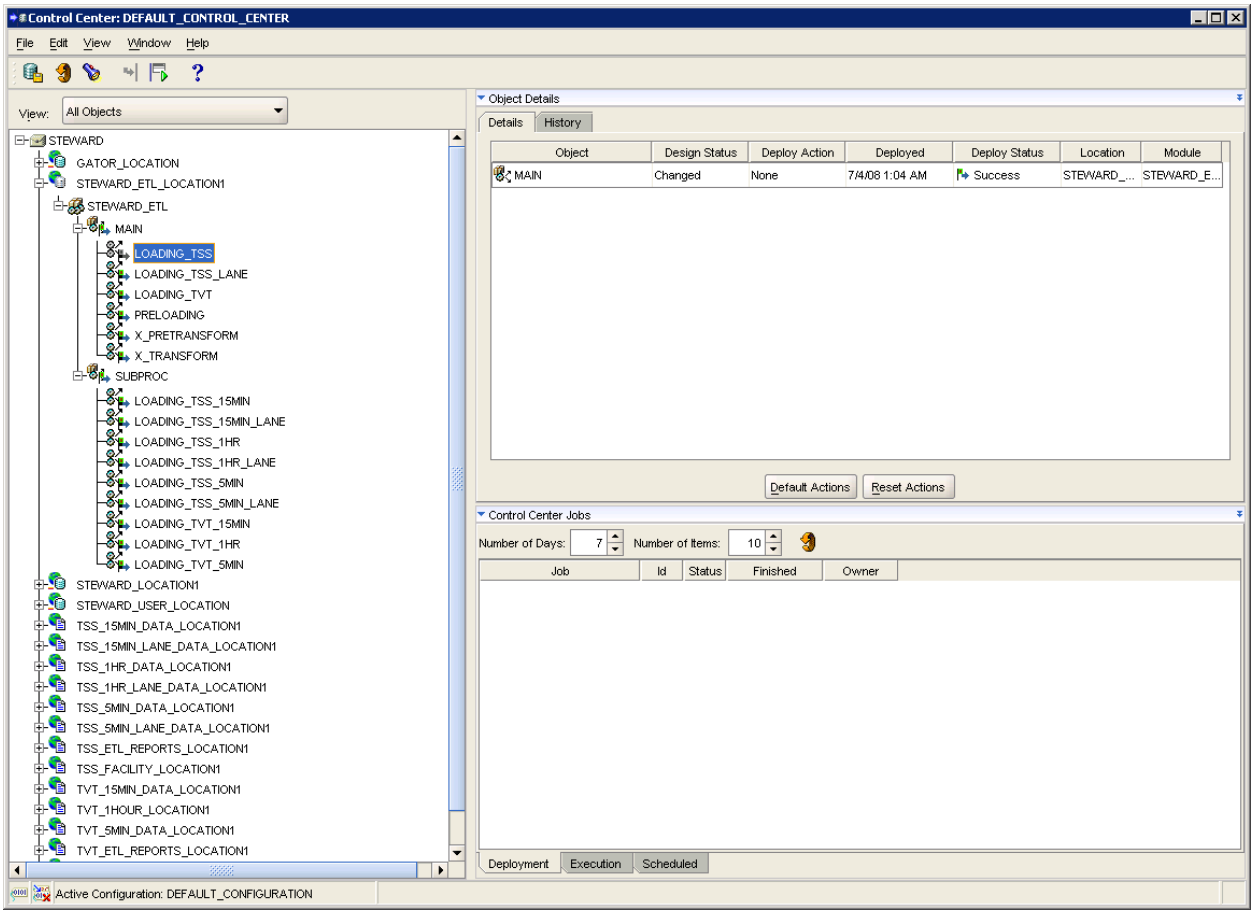


Figure 136 Oracle Warehouse Builder Control Center

5.1.4 Refresh Configuration for Materialized Views in the STEWARD Database

After all of the processed data are loaded into STEWARD, the materialized views for the reports need to be updated. There are two issues on this update. One is the *refresh-on* method and the other is *refresh* method.

As shown in Figure 137, there are two *refresh-on* methods for the materialized views - Demand or Commit. These are how to start the materialized views update when the master tables are changed.

The commit method updates materialized view every time the master tables are created or updated. The demand method does not update materialized view until the STEWARD users update them manually.

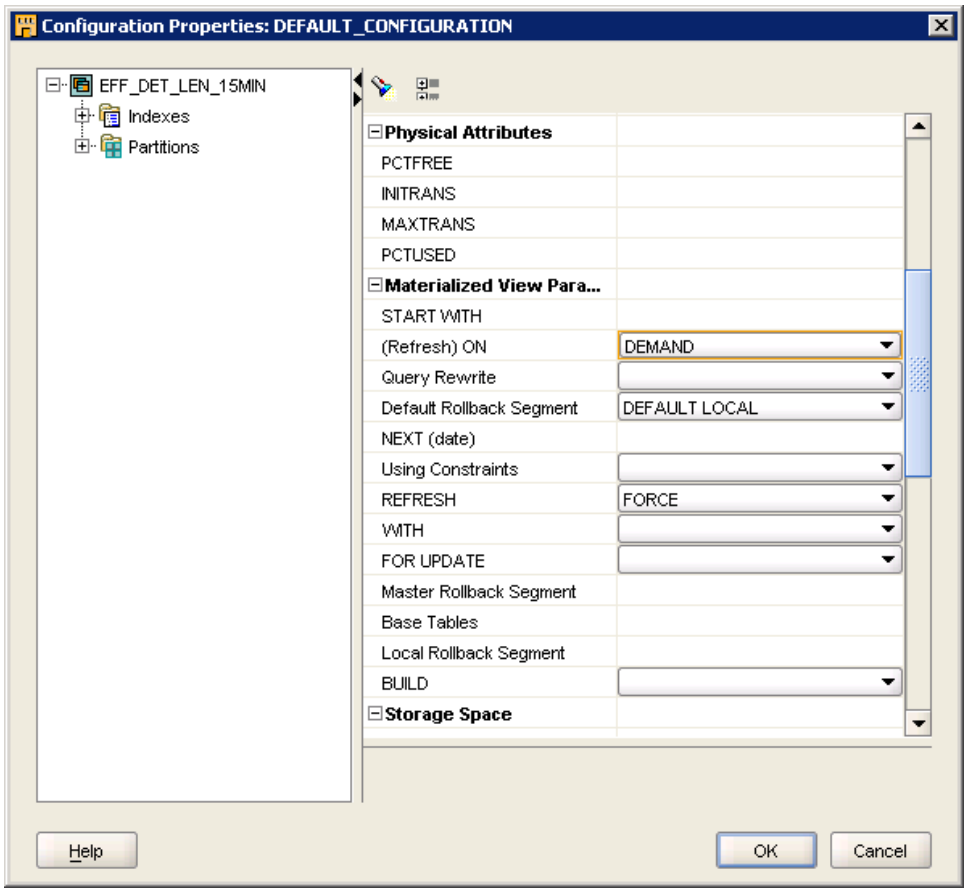


Figure 137 Refresh settings for materialized views

At this time, the refresh-on methods for all of the materialized views are set with the on-demand method. The reasons are that the on-commit method updates materialized views every time there are changes in the master tables. Therefore loading TSS data files take more time than the on-demand method. The On-demand method updates the materialized views after all of the data are loaded into the master table. The second reason is that the Oracle database does not allow using on-commit refresh-on method when the materialized views use aggregate functions, such as AVG, SUM (ORA-12054). This is because these aggregate functions need to be recalculated every time after each update on the master table and this causes severe overhead problems. Materialized view "PM_5MIN" uses the aggregate function SUM in queries and would not be allowed to use the on-commit query.

STEWARD materialized views might be expected to change from on-demand method into on-commit refresh method when the STEWARD system is migrated into a high performance server.

The refresh method would be another issue for the materialized views. There is one of four options for each materialized view, which are *Fast*, *Complete*, *Force* or *Never*. The details are described in Oracle Warehouse Design Center as follows.

> *Specify FAST to indicate the incremental refresh method. Specify COMPLETE to indicate the complete refresh method, which is implemented by executing the defining query of the materialized view. Specify FORCE to indicate that when a refresh occurs, the Oracle*

Database will perform a fast refresh if one is possible or a complete refresh otherwise. FORCE is the default. Specify NEVER to prevent the materialized view from being refreshed with any Oracle Database refresh mechanism or packaged procedure.

Even though STEWARD selects the FORCE option for all of the materialized views, it does not do the incremental refresh and performs the complete refresh all the times. Therefore, when there are changes in the master files, the corresponding materialized views always need to be rebuilt entirely. These issues need be reviewed carefully to improve the system performance.

5.1.5 Update the Materialized Views in the STEWARD Database

The materialized views can be updated using the Oracle Data Warehouse Control Center. As Figure 138 shows, each materialized view can be deployed with the deploy action, "Replace." It will delete the previous one and create new tables as defined.

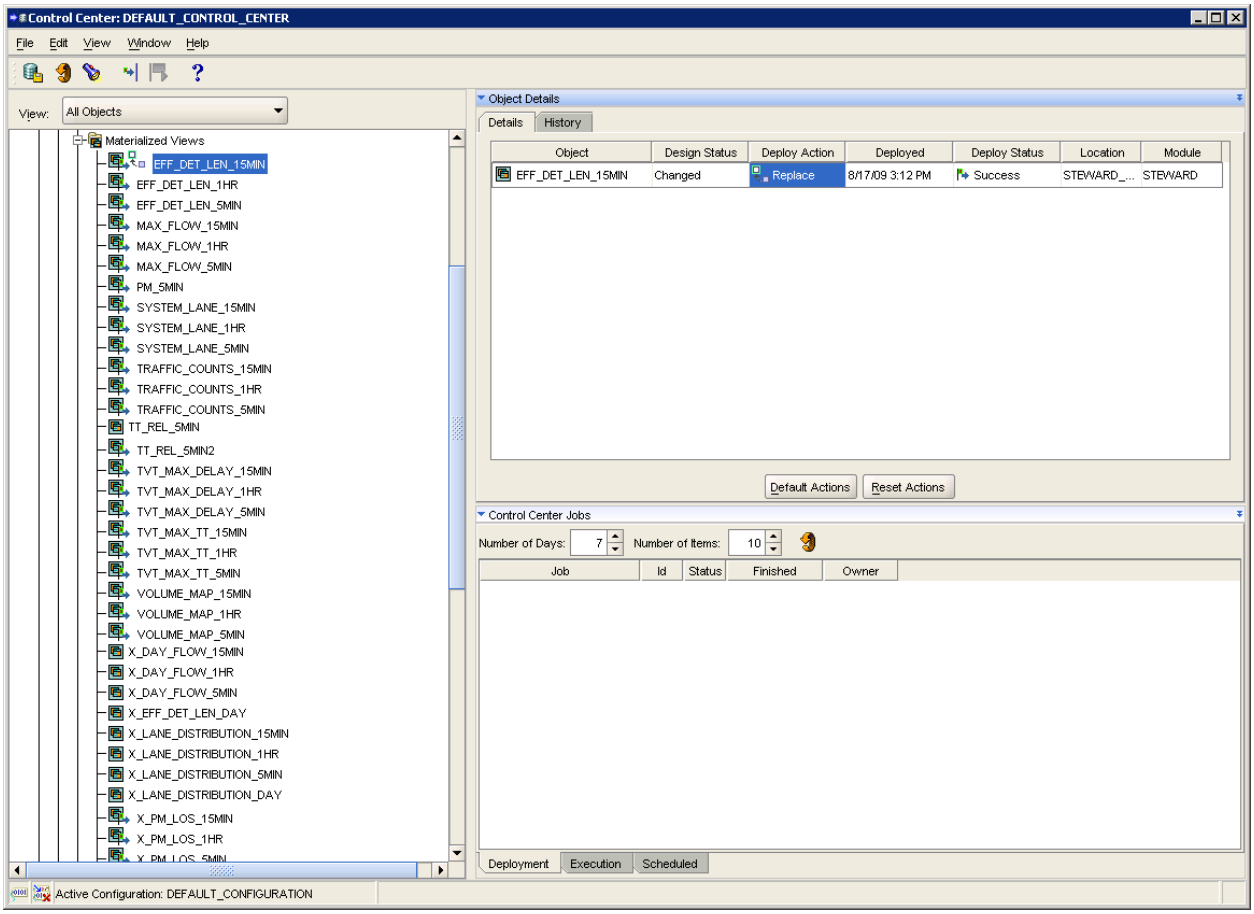


Figure 138 Creating STEWARD materialized views

5.1.6 Backup Plan and Procedure for the STEWARD Database and Web

STEWARD has two types of data to back up. One is for the traffic data from each district, which are processed and archived in each step. The other is the STEWARD system data which are archived for system maintenance and migration.

Traffic data backup

Three types of traffic data - raw format data from each district, processed data from the ETL utility and the Oracle database are backed up and archived. These data backups are done automatically using the system or Oracle Scheduler automatically.

- The raw data and processed data are archived using VB script into the data backup storage (U:\FTP_Backup for raw data, U:\DISTRICT_DATA\DistrictX\ProcessedData for processed data).
- The processed data are archived to support requests from other research projects. These data include 1-minute station or lane data which are not published on the STEWARD website but are useful for research purposes. Also, when STEWARD would migrate into new database server, these data are expected to be reloaded.
- Oracle database are backed up daily using the Oracle Enterprise Manager (EM). It is executed every day using the Oracle EM backup function. Figure 139 shows a screen capture of the Oracle database back reports. (D:\Steward_backup\DB\flash_recovery_area\STEWARD\DATAFILE)

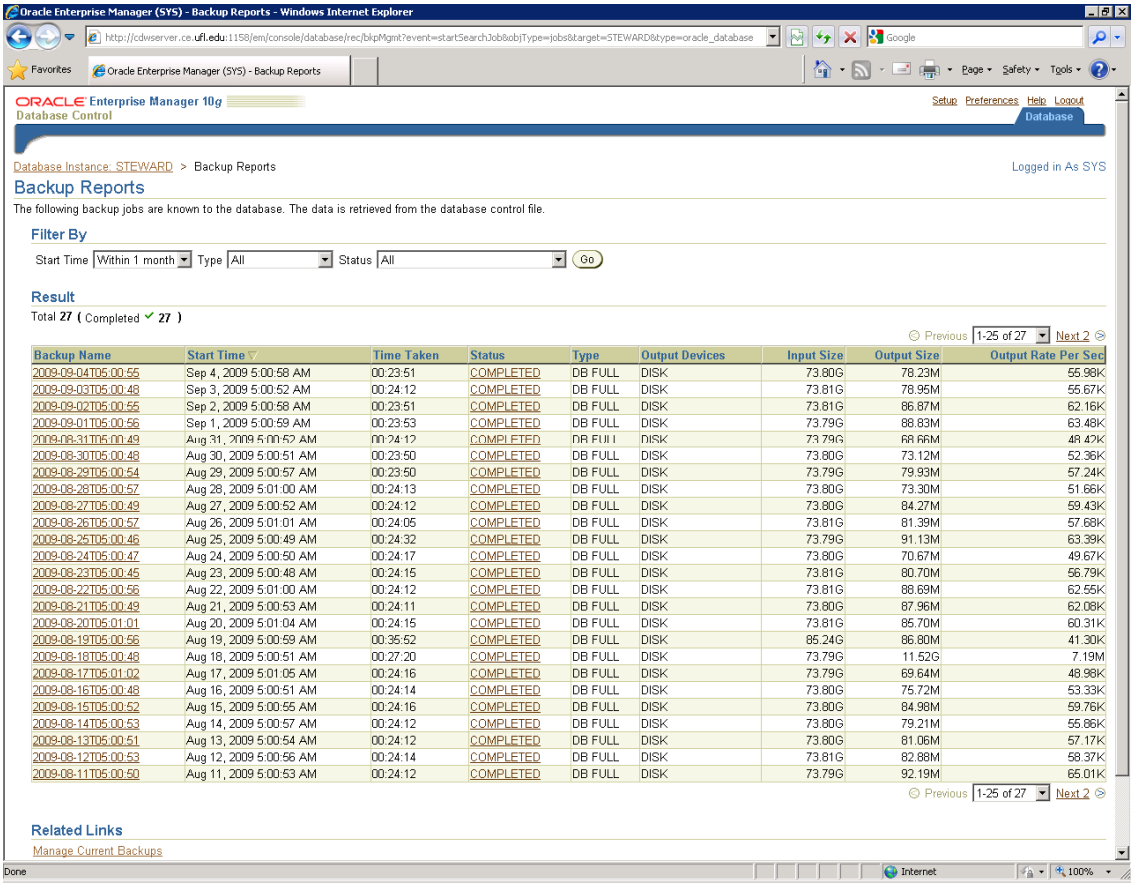


Figure 139 STEWARD database backup reports

STEWARD system backup

Three types of STEWARD system components - STEWARD FTP directory, STEWARD warehouse configuration and STEWARD web pages are backed up. These backups are executed manually when required.

- When there are structural changes in the STEWARD FTP directory, the entire directory is manually backed up before the changes. This would create a duplicate copy of theraw data from the districts (D:\Steward_backup\FTP).
- STEWARD system configuration – DB configuration and operations, process and schedules are designed and managed by the Oracle Warehouse Builder. All of this information is archived using the *Export Warehouse Builder Metadata* function in the Oracle Warehouse Builder. This procedure can be executed manually when required as shown in Figure 140. These files are archived in the folder, D:\Steward_backup\OWB

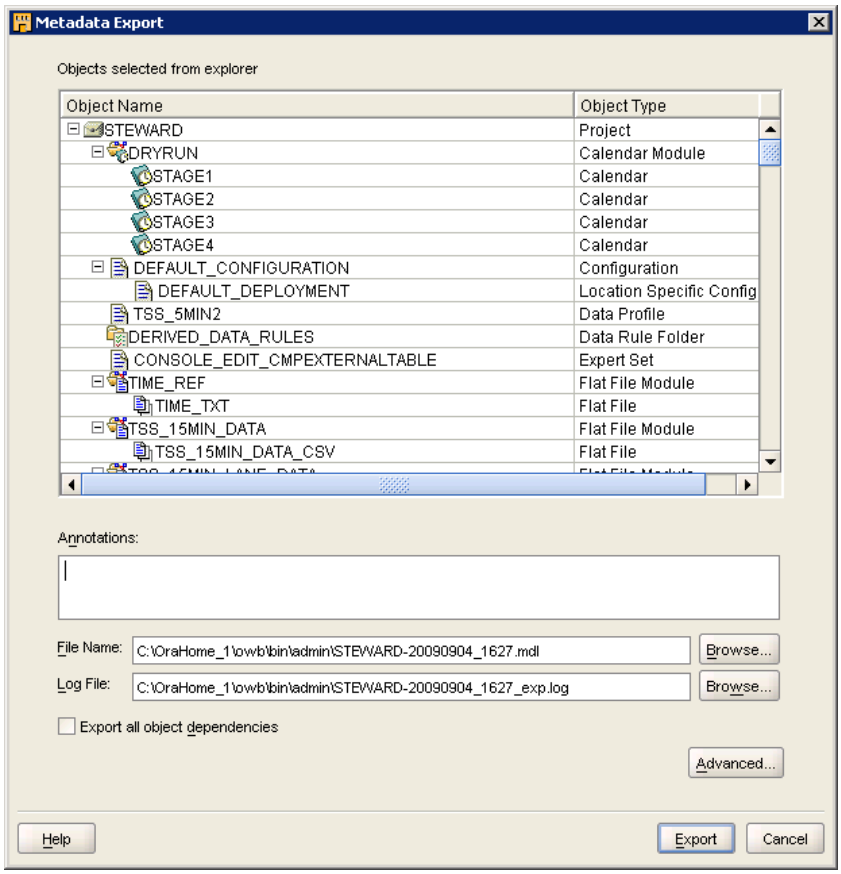


Figure 140 STEWARD warehouse builder metadata export

- The STEWARD web server was developed using ASP, JavaScript, and HTML codes and the entire directories are backed up manually when required to (D:\Steward_backup\Website)

Most of the data backups are done using the two dedicated data storage devices. Most of the traffic data are archived in the U: drive which is shared among the STEWARD computers. The main uses of the U: drive are to back up all of the TSS data before they are loaded into

STEWARD and to share the processed data between the FTP server and DB/Web server. It has a RAID 1 scheme and therefore, mirrors V drive.

The Oracle database and STEWARD system files are archived the local drive, D drive of the DB/Web server. This is a dedicated drive for the STEWARD system operation and data back up.

5.2 STEWARD Operation- Adding a New District Facility

A new district or facility can be easily added to the system. In this section, the procedure to add a new district or facility will be described in three stages. It is assumed that the district/facility data (TSS data) have good quality and their facility information (Station/lane meta data) are ready to create/update the facility data file for the SUN ETL utility.

5.2.1 Changes in ETL Process

The main impact of adding new district or facility would be to guarantee the quality of TSS data and to update the facility data file into good working condition. In general, there needs to be several communications between the TRC and district to receive the new district data. Therefore, the STEWARD administrator can set up to receive the new district data. But in case of new facility data, they could be found in the TSS data files without notification. In this case, the increased number of orphan stations in the ETL report file could be used as the indicator of new facility data in the TSS data file.

There were several issues on the ETL process to include the new district data into STEWARD:

- Basic quality checks are executed on new data. If it fails, STEWARD holds the data until this problem is resolved. One of the issues is negative or zero scan intervals for its TSS data. District 2, 6 and 5 had more than 10% of negative or zero scan intervals for each day. TSS data files are expected to have all the traffic data logged in time order but some records have older time stamps than the previous record (negative scan) or there are duplicated time stamps from the same location at the same time (zero scan). It is important to resolve this issue. In the case of District 2 or 6, updating RTMS driver appears to resolve the issue.

- Facility data files might need to be created, updated and verified. These data would have a great impact on the quality of the STEWARD data. Therefore, roadway configuration data from the FDOT Roadway Characteristic Inventories (RCI) and Google satellite maps have been used as necessary to verify the facility data.

Adding a new district will involve the following actions:

- Updates in the FTP server: Currently, STEWARD is set up for the District 1 to 7 and the Turnpike. It might be necessary to update this configuration.
- Daily archive data transmission scripts (STEWARD_ETL_Dx.vbs): will need to be created for each new district.
- Each district must have a configuration workbook (TSSMaster-Dx.xls) file for its facility data. This file includes all of the station and lane information for the district TMC. This file needs to be created and verified with other data sources, such as FDOT RCI database or Google satellite maps. The detail of the facility data are explained in the Chapter 4 of thePhase 3 final report.

- A new ETL script must be run using the Windows Scheduler

Adding new facility will require the following action:

- Update the facility data file (TSSMaster-Dx.xls) to accommodate all of the stations and lanes in the new facility.
- Run a new or modified ETL script using the Windows Scheduler.

5.2.2 Updates to the STEWARD Database

There are no design changes in the STEWARD database scheme. Updating the TSS facility table (TSS_STATION) would involve adding the new stations and loading the new district or facility data without any modifications.

5.2.3 Updates to the STEWARD Web Interface

There might be two types of changes to incorporate the new district or facility to update the STEWARD web interface and to update the STEWARD web report interface.

Updating the STEWARD web interface would be relatively easy by updating an HTML/ASP program as follows.

- Adding or replacing new/update facility file in the resources page
  Facility file location: `C:\Inetpub\wwwroot\steward\resources`
  Resource web page: `C:\Inetpub\wwwroot\steward\resources.html`
- Adding or replacing new/update facility file in the maps page
  Map file location: `C:\Inetpub\wwwroot\steward\maps`
  Maps web page: `C:\Inetpub\wwwroot\steward\maps.html`
- Activating the new district in the floating menu
  Left pane web page: `C:\Inetpub\wwwroot\steward\left_menu.asp`

Updating STEWARD web reports might some new code in the following folders.

- Update TSS facility level reports
  Main folder for TSS section level reports
  `C:\Inetpub\wwwroot\steward\index_main\TSS_Facility_Level_Reports`
  TSS facility level reports folder
  `C:\Inetpub\wwwroot\steward\index_main\TSS_Facility_Level_Reports\volume_map`
  `C:\Inetpub\wwwroot\steward\index_main\TSS_Facility_Level_Reports\traffic_counts`
  `C:\Inetpub\wwwroot\steward\index_main\TSS_Facility_Level_Reports\lane_volumes`
- Update TSS section level reports
  Main folder for TSS section level reports
  `C:\Inetpub\wwwroot\steward\index_main\TSS_Section_Level_Reports`
  TSS section level reports folder
  `C:\Inetpub\wwwroot\steward\index_main\TSS_Section_Level_Reports\performance_measure`
  `C:\Inetpub\wwwroot\steward\index_main\TSS_Section_Level_Reports\TT_Reliability`

- Update TSS station level reports
  Main folder for TSS section level reports
  `C:\Inetpub\wwwroot\steward\index_main\TSS_Station_Level_Reports`
  TSS section level reports folder
  `C:\Inetpub\wwwroot\steward\index_main\TSS_Station_Level_Reports\volume_data`
  `C:\Inetpub\wwwroot\steward\index_main\TSS_Station_Level_Reports\traffic_counts`
  `C:\Inetpub\wwwroot\steward\index_main\TSS_Station_Level_Reports\max_flow_rate`
  `C:\Inetpub\wwwroot\steward\index_main\TSS_Station_Level_Reports\effective_detector_length`

5.3 Known Issues

There might be several challenges in STEWARD system for the next phase.

- *Dynamic speed limit:* To incorporate the dynamic speed limits for District 5 data. Roadway speed limits are used in STEWARD to calculate the delay and performance but STEWARD does not now accommodate dynamic speed limits.
- *HOT/HOV lanes:* STEWARD might need to handle HOT lane in District 6 or HOV lanes separately for analysis and data processing.
- *Surface roads:* District 5 began to archive surface road data and STEWARD needs to develop the capability to handle the surface road data.
- *System configuration changes:* As the roadway detector configurations are changed (new installation, migration, deactivation, etc), STEWARD needs to follow these changes explicitly.

Appendix 3
STEWARD Web Interface

Table of Contents

1 Introduction 1
2 Overview of the STEWARD Web Interface 1
3 STEWARD Web Architecture 1
4 STEWARD Web Pages 2
4.1 STEWARD Main page –STEWARD Overview 2
4.2 STEWARD Resources 2
4.2.1 Summary of STEWARD utilities 2
4.2.2 Summary of STEWARD traffic volume report 4
4.3 Maps 5
4.3.1 GIS maps 6
4.3.2 Interactive map using Google maps. 7
4.4 District Data and Reports 8
4.4.1 Facility level reports 8
4.4.2 Section level reports 13
4.4.3 Station level reports 14
4.4.4 TSS ETL Report 15
4.4.5 Report archive 16
4.5 Report details 18
4.5.1 All Data Fields in TSS Facility Level Report 18
4.5.2 Volume map and I/O balance in TSS Facility Level Report 19
4.5.3 Traffic counts in TSS Facility Level Report 19
4.5.4 Performance measure in TSS Section Level Report 19
4.5.5 Travel Time Reliability in TSS Section Level Report 20
4.5.6 All data fields in TSS Station Level Report 21
4.5.7 Traffic counts in TSS Station Level Report 21
4.5.8 Max Flow Rate in TSS Station Level Report 22
4.5.9 Effective Vehicle Length in theTSS Station Level Report 22
4.5.10 Contact Information 23
5 Examples 24
5.1 Example1 Retrieve all data fields report from TSS facility-level report 24
5.2 Example2 Retrieve Travel time reliability report from TSS section-level report 28

List of Figures

Figure 1: STEWARD web architecture 1
Figure 2: STEWARD overview web page 3
Figure 3: STEWARD resources web page 4
Figure 4: STEWARD maps page 5
Figure 5: STEWARD TSS detector map for District 2 6
Figure 6: STEWARD TSS interactive detector map for District 2 I-95 7
Figure 7: STEWARD Facility level report page 9
Figure 8: STEWARD Facility level report page (interactive) 10
Figure 9: STEWARD Traffic count report page (on-screen data window) 11
Figure 10: STEWARD Traffic count report page (CSV file archives) 11
Figure 11: STEWARD CSV file list report web page 12
Figure 12: STEWARD Section level report page 13
Figure 13: STEWARD station level report page 15
Figure 14: STEWARD TSS ETL report web page 16
Figure 15: STEWARD Report archive web page 17
Figure 16: STEWARD contact information web page 23
Figure 17: Change the District from the pull-down menu 24
Figure 18: Select the TSS facility-level report from the left pane 25
Figure 19: Set the input parameters for data retrieval 26
Figure 20: Output results to create the *All Data Fields* report 27
Figure 21: Newly-created report in the report archive 28
Figure 22: Change the district from the pull-down menu 29
Figure 23 Select the TSS section-level report from the left pane 30
Figure 24: Set the input parameters for data retrieval 31
Figure 25: Output results to create travel time reliability report 32

List of Tables

Table 1: TSS ETL Report Table 16
Table 2: Contents of the *All Data Fields* Report 18
Table 3: Contents of theVolume Map and I/O Balance Report 19
Table 4: Contents of Traffic Counts Report 19
Table 5: Contents of Performance Measure and LOS Report Table 20
Table 6: Contents of the Travel Time Reliability Report 20
Table 7: Contents of the *All Data Fields* Report 21
Table 8: Contents of Traffic Counts Report 21
Table 9: Contents of Maximum Flow Report 22
Table 10: Contents of Effective Vehicle Length Report 22

1 Introduction

While some use of the Statewide Traffic Engineering Warehouse for Regionally Archived Data (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 document describes the Internet based features of STEWARD from the perspective of a user who seeks to query the database and produce reports via the Internet.

2 Overview of the STEWARD Web Interface

The STEWARD web site has been developed for general users to access and retrieve the STEWARD 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.

At this time, the STEWARD web site can be accessed from the following address:
http://cdwserver.ce.ufl.edu/steward/index.html

3 STEWARD Web Architecture

The STEWARD web site consists of three main categories, including overview, resources and district data reports. The overall architecture is shown in Figure 1.

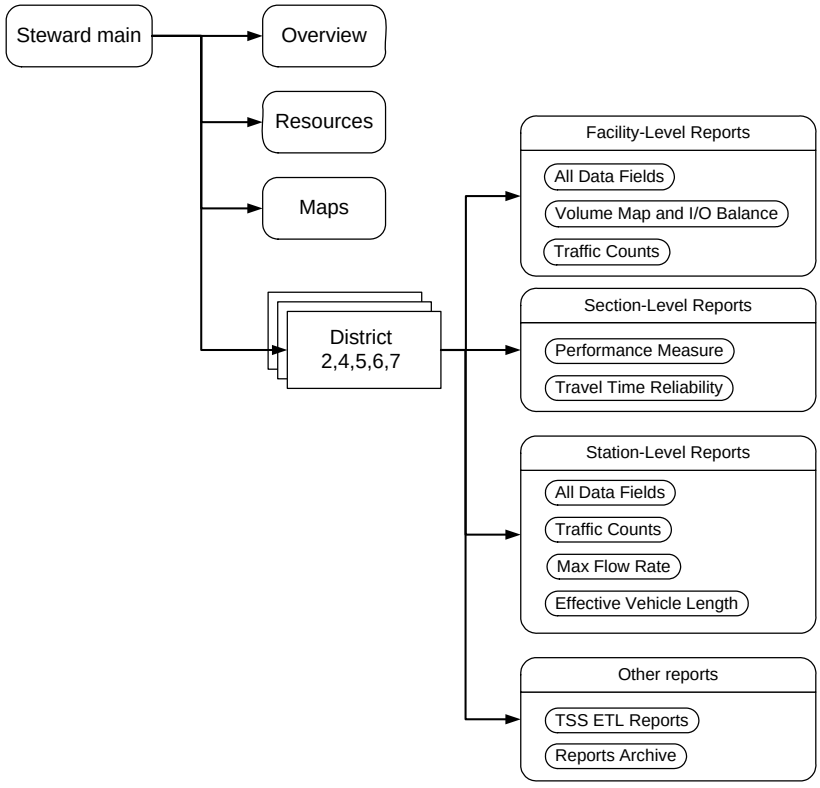


Figure 1: STEWARD web architecture

The Overview page provides a general description of the STEWARD project. The *Resource* page includes STEWARD project reports and utility programs. The *SunETLUtility* transforms

data files from the SunGuide data archive to the STEWARD database format. It executes data verification, data aggregation in time (5, 15 and 60 min intervals) and in space (lane-level and station-level data) and also generates reports. The *MPConverter* utility converts longitude/latitude data into state/county mile post or vice- versa. The results were validated with GPS data on I-95. It supports I-95 and I-295 data conversion in Florida only.

From the traffic data and reports section, the user can access traffic data (volume/speed/occupancy), travel time data as well as several reports and maps. STEWARD supports four types of reports:

- Facility level reports: *All Data Fields*, *Volume Map/IO Balance*, and *Traffic Counts*
- Section level reports: *Performance Measures* and *Travel Time Reliability*
- Station level reports: *All Data Fields*, *Traffic Counts*, *Maximum Flow Rates* and *Effective Vehicle Length*
- Other reports: *TSS ETL* reports, reports archives

4 STEWARD Web Pages

This section will review each STEWARD web page and describe its specific functions.

4.1 STEWARD Main Page –STEWARD Overview

The STEWARD main page includes two panes as shown in Figure 2. The main pane displays a brief description of the STEWARD project, objectives and tasks. The left pane is used to navigate to the STEWARD Overview, Resources and District Data/Reports sections.

4.2 STEWARD Resources

The STEWARD resource page includes the report section, the utilities section, and the traffic volume report section as shown in Figure 3. The report section includes the Phase II Final Report, progress reports and presentation material. In the utility section, seven utility programs, SunETLUtility, MPConverter, ITSCounts, SunVol, Hotter, SimTMC, and FTP Scripts are included. The traffic volume report section provides links to the traffic volume reports for all of the detectors in Districts 2, 4 and 6 for 2008.

4.2.1 Summary of STEWARD utilities

These programs have installation packages and user guides. All of the utilities are configured to be executable as stand-alone programs in the Microsoft Windows environment. The documentation for these programs is presented in other appendices to the STEWARD final report.

- The SunETLUtility performs the transformation process for STEWARD. The program reads the data files from SunGuide archive files, verifies and aggregates the data and generates reports. The output of the program is used as input to the STEWARD database.

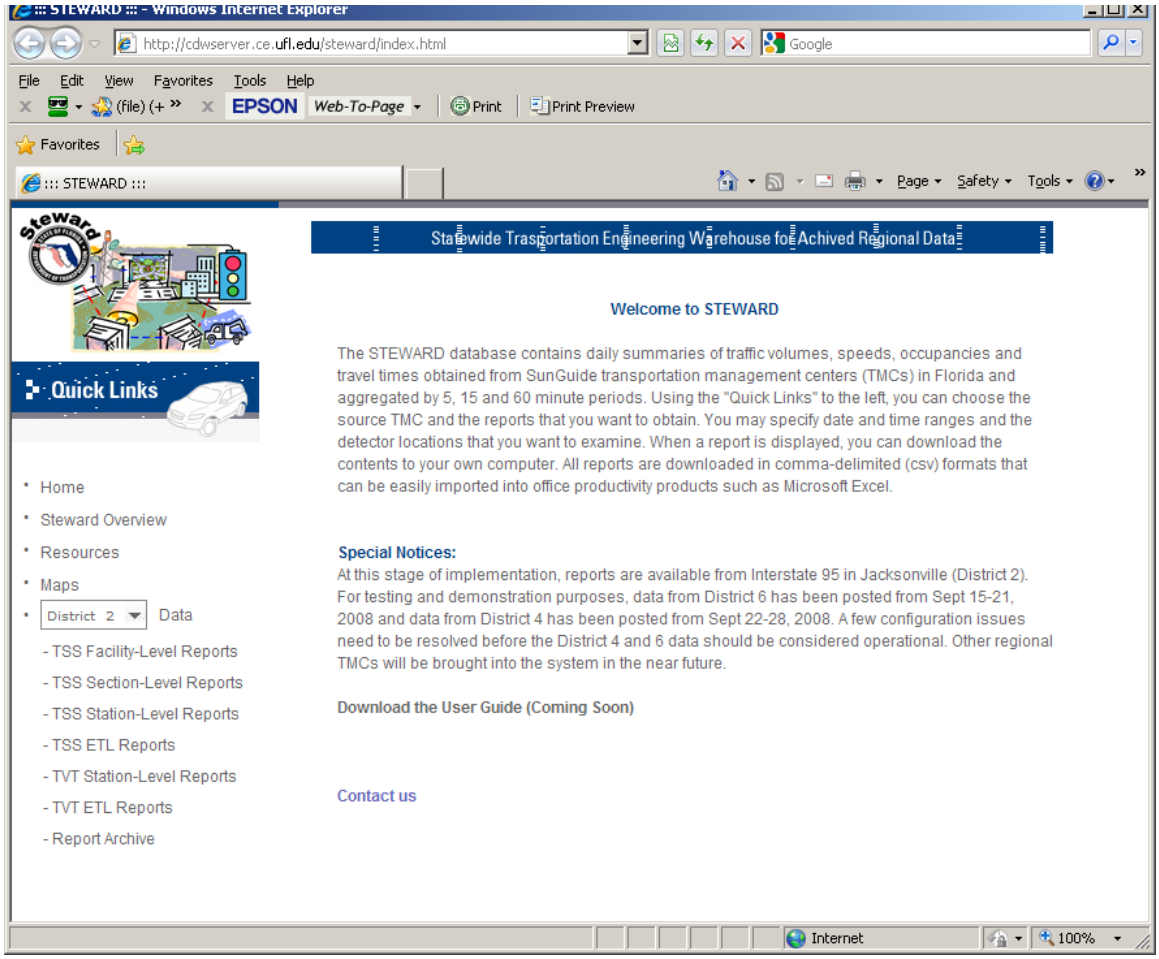


Figure 2: STEWARD overview web page

- The MPConverter converts longitude/latitude data into state/county milepost or vice-versa. This program was validated using GPS data on I-95 in Volusia County. The data conversion is limited to I-95 and I-295 in Florida only.

- The ITSCounts program converts the traffic counts in the SunGuide archives into the traffic count format used by the FDOT Statistics Office and District Planning Office. Traffic count reports from facility-level and station-level reports are used for the program input.

- The SunVol program examines the variability of data from day to day and identifies questionable days from the analysis of traffic counts.

- The HOTTER program analyzes the operation of managed lanes within SunGuide traffic management systems in Florida. It also can be used to analyze the operation of high occupancy vehicle (HOV) lanes and high occupancy toll (HOT) lanes.

- The SimTMC program was developed to simulate the SunGuide data archives using the CORSIM traffic network simulation model.

- A set of FTP Scripts was developed by District 4 and 6 to transfer SunGuide archive data into the STEWARD FTP server.

4.2.2 Summary of STEWARD traffic volume report

The STEWARD traffic counts from 2008 were used to create a set of traffic volume reports. All of the traffic counts were created using the SunVol utility and the results were plotted as graphics. Traffic volume reports were created for each facility in Districts 2, 4 and 6.

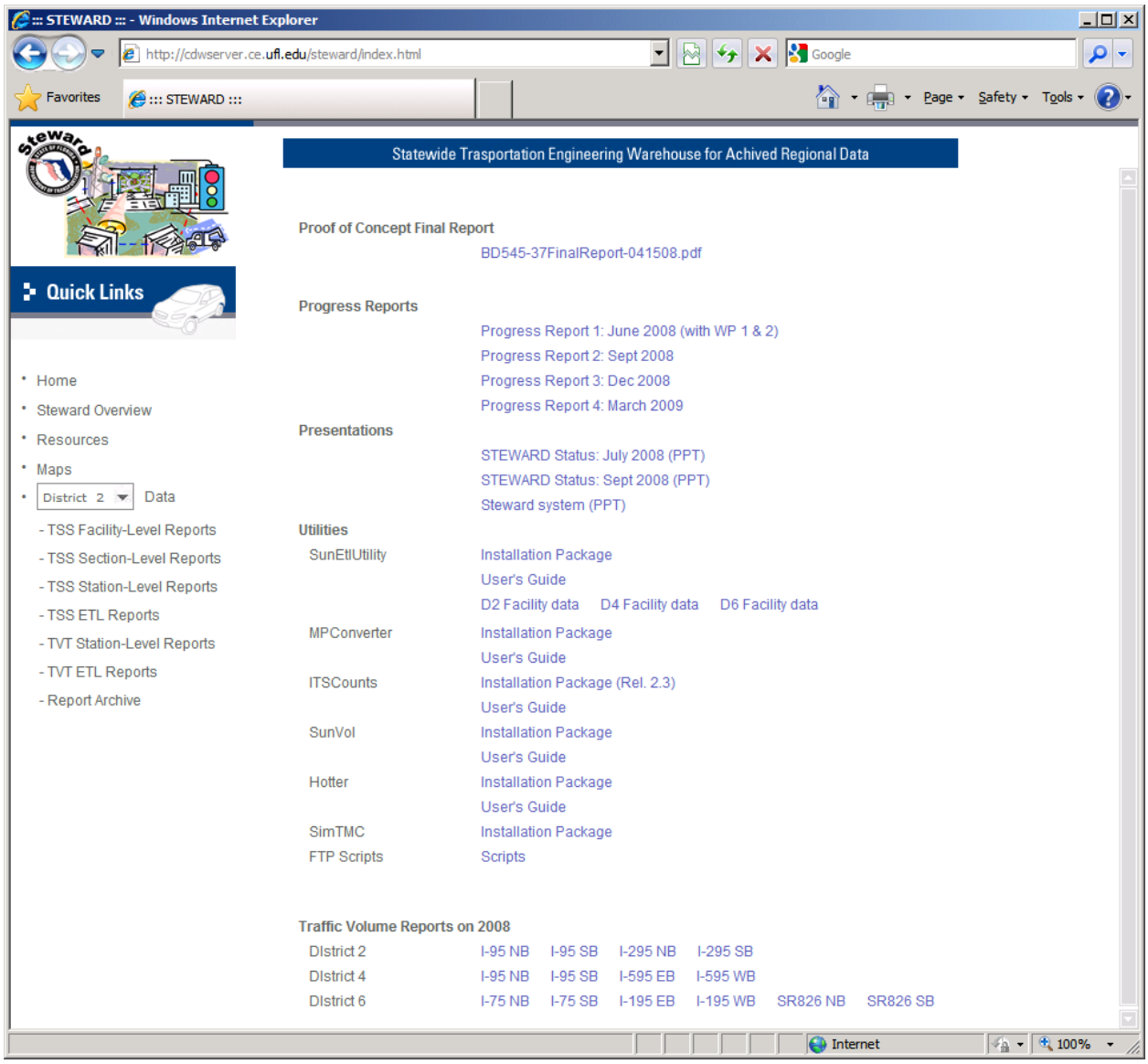


Figure 3: STEWARD resources web page

4.3 Maps

Detector maps for Districts 2, 4 and 6 were created and posted on a STEWARD web page. As Figure 4 shows, this page consists of two sections: statewide maps and district maps. District maps include GIS maps for each district and interactive Google maps.

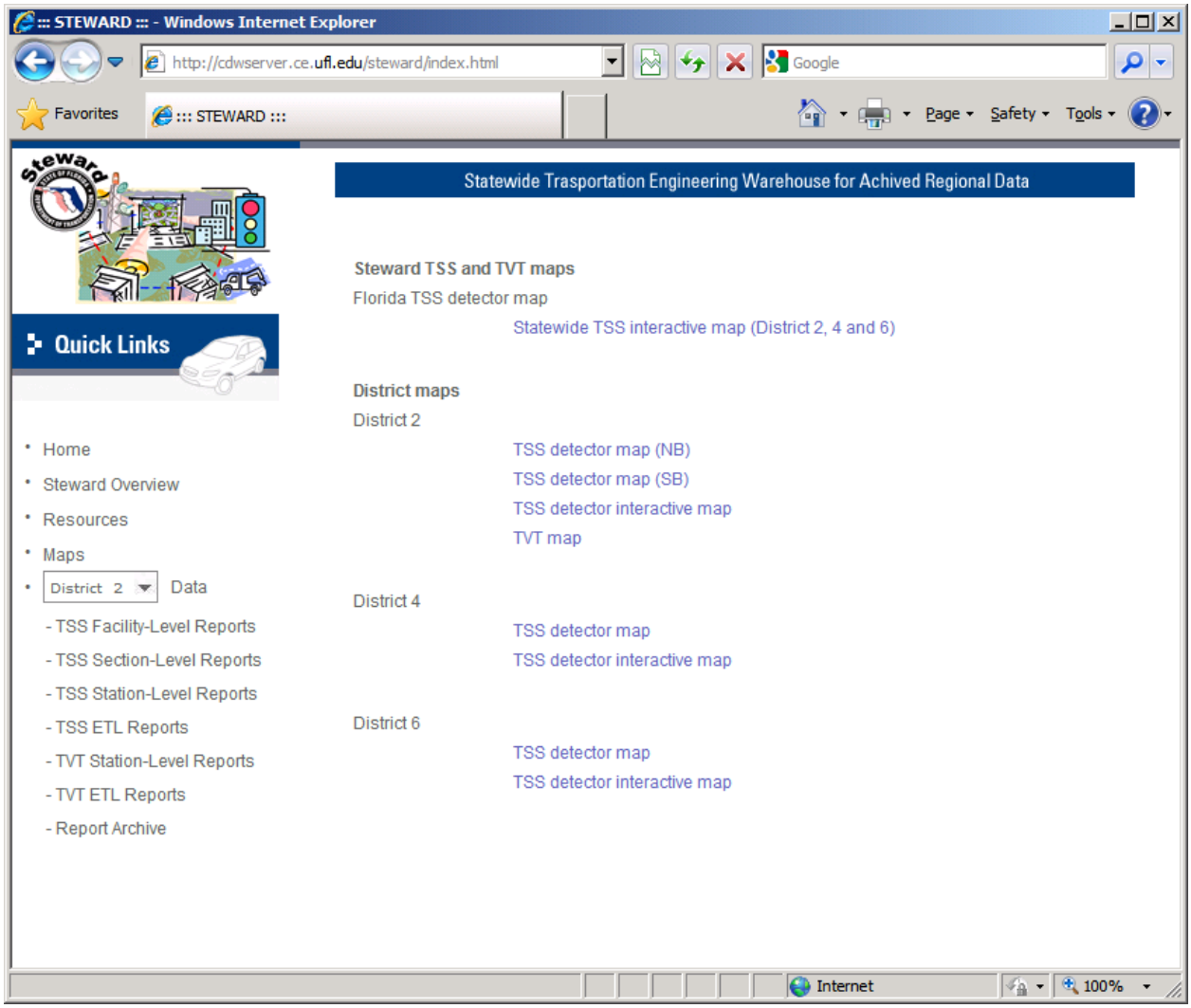


Figure 4: STEWARD maps page

4.3.1 GIS maps

STEWARD users can download TSS detector maps in GIF format. An examples of a GIS map is presented in Figure 5.

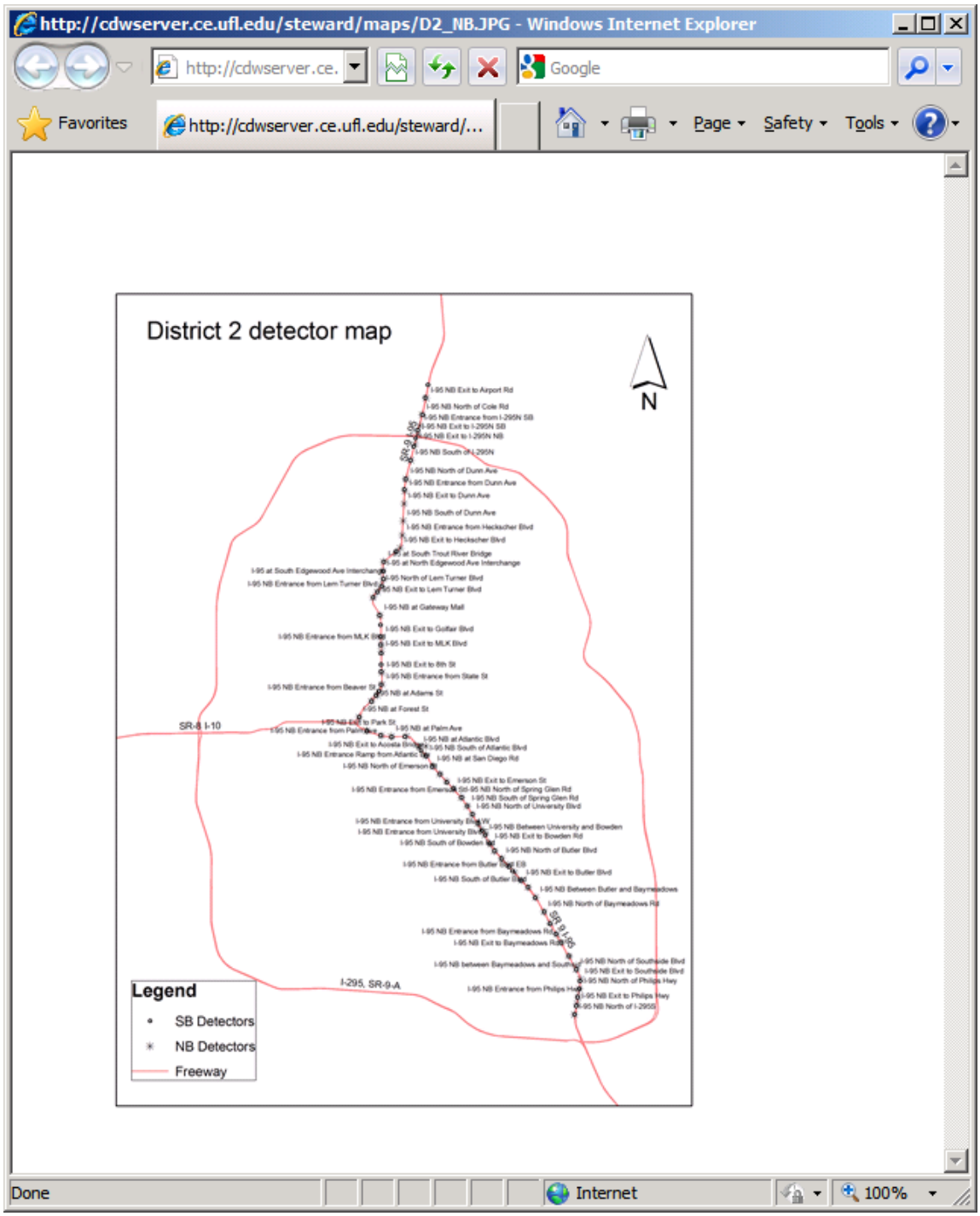


Figure 5: STEWARD TSS detector map for District 2

4.3.2 Interactive map using Google maps.

STEWARD users can retrieve detector information from the map interactively. An example of a Google map is presented in Figure 6. From this map, each station is marked as a red balloon and provides the station information using tool tips. For example, the station tool tip shows the station ID (200331), number of lanes (two lanes), latitude/longitude and the station description. These maps were developed using the Google map APIs and have the same interface as the Google maps. Therefore, the user can zoom in or out of a map using the scale on the left side or change the base map type using right side selection buttons.

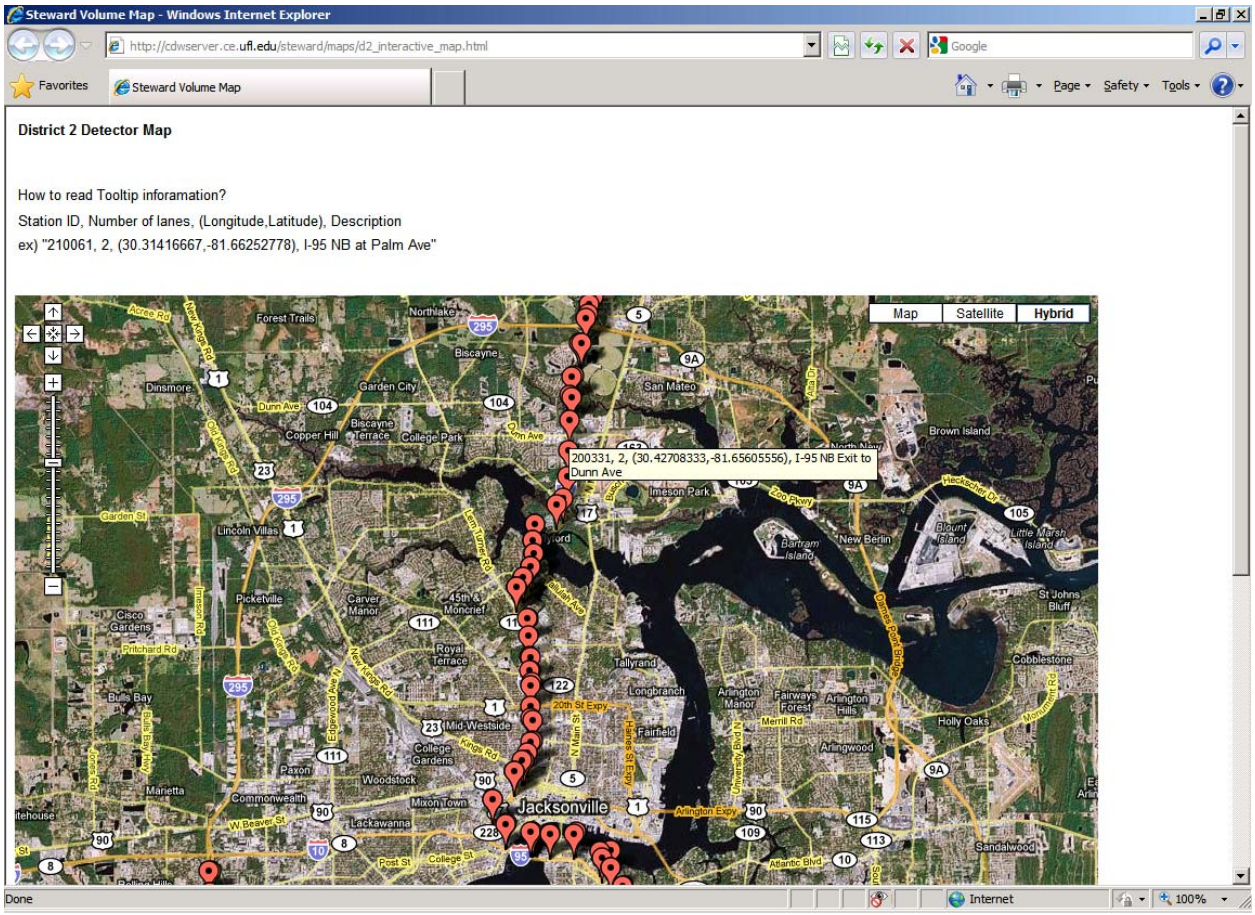


Figure 6: STEWARD TSS interactive detector map for District 2 (I-95)

4.4 District Data and Reports

When the District is selected in the left pane, the user can access one of four items (facility level reports, section level reports, station level reports, and other reports) for the selected district. At this point, District 2, 4 and 6 data are available.

The structures of each web page are almost identical. In this document, the first page, *Facility level reports* will be described in detail and the other pages will be treated in less detail when their content overlaps the *Facility level reports* pages.

4.4.1 Facility level reports

Facility level reports generate system-wide reports in the specific date. As shown in Figure 7, facility level reports generate three types of reports: *All Data Fields, Volume Map and I/O Balance,* and *Traffic Counts*. STEWARD users can select the date, time range, facility and direction. The reports can be retrieved from the aggregated data in 5min, 15min and 1hr resolution from the STEWARD database. The output reports have two formats: on-screen report or CSV file format.

The user can query the traffic data with selection criteria as follows:

- In the date range, a STEWARD user can select the specific date. The last month that the traffic data are available is displayed in the calendar. The date in calendar has a bold font if the data on that day are available. For example, STEWARD database archived traffic volume data from April 1st, 2009 to April 28th, 2009 and the calendar shows bold fonts for those dates. The date can be selected by clicking on that date on the calendar. Then the color of the selected date will be changed to red.
- The daily time range can be specified by the starting and ending hour.
- Four types of aggregation level are supported: 5min, 15min, 1hr and full-period. The full-period report generates the summary of the daily time range from 5min data for the selected day. The summary is the sum (in case of volume), average (occupancy), or weighted average (speed), depending on the item.
- Facility and direction can also be selected. In case of District 2, I-95 and NB is the default set.
- In the report section, the user can select the output report format as on-screen or CSV file format. The on-screen format displays the report on screen and the CSV file generates the report for the data download. The output reports are same for both formats.

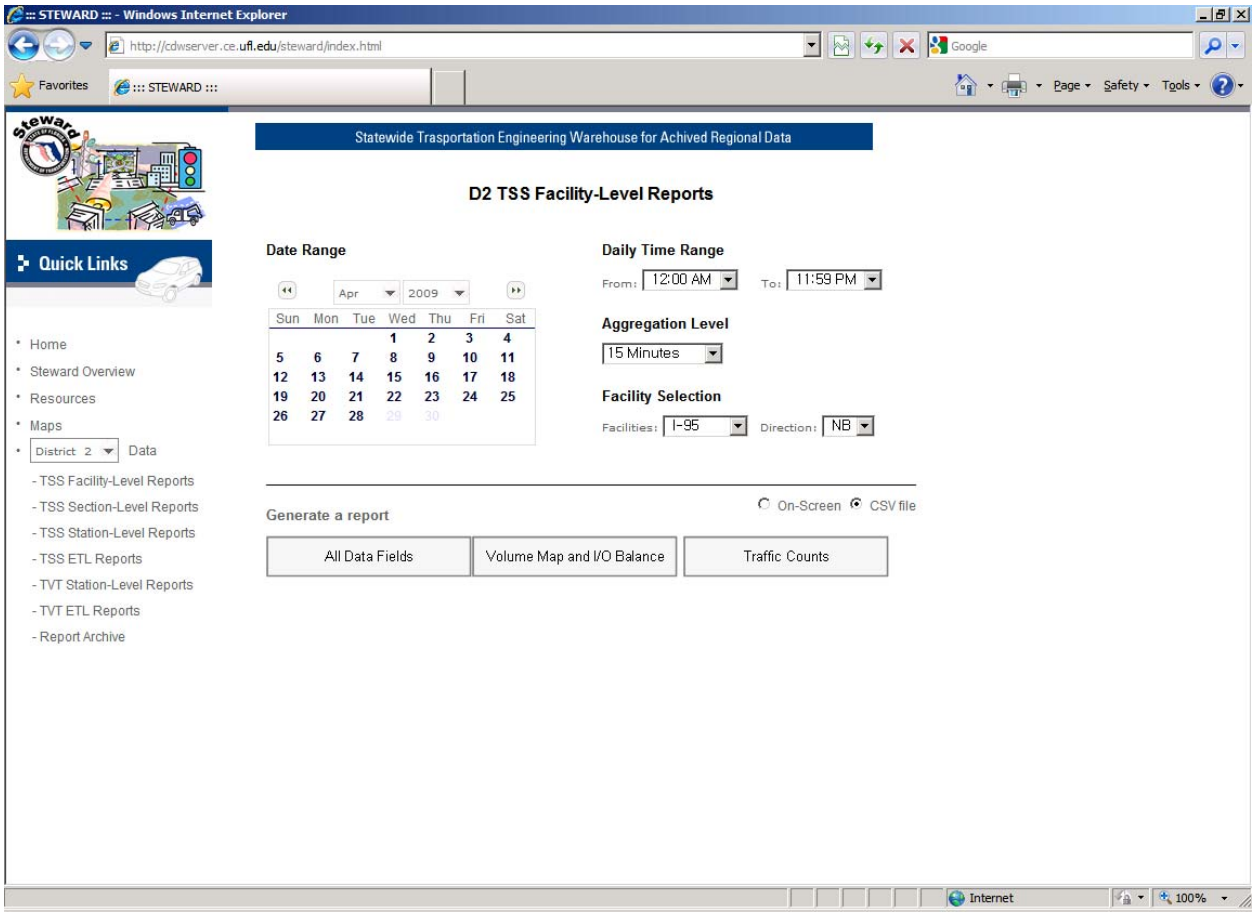


Figure 7: STEWARD Facility level report page

Figure 8 shows the resulting screen when the user selects the following criteria:

- Date: April 16th, 2008
- Time range: 8:00AM ~8:59AM
- Aggregation level: 1hour
- Facility: I-95
- Direction: SB
- Output format: On-Screen

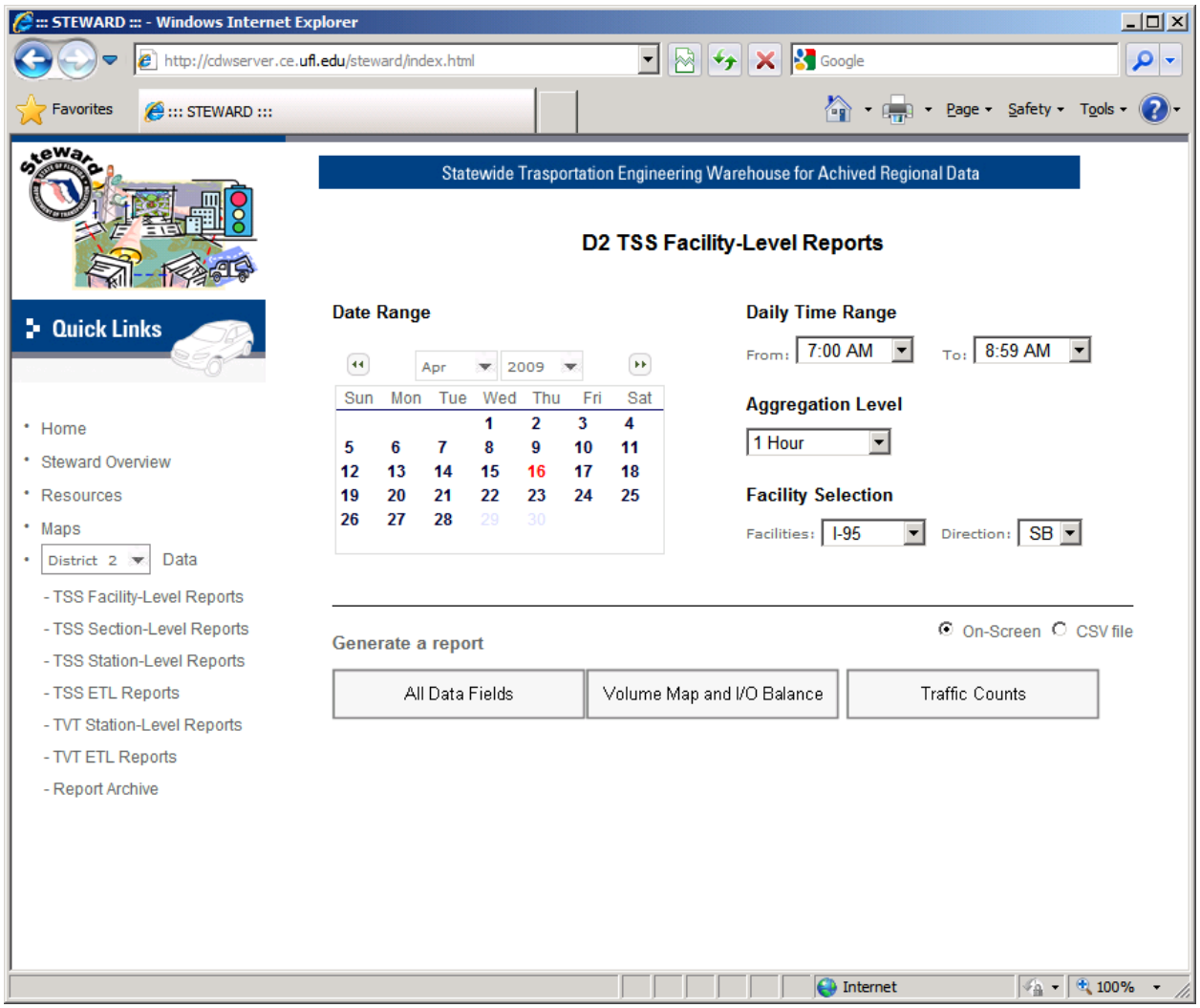
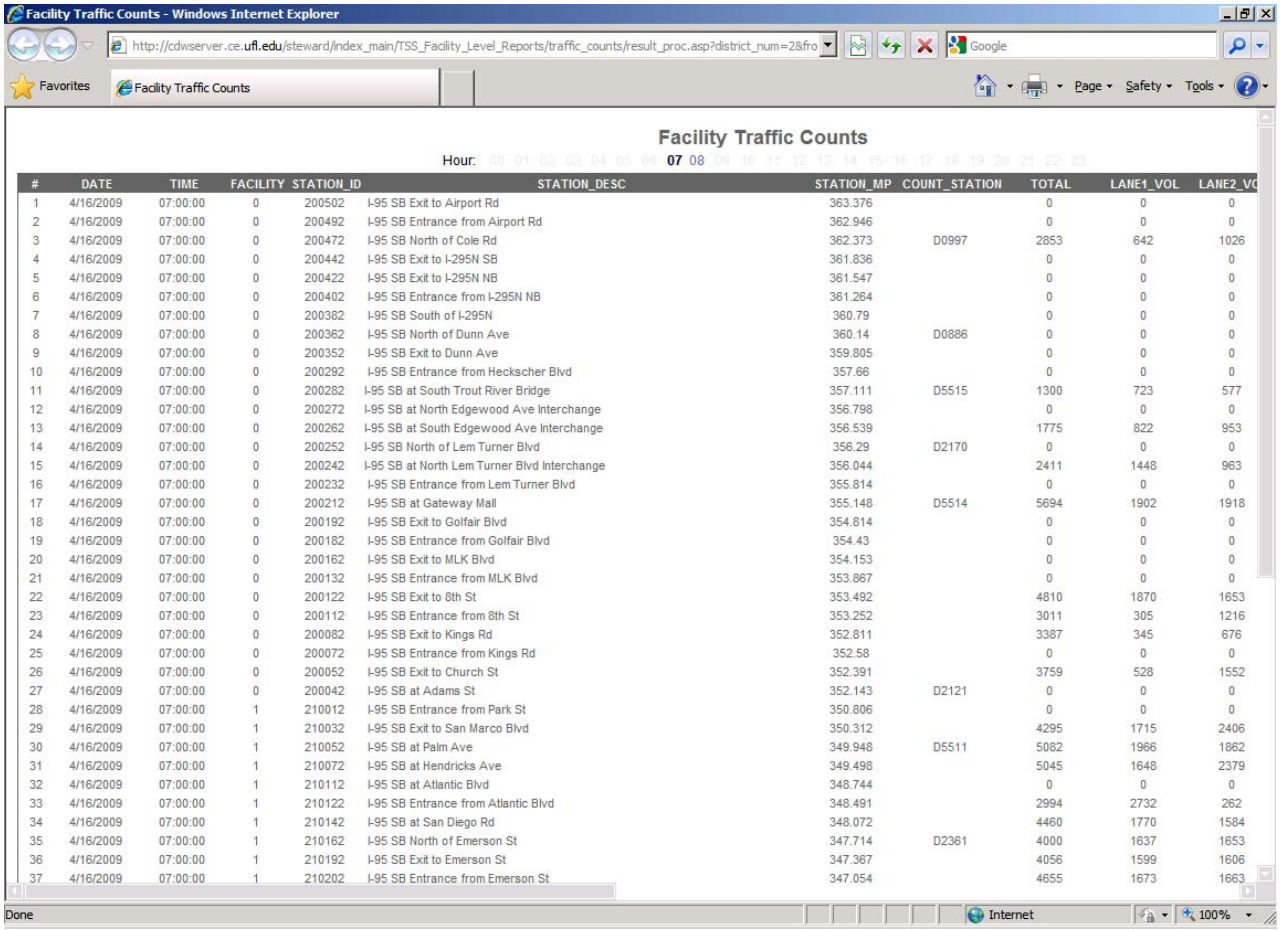


Figure 8: STEWARD Facility level report page (interactive)

When one of three buttons (*All Data Fields*, *Volume Map and I/O Balance*, or *Traffic Counts*) is clicked, a new popup window will appear as shown in Figure 9 with the report. In this case, the *Traffic Counts* report is illustrated. The user can check the results on screen. If the "CSV file" output option is selected and *Traffic Counts* was clicked, a newly generated file is archived at the top of the *Report Archives* list. The file name is generated using a combination of district name, facility name, direction, the selected data and the creation date.



#	DATE	TIME	FACILITY	STATION_ID	STATION_DESC	STATION_MP	COUNT_STATION	TOTAL	LANE1_VOL	LANE2_VO
1	4/16/2009	07:00:00	0	200502	I-95 SB Exit to Airport Rd	363.376		0	0	0
2	4/16/2009	07:00:00	0	200492	I-95 SB Entrance from Airport Rd	362.946		0	0	0
3	4/16/2009	07:00:00	0	200472	I-95 SB North of Cole Rd	362.373	D0997	2853	642	1026
4	4/16/2009	07:00:00	0	200442	I-95 SB Exit to I-295N SB	361.836		0	0	0
5	4/16/2009	07:00:00	0	200422	I-95 SB Exit to I-295N NB	361.547		0	0	0
6	4/16/2009	07:00:00	0	200402	I-95 SB Entrance from I-295N NB	361.264		0	0	0
7	4/16/2009	07:00:00	0	200382	I-95 SB South of I-295N	360.79		0	0	0
8	4/16/2009	07:00:00	0	200362	I-95 SB North of Dunn Ave	360.14	D0886	0	0	0
9	4/16/2009	07:00:00	0	200352	I-95 SB Exit to Dunn Ave	359.805		0	0	0
10	4/16/2009	07:00:00	0	200292	I-95 SB Entrance from Heckscher Blvd	357.66		0	0	0
11	4/16/2009	07:00:00	0	200282	I-95 SB at South Trout River Bridge	357.111	D5515	1300	723	577
12	4/16/2009	07:00:00	0	200272	I-95 SB at North Edgewood Ave Interchange	356.798		0	0	0
13	4/16/2009	07:00:00	0	200262	I-95 SB at South Edgewood Ave Interchange	356.539		1775	822	953
14	4/16/2009	07:00:00	0	200252	I-95 SB North of Lem Turner Blvd	356.29	D2170	0	0	0
15	4/16/2009	07:00:00	0	200242	I-95 SB at North Lem Turner Blvd Interchange	356.044		2411	1448	963
16	4/16/2009	07:00:00	0	200232	I-95 SB Entrance from Lem Turner Blvd	355.814		0	0	0
17	4/16/2009	07:00:00	0	200212	I-95 SB at Gateway Mall	355.148	D5514	5694	1902	1918
18	4/16/2009	07:00:00	0	200192	I-95 SB Exit to Golfair Blvd	354.814		0	0	0
19	4/16/2009	07:00:00	0	200182	I-95 SB Entrance from Golfair Blvd	354.43		0	0	0
20	4/16/2009	07:00:00	0	200162	I-95 SB Exit to MLK Blvd	354.153		0	0	0
21	4/16/2009	07:00:00	0	200132	I-95 SB Entrance from MLK Blvd	353.867		0	0	0
22	4/16/2009	07:00:00	0	200122	I-95 SB Exit to 8th St	353.492		4810	1870	1653
23	4/16/2009	07:00:00	0	200112	I-95 SB Entrance from 8th St	353.252		3011	305	1216
24	4/16/2009	07:00:00	0	200082	I-95 SB Exit to Kings Rd	352.811		3387	345	676
25	4/16/2009	07:00:00	0	200072	I-95 SB Entrance from Kings Rd	352.58		0	0	0
26	4/16/2009	07:00:00	0	200052	I-95 SB Exit to Church St	352.391		3759	528	1552
27	4/16/2009	07:00:00	0	200042	I-95 SB at Adams St	352.143	D2121	0	0	0
28	4/16/2009	07:00:00	1	210012	I-95 SB Entrance from Park St	350.806		0	0	0
29	4/16/2009	07:00:00	1	210032	I-95 SB Exit to San Marco Blvd	350.312		4295	1715	2406
30	4/16/2009	07:00:00	1	210052	I-95 SB at Palm Ave	349.948	D5511	5082	1966	1862
31	4/16/2009	07:00:00	1	210072	I-95 SB at Hendricks Ave	349.498		5045	1648	2379
32	4/16/2009	07:00:00	1	210112	I-95 SB at Atlantic Blvd	348.744		0	0	0
33	4/16/2009	07:00:00	1	210122	I-95 SB Entrance from Atlantic Blvd	348.491		2994	2732	262
34	4/16/2009	07:00:00	1	210142	I-95 SB at San Diego Rd	348.072		4460	1770	1584
35	4/16/2009	07:00:00	1	210162	I-95 SB North of Emerson St	347.714	D2361	4000	1637	1653
36	4/16/2009	07:00:00	1	210192	I-95 SB Exit to Emerson St	347.367		4056	1599	1606
37	4/16/2009	07:00:00	1	210202	I-95 SB Entrance from Emerson St	347.054		4655	1673	1663

Figure 9: STEWARD Traffic count report page (on-screen data window)

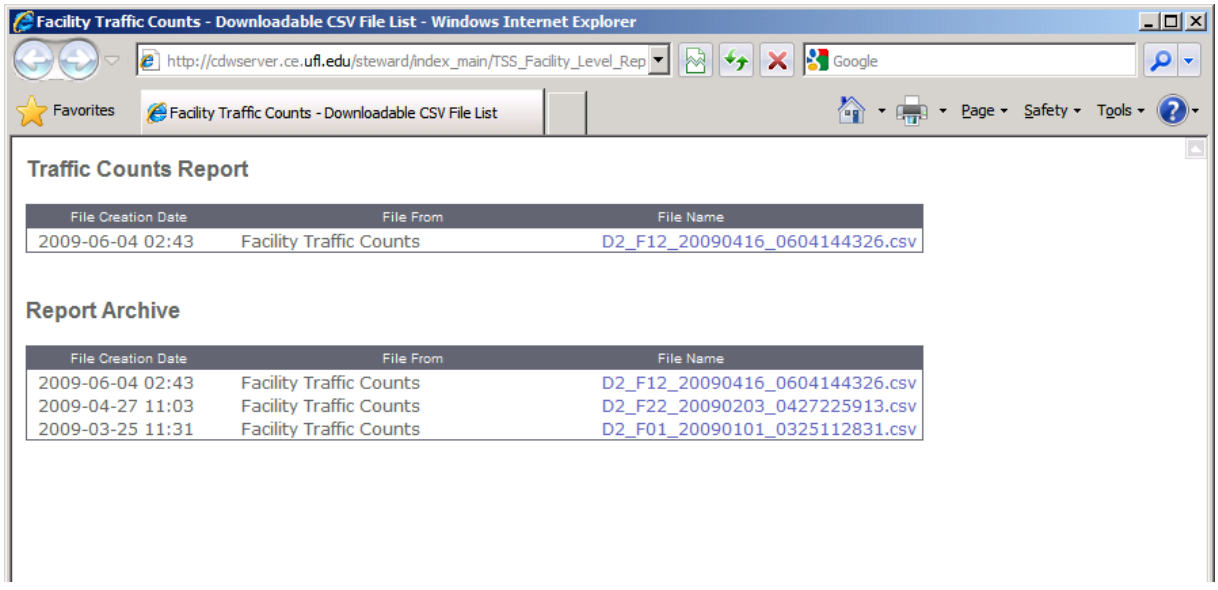


Figure 10: STEWARD Traffic count report page (CSV file archives)

When a new CSV file is generated, it is also registered to the report archive page. Therefore, the user can access this CSV file later without a repeated data entry. For example, when a new CSV file (D2_F12_20090416_0604144326.csv in Figure 10) is generated, the same CSV file is archived in CSV Report page as shown in Figure 11.

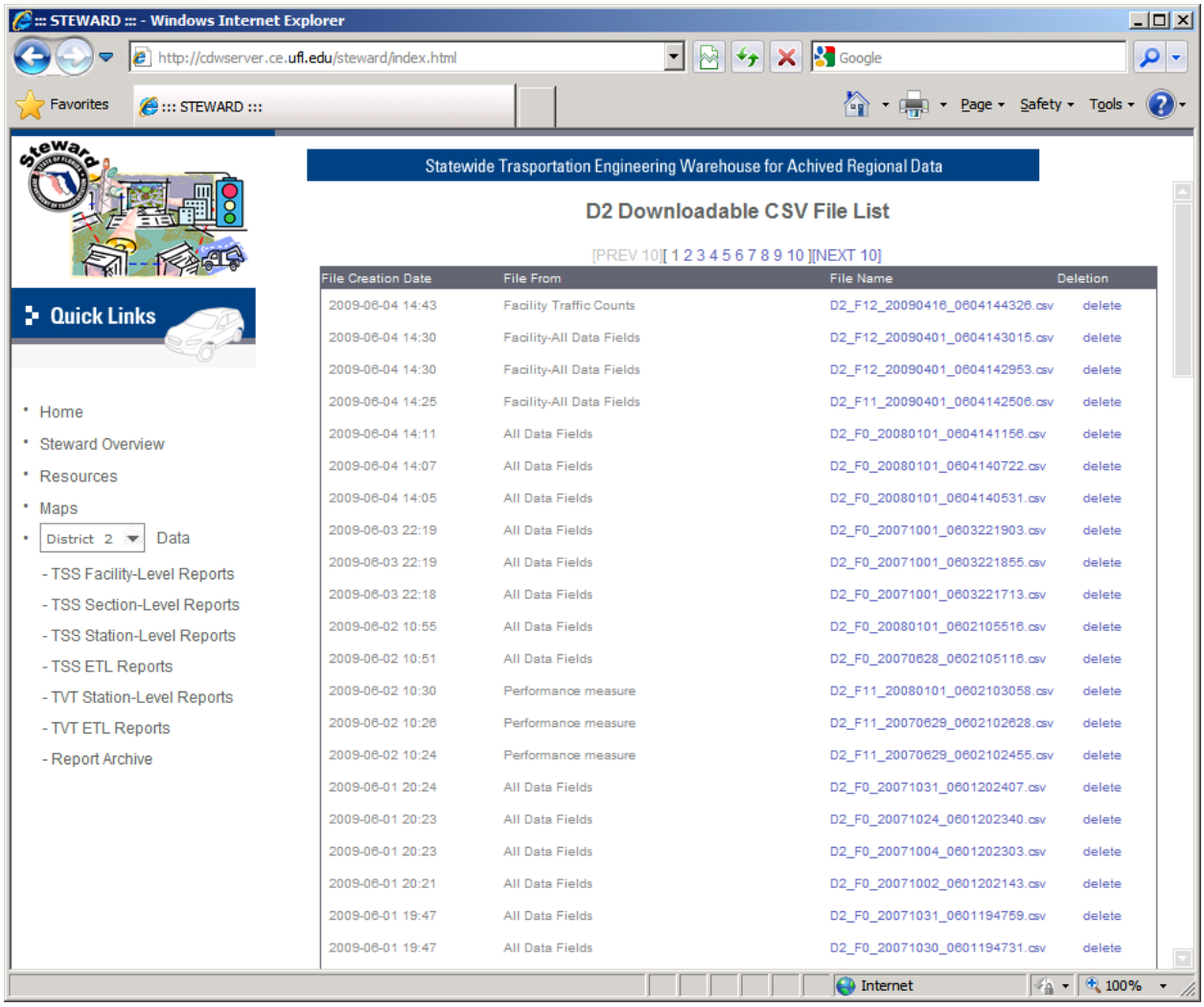


Figure 11: STEWARD CSV file list report web page

4.4.2 Section level reports

Section level reports generate reports from a specified section of the facility. As shown in Figure 12, there are two types of reports: *Performance Measures* and *Travel Time Reliability*. STEWARD users can select the start date, end date, time range, aggregation level, facility, direction, and start and end stations. The reports can be retrieved from the aggregated data in the STEWARD database. The output reports have two formats: on-screen report or CSV file format.

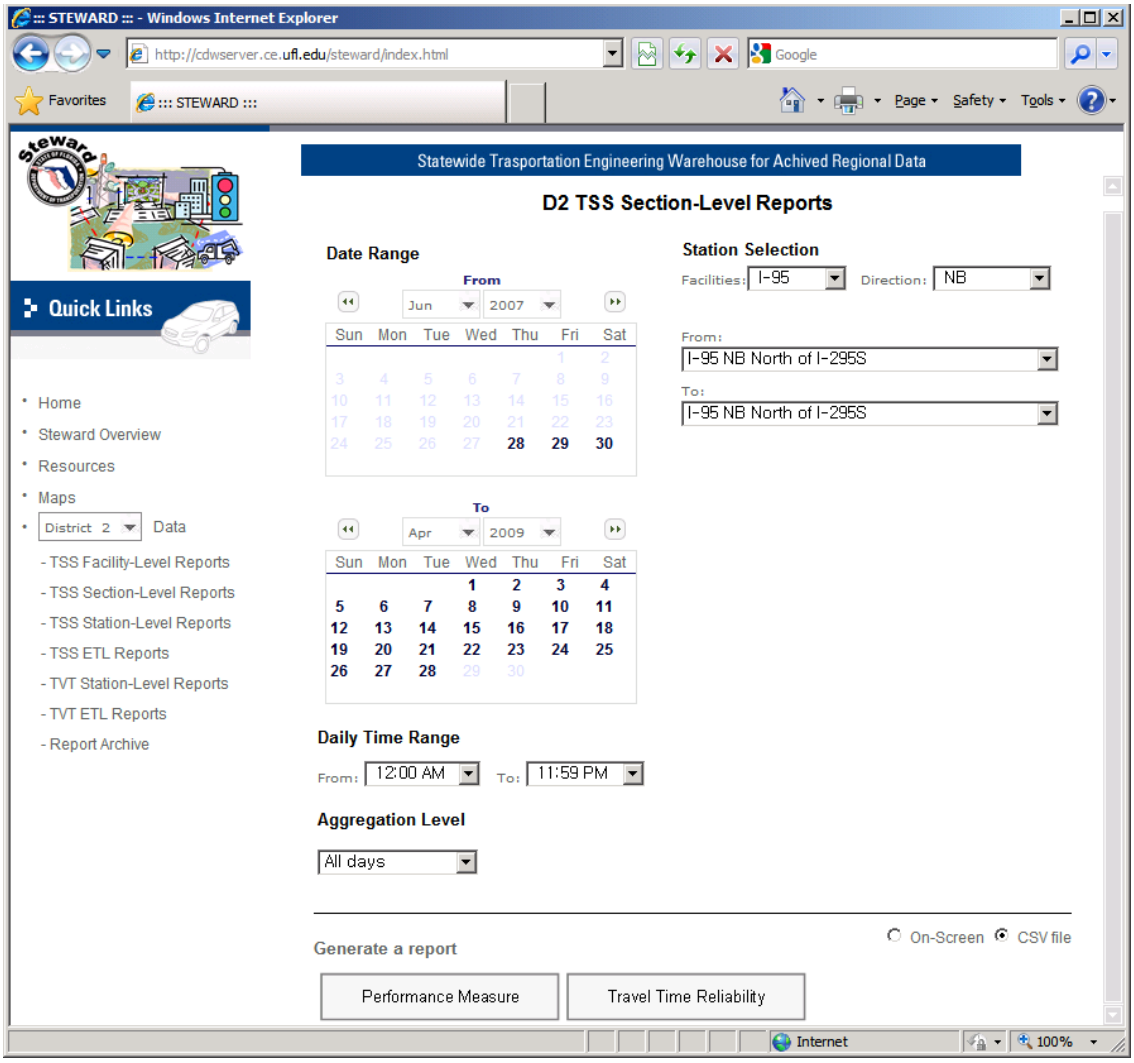


Figure 12: STEWARD Section level report page

The user can query the traffic data with the specific conditions as follows:

- In the date range, users can select the start and end date. The first and last month that the traffic data are available are displayed in the *from* and *to* calendars. The date in calendars have a bold font if the data on that day are available. For example, STEWARD database archived traffic volume data are available from June 28th, 2009 to April 28th, 2009 and Figure 12he calendar shows bold fonts for those dates. The start and end date can be selected by

clicking on those dates on the calendar. Then the color of the selected date will be changed to red.
- The daily time range can be specified by the starting and ending hour.
- In the aggregation level, the user can select the day of the week (weekday, weekend, all days, Mondays, Tuesdays, Wednesdays, Thursdays, Fridays, Saturdays and Sundays).
- Facility and direction can be selected first and then the start station and end station can be selected. In the case of District 2, I-95, NB, and I-95 NB North of I-295S are the default set.
- In the report section, the user can select the output report format as on-screen or CSV file format. The on-screen format displays the report on screen and the CSV file generates the reports for the data download. The output reports are same for both formats.

4.4.3 Station level reports

Station level reports may be generated from the selected stations on the facility. As shown in Figure 13, there are four types of reports: *All data fields*, *Traffic Counts*, *Maximum flow rate*, and *Effective vehicle length*. Users can select the start date, end date, time range, aggregation level, facility, direction, and stations. The reports can be retrieved from the aggregated data that are in 5min, 15min and 1hr resolution in the STEWARD database. The output reports have two formats: On-screen report or CSV file format.

The user can query the traffic data with specific conditions as follows:
- In the date range, the user can select the start and end date. The first and last month that the traffic data are available are displayed in the *from* and *to* calendar. The date in calendar has a bold font if the data on that day are available. For example, STEWARD database archived traffic volume data from June 28th, 2009 to April 28th, 2009 and the calendar shows bold fonts for those dates. The start and end date can be selected by clicking on those dates on the calendar. Then the color of the selected date will be changed to red.
- The daily time range can be specified by the starting and ending hour.
- In the aggregation level, the user can select days of the week (weekday, weekend, all days, Mondays, Tuesdays, Wednesdays, Thursdays, Fridays, Saturdays and Sundays) and aggregation level (5min, 15min, 1hr and the full period).
- Facility and Direction can be selected first and then multiple stations can be selected. In case of District 2, I-95 and NB are the default set.
- In the report section, the user can select the output report format as on-screen or CSV file format. The on-screen format display the report on screen and the CSV file generates the reports for the data download. The output reports are same for both of these formats.

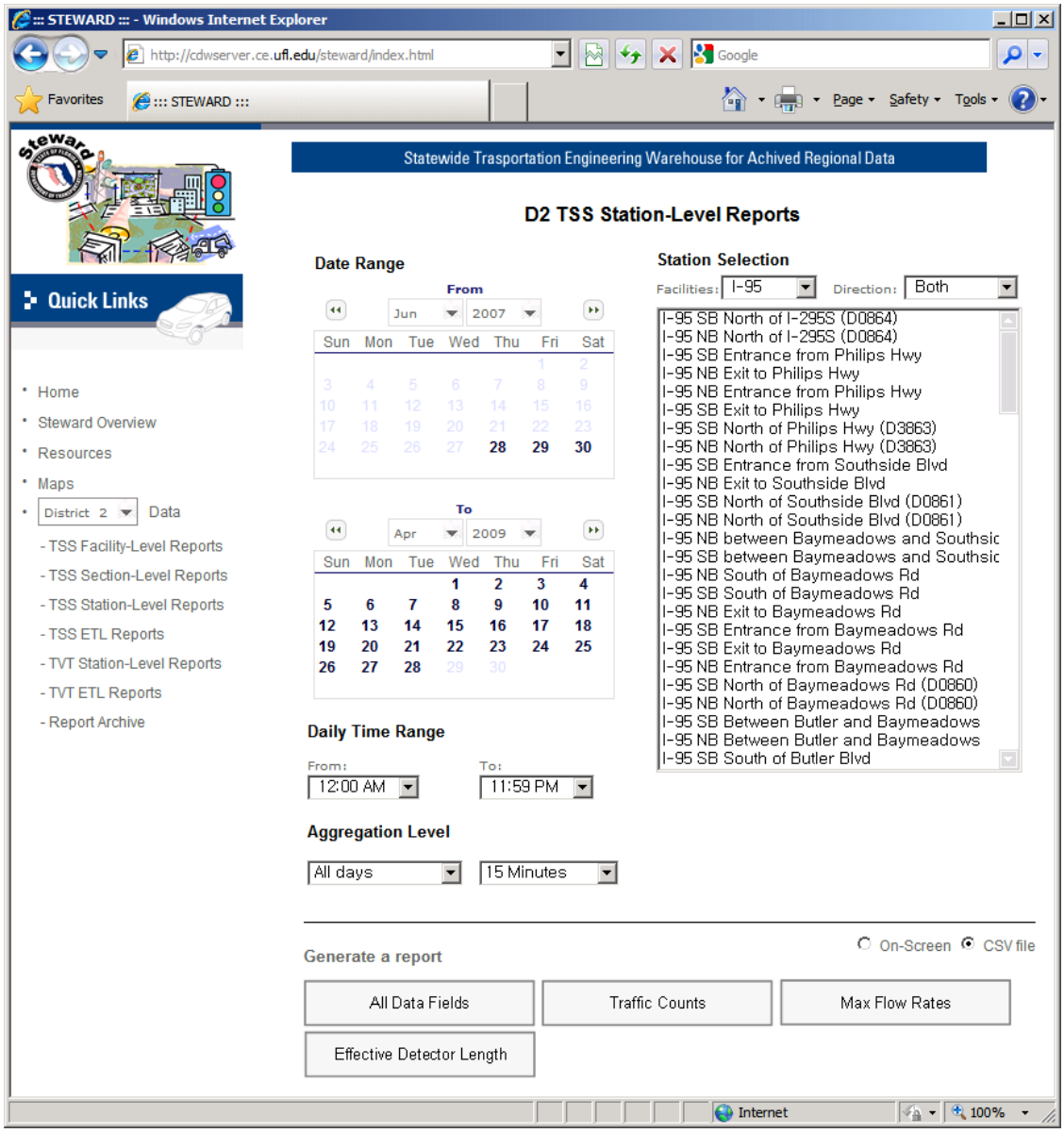


Figure 13: STEWARD station level report page

4.4.4 TSS ETL Report

As shown in Figure 14, users can access the TSS ETL reports. These reports are generated while the ETL program processes the input data file. One log file is generated for every input data file and one summary row is appended into the conversion log file. The conversion log file has the name of TSSConversionHistory-Dx.csv, where x stands for district number. The log file has the name of Dx_TSS-mmddyyyy-DailyReport.Log, where x stands for district number and mmddyyyy stands for the date data are collected. For example, D2_TSS-10042007-DailyReport.Log is the log file for District 2 TSS data collected on Oct. 4th, 2007. In these log files, the status of each station and diagnostic results are summarized.

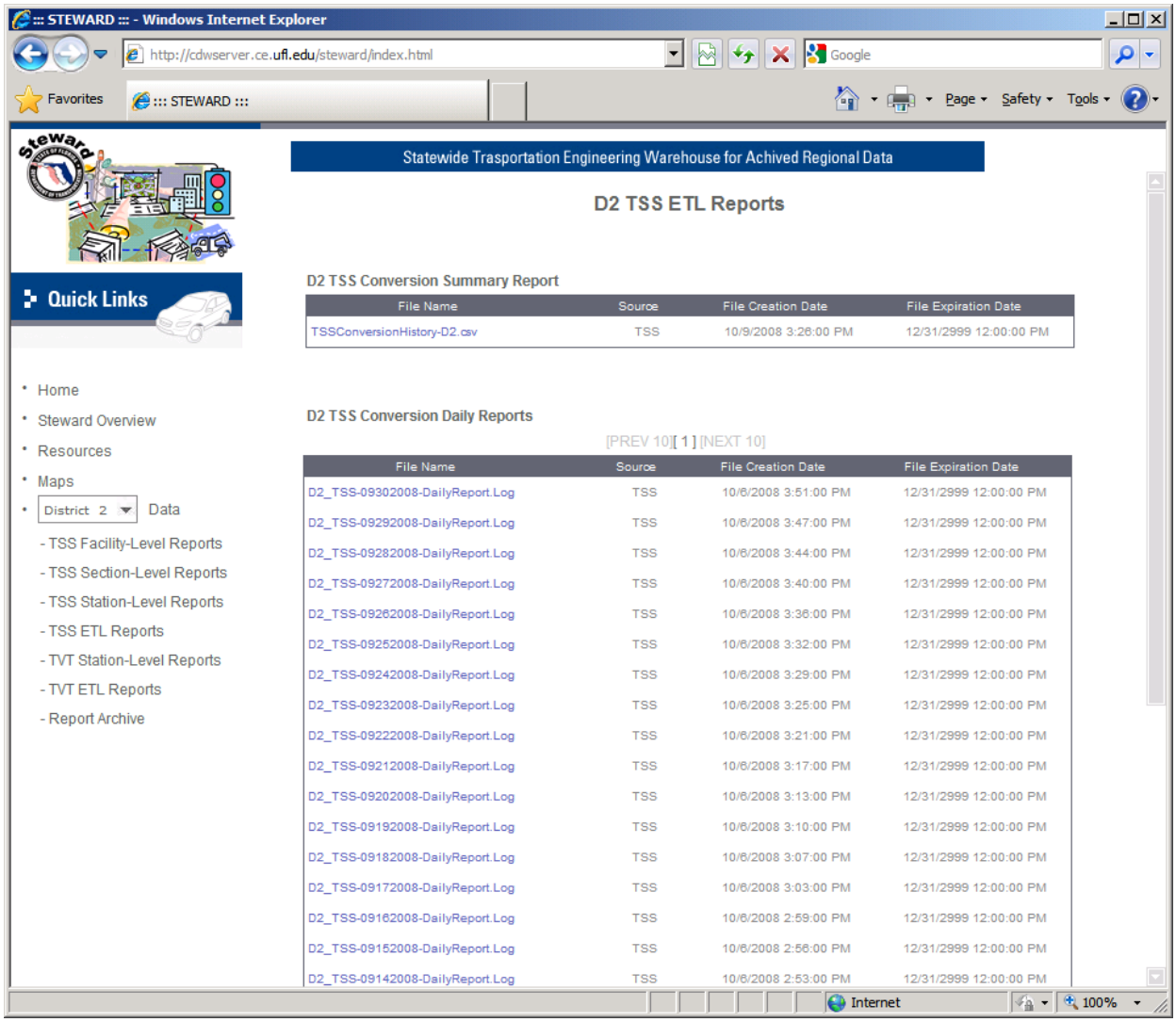


Figure 14: STEWARD TSS ETL report web page

Table 1 describes each column of the TSS ETL Report table.

Table 1: TSS ETL Report Table

Column	Data Type
File name	Log file name
Source	Data source
File creation date	Date the log file was generated
File expiration date	Date the log file could be removed

4.4.5 Report archive

As described in the *TSS Facility level report section*, the CSV files for each report are generated to the STEWARD server. As shown in Figure 15, these files are accessible from the CSV file list page. From this page, users can delete the CSV files from the STEWARD server or download the CSV files later after queries were made.

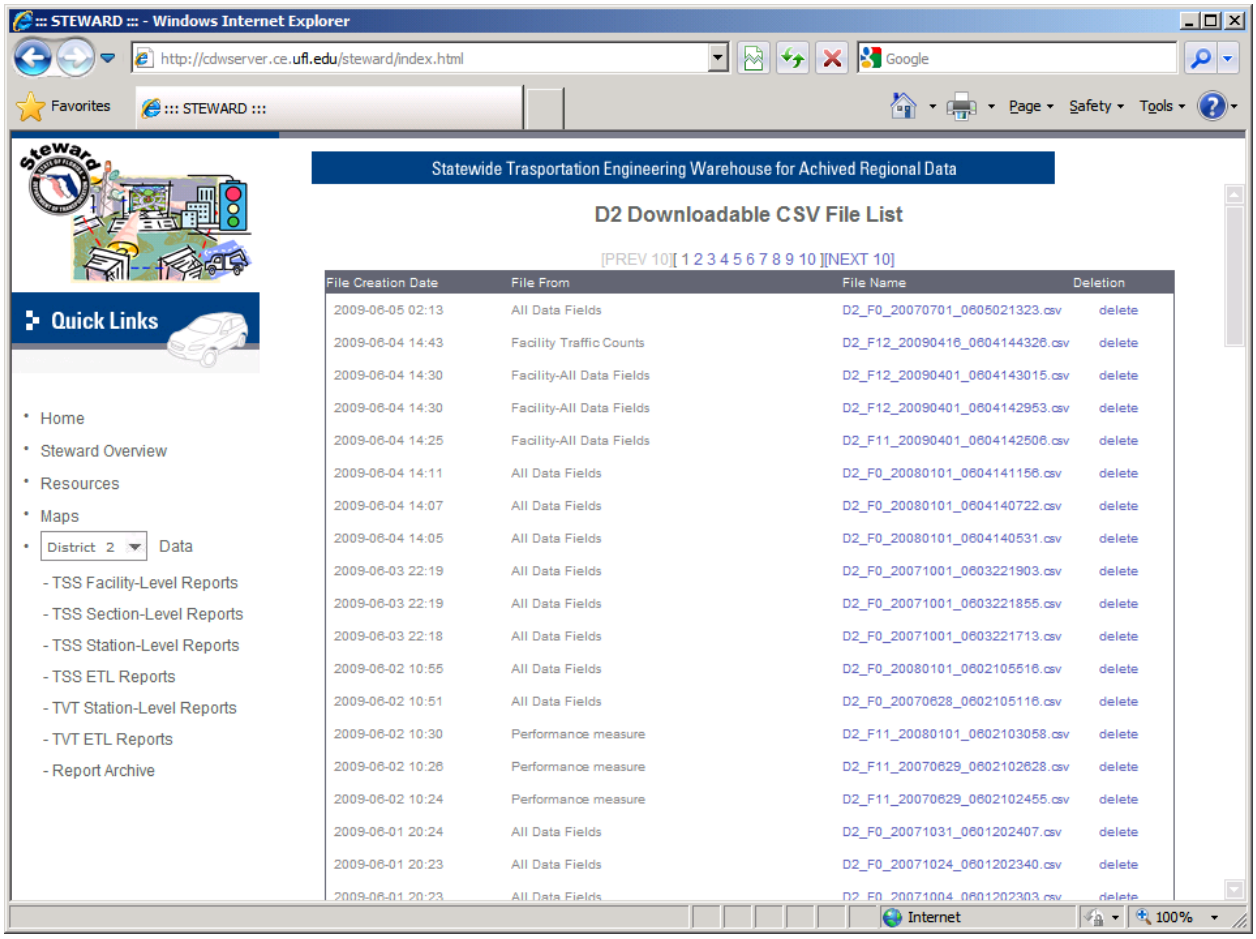


Figure 15: STEWARD report archive web page

4.5 Report details

In this section, the contents of each output report are summarized. The detail description and its contents will be described in the STEWARD main reports.

4.5.1 All Data Fields in TSS Facility Level Report

Table 2 describes each column of the *All Data Fields* report.

Table 2: Contents of the *All Data Fields* Report

Column	Data Type
DATE	Date of collection (MM/DD/YYYY)
TIME	Time Stamp (HH24:MM:SS)
Staiton_ID	Station ID
Station Description	Station Description
Station_MP	Station Milepost
Lane1 Vol	Lane1 Volume
Lane1 Occ	Lane1 Occupancy
Lane1 Spd	Lane1 Speed
Lane2 Vol	Lane2 Volume
Lane2 Occ	Lane2 Occupancy
Lane2 Spd	Lane2 Speed
Lane3 Vol	Lane3 Volume
Lane3 Occ	Lane3 Occupancy
Lane3 Spd	Lane3 Speed
Lane4 Vol	Lane4 Volume
Lane4 Occ	Lane4 Occupancy
Lane4 Spd	Lane4 Speed
Lane5 Vol	Lane5 Volume
Lane5 Occ	Lane5 Occupancy
Lane5 Spd	Lane5 Speed
Lane6 Vol	Lane6 Volume
Lane6 Occ	Lane6 Occupancy
Lane6 Spd	Lane6 Speed
Onramp1 Vol	Onramp1 Volume
Onramp1 Occ	Onramp1 Occupancy
Onramp1 Spd	Onramp1 Speed
Onramp2 Vol	Onramp2 Volume
Onramp2 Occ	Onramp2 Occupancy
Onramp2 Spd	Onramp2 Speed
Onramp3 Vol	Onramp3 Volume
Onramp3 Occ	Onramp3 Occupancy
Onramp3 Spd	Onramp3 Speed
Offramp1 Vol	Offramp1 Volume
Offramp1 Occ	Offramp1 Occupancy
Offramp1 Spd	Offramp1 Speed
Offramp2 Vol	Offramp2 Volume
Offramp2 Occ	Offramp2 Occupancy
Offramp2 Spd	Offramp2 Speed
Offramp3 Vol	Offramp3 Volume
Offramp3 Occ	Offramp3 Occupancy
Offramp3 Spd	Offramp3 Speed

4.5.2 Volume map and I/O balance in TSS Facility Level Report

Table 3 describes each column of the *Volume Map and I/O Balance* report.

Table 3: Contents of theVolume Map and I/O Balance Report

Column	Data Type
DATE	Date of collection (MM/DD/YYYY)
TIME	Time Stamp (HH24:MM:SS)
Staiton_ID	Station ID
Station_MP	Station Milepost
Upnode_ID	Upstream node station ID
Entry_Volume	Entrance ramp volume from the current node
FWY_Volume	Freeway volume from the current node
Exit_Volume	Exit ramp volume from the current node
Link_input	Sum of upnode freeway volume and upnode entrance ramp volume
Link_output	Sum of freeway volume and exit ramp volume at current node
Difference	Link_input – Link_output
Pcnt_Diff	100 * 2*(Link_input – Link_output) /(Link_input + Link_output)

4.5.3 Traffic counts in TSS Facility Level Report

Table 4 describes each column of the *Traffic Counts* report.

Table 4: Contents of Traffic Counts Report

Column	Data Type
DATE	Date of collection (MM/DD/YYYY)
TIME	Time Stamp (HH24:MM:SS)
Staiton_ID	Station ID
Station Description	Station Description
Station_MP	Station Milepost
Count Station	Count Station ID
Total	Total thru volume (Sum of lane1~lane6 volume)
Lane1 Vol	Lane1 Volume
Lane2 Vol	Lane2 Volume
Lane3 Vol	Lane3 Volume
Lane4 Vol	Lane4 Volume
Lane5 Vol	Lane5 Volume
Lane6 Vol	Lane6 Volume
FWY_QA	Freeway QA
ON_RAMP1	On Ramp1 Volume
ON_RAMP2	On Ramp2 Volume
ON_RAMP3	On Ramp3 Volume
ON_RAMP_QA	On Ramp QA
OFF_RAMP1	Off Ramp1 Volume
OFF_RAMP1	Off Ramp2 Volume
OFF_RAMP1	Off Ramp3 Volume
OFF_RAMP_QA	On Ramp QA
COUNTY	County number

4.5.4 Performance measure in TSS Section Level Report

Table 5 describes each column of the *Performance Measures* report.

Table 5: Contents of Performance Measure and LOS Report Table

Column	Data Type
Segment	Segment description
MP	Milepost
Length	Distance between upnode and current node
Average volume	Average volume per segment
Lanes	Number of lanes
Vol per Lane	Average of Link_In* and Link_out volume* *Link_In volume = Sum of upnode freeway volume and upnode entrance ramp volume *Link_out volume = Sum of freeway volume and exit ramp volume at current node
Veh-Miles	Link_volume * Length
Veh-Hours	Link_volume * Length/(freeway speed at current node)
Speed	Volume-weighted average speed
Delay (Veh-Hr)	When freeway speed is below (speed limit/1.5), delay = sum of ((link volume) * length *(1/freeway_speed – 1.5/(freeway_speed))
Kinetic Energy	Link_volume * freeway_speed
Percent Observations	(Number of observation)/(number of observations expected)
Density	(Link_volume/Lanes) / freeway_speed
V/C Ratio	100 * Density /Lane_Capacity In freeway, 2200 vplph is used as a capacity.
LOS	F when V/C > 100 E when Density > 35 D when Density > 26 C when Density > 18 B when Density > 11 A otherwise

4.5.5 Travel Time Reliability in TSS Section Level Report

Table 6 describes each column of the *Travel Time Reliability* report.

Table 6: Contents of the Travel Time Reliability Report

Column	Data Type
Segment	Segment description
MP	Milepost
Length	Distance between upnode and current node
Average volume	Average volume per segment
Lanes	Number of lanes
Speed	Speed
AV_TT	Length / Speed
TT Index	Av_TT/(TT at the speed limit)
PCNT ON-TIME	The ratio of (Av TT =< (TT@(SpeedLimit-10mph)))/(Entire occasion)
95% TT	95th percentile of Link AV_TT

Buffer Index	If (95% TT < LINK AV_TT) then 0
	Else
	((95% TT)- (Link AV_TT))/ (LINK AV_TT)

Ontime Delay (Min/Veh)	When AvTT > TT@(SpeedLimit-10mph) Sum of ((AvTT-TT@(SpeedLimit-10mph))*Vol)
Congestion Delay (Min/Veh)	When AvTT > 1.5*TT@(SpeedLimit) Sum of ((AvTT-1.5*TT@(SpeedLimit))*5min Vol)

4.5.6 All data fields in TSS Station Level Report

Table 7 describes each column of the *All Data Fields* report.

Table 7: Contents of the *All Data Fields* Report

Column	Data Type
DAY	Date of collection (MM/DD/YYYY)
TIME	Time Stamp (HH24:MM:SS)
STATION_ID	Station ID (5 characters)
FWY_SPD	Volume weighted average speed for mainline
FWY_VOL	Total count for mainline
FWY_OCC	Average occupancy for mainline (unweighted)
SPD_CV	Speed standard deviation / mean for all mainline lanes
VOL_RATIO	Highest lane volume / lowest lane volume for mainline lanes
SPD_RATIO	Highest lane speed / lowest lane speed for mainline lanes
ENTRY_VOL	Sum of entrance ramp counts
EXIT_VOL	Sum of exit ramp counts
FWY_QA	Currently the percent of expected observations that were received during the time period.
ENTRY_QA	
EXIT_QA	
HOV_SPD	Corresponding measures of freeway speed, volume occupancy and data quality in the HOV lanes, if present
HOV_VOL	
HOV_OCC	
HOV_QA	

4.5.7 Traffic counts in TSS Station Level Report

Table 8 describes each column of the *Traffic Counts* report.

Table 8: Contents of Traffic Counts Report

Column	Data Type
DATE	Date of collection (MM/DD/YYYY)
TIME	Time Stamp (HH24:MM:SS)
Facility	Facility
Staiton_ID	Station ID
Num_of_Lanes	Number of lanes
Direction	Direction
Station_desc	Station Description
Station_MP	Station Milepost
Count_Station	Count Station ID
Total	Total thru volume (Sum of lane1~lane6 volume)
Lane1 Vol	Lane1 Volume
Lane2 Vol	Lane2 Volume
Lane3 Vol	Lane3 Volume
Lane4 Vol	Lane4 Volume
Lane5 Vol	Lane5 Volume
Lane6 Vol	Lane6 Volume
Balance	(Max lane volume)/(Min non-zero lane volume)
FWY_QA	Freeway QA
ON_RAMP1	On Ramp1 Volume
ON_RAMP2	On Ramp2 Volume
ON_RAMP3	On Ramp3 Volume
ON_RAMP_QA	On Ramp QA

OFF_RAMP1	Off Ramp1 Volume
OFF_RAMP1	Off Ramp2 Volume
OFF_RAMP1	Off Ramp3 Volume
OFF_RAMP_QA	On Ramp QA
COUNTY	County number

4.5.8 Maximum Flow Rate in TSS Station Level Report

Table 9 describes each column of the *Maximum Flow Rate* report.

Table 9: Contents of Maximum Flow Report

Column	Data Type
DATE	Date of collection (MM/DD/YYYY)
Facility	Facility ID
Staiton_ID	Station ID
Direction	Direction
Station Description	Station Description
Station_MP	Station Milepost
Lane num	Lane number with max flow
Max flow	Max flow
Max time	Time when max flow occurs

4.5.9 Effective Vehicle Length in theTSS Station Level Report

Table 10 describes each column of the *Effective Vehicle Length* report.

Table 10: Contents of Effective Vehicle Length Report

Column	Data Type
#	Column number
DATE	Date of collection (MM/DD/YYYY)
TIME	Time Stamp (HH24:MM:SS)
Facility	Facility ID
Staiton_ID	Station ID
Direction	Direction
Station Desc	Station Description
Station_MP	Station Milepost
Lane1 Vol	Lane1 Volume
Lane1 Spd	Lane1 Speed
Lane1 Occ	Lane1 Occupancy
Lane1 Eff Der Length	Lane1 effective vehicle length (SPD * 5280) * (OCC /100) / VOL
Lane2 Vol	Lane2 Volume
Lane2 Spd	Lane2 Speed
Lane2 Occ	Lane2 Occupancy
Lane2 Eff Der Length	Lane2 effective vehicle length (SPD * 5280) * (OCC /100) / VOL
Lane3 Vol	Lane3 Volume
Lane3 Spd	Lane3 Speed
Lane3 Occ	Lane3 Occupancy
Lane3 Eff Der Length	Lane3 effective vehicle length (SPD * 5280) * (OCC /100) / VOL
Lane4 Vol	Lane4 Volume

Lane4 Spd	Lane4 Speed
Lane4 Occ	Lane4 Occupancy
Lane4 Eff Der Length	Lane4 effective vehicle length (SPD * 5280) * (OCC /100) / VOL
Lane5 Vol	Lane5 Volume
Lane5 Spd	Lane5 Speed
Lane5 Occ	Lane5 Occupancy
Lane5 Eff Der Length	Lane5 effective vehicle length (SPD * 5280) * (OCC /100) / VOL
Lane6 Vol	Lane6 Volume
Lane6 Spd	Lane6 Speed
Lane6 Occ	Lane6 Occupancy
Lane6 Eff Der Length	Lane6 effective vehicle length (SPD * 5280) * (OCC /100) / VOL

4.5.10 Contact Information

As shown in Figure 16, the project contact information is presented on this page.

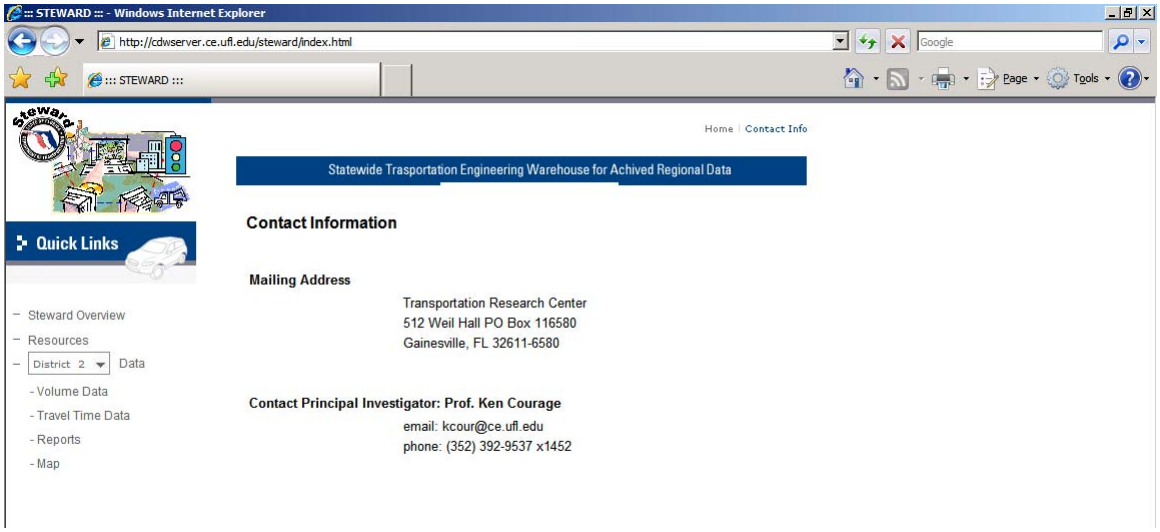


Figure 16: STEWARD contact information web page

5 Examples

Three examples are provided to illustrate how to use the STEWARD web site.

5.1 Example 1 Retrieve all data fields report from the TSS facility-level report

This example demonstrates how to retrieve all data field report from TSS facility-level report. The requirements are given as follows:

- Retrieve the lane-by-lane traffic volume data.
- Time period: One day data (Jan 23rd, 2009) at morning rush hour (7:00AM to 10:00AM)
- Location: District4 I-95 NB

Step 1 Open the STEWARD web site (http://cdwserver.ce.ufl.edu/steward/index.html) and select District 4 from the District pull-down menu (Figure 17).

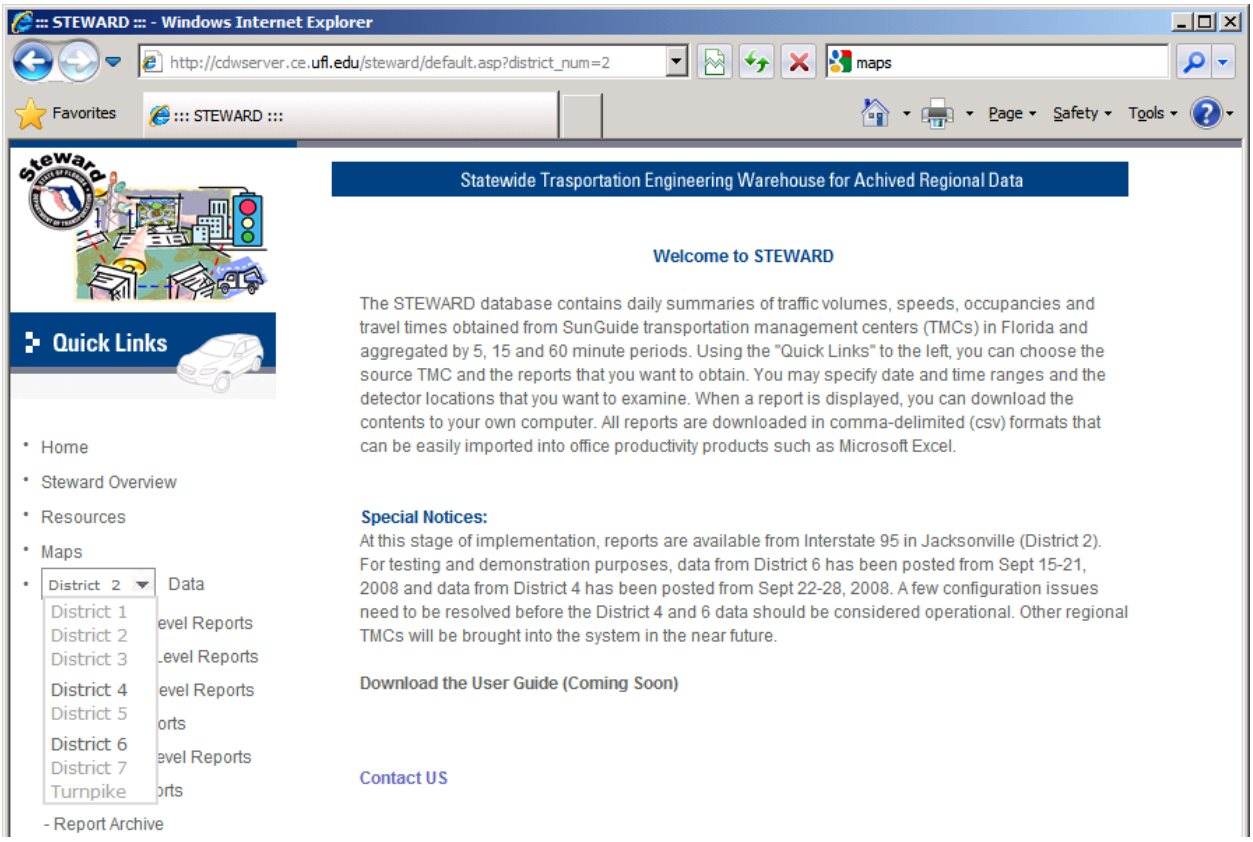


Figure 17: Change the district from the pull-down menu

Step 2 Select the TSS Facility-level report from the left pane. Click the TSS Facility-level report in the left pane and check the web pate shows as like Figure 18.

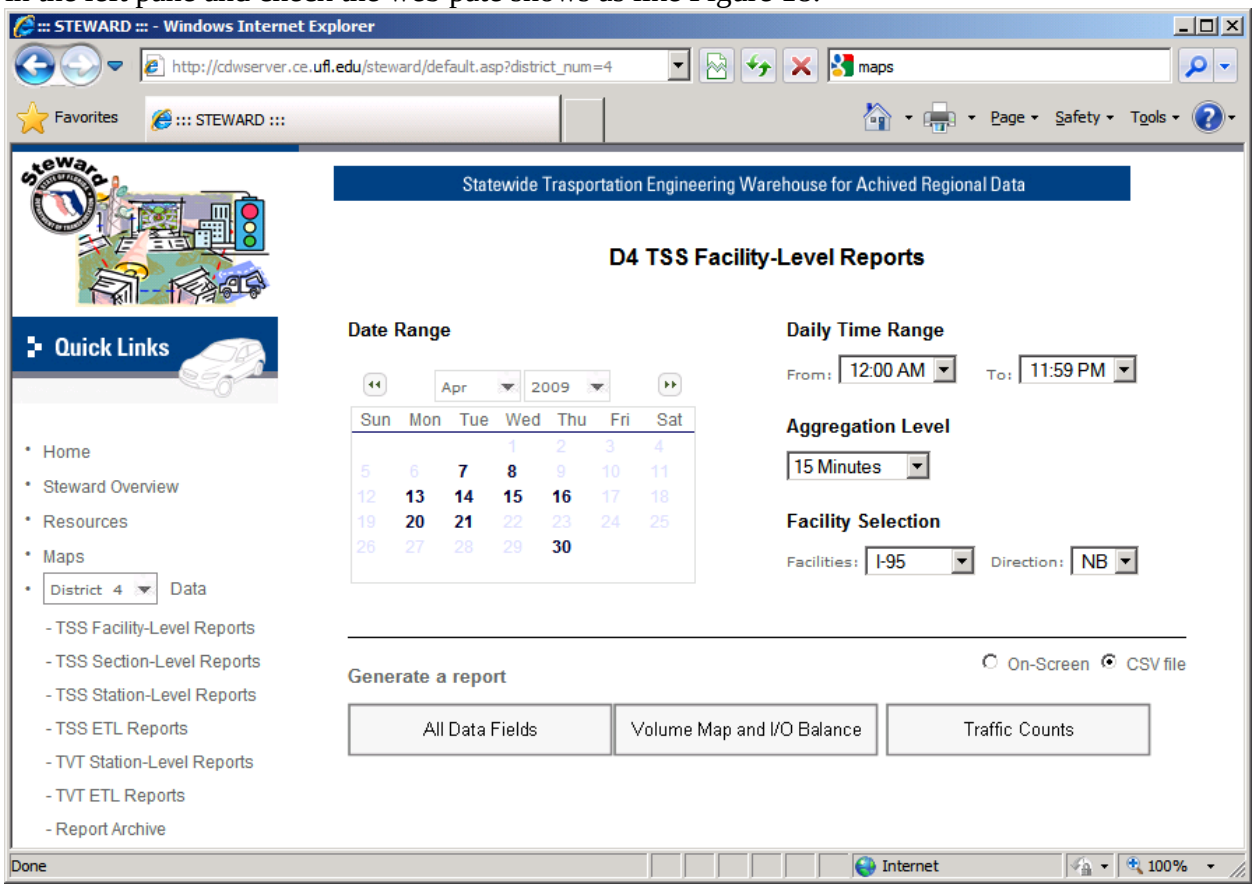


Figure 18: Select the TSS facility-level report from the left pane

Step3 Set the input parameters: Data Range, Daily Time Range, Aggregation Level and Facility Selection as like Figure 19.

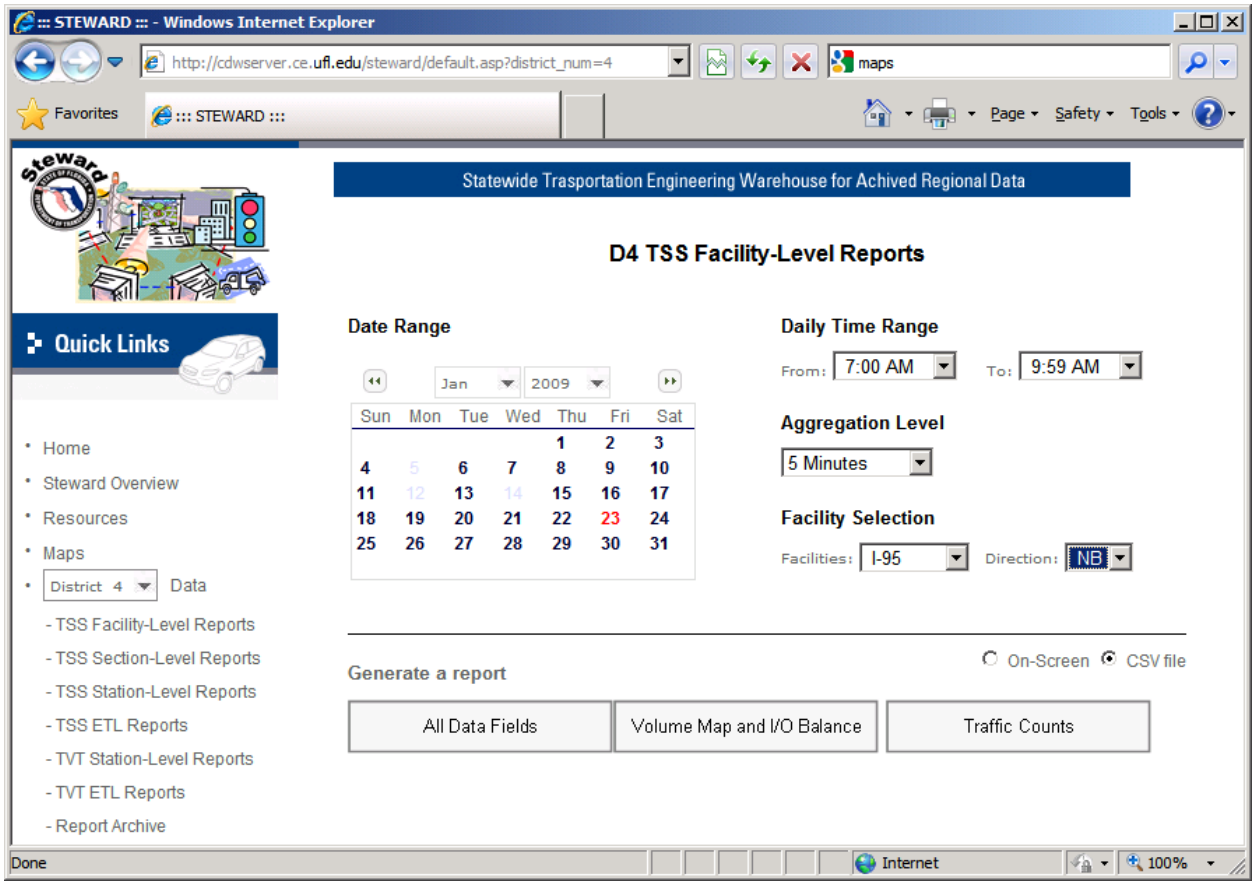


Figure 19: Set the input parameters for data retrieval

Step4 Click the *All Data Fields* button to create the reports. It will take several seconds to minutes to create the reports. The output window will be displayed as shown in Figure 20. The newly created report will be found in the facility level -*All Data Fields* list. Also the file name will be located at the top of the list. This list displays the generated reports in reverse chronological order. The output file can be displayed by clicking the file name or saved to a local computer by clicking the right mouse button.

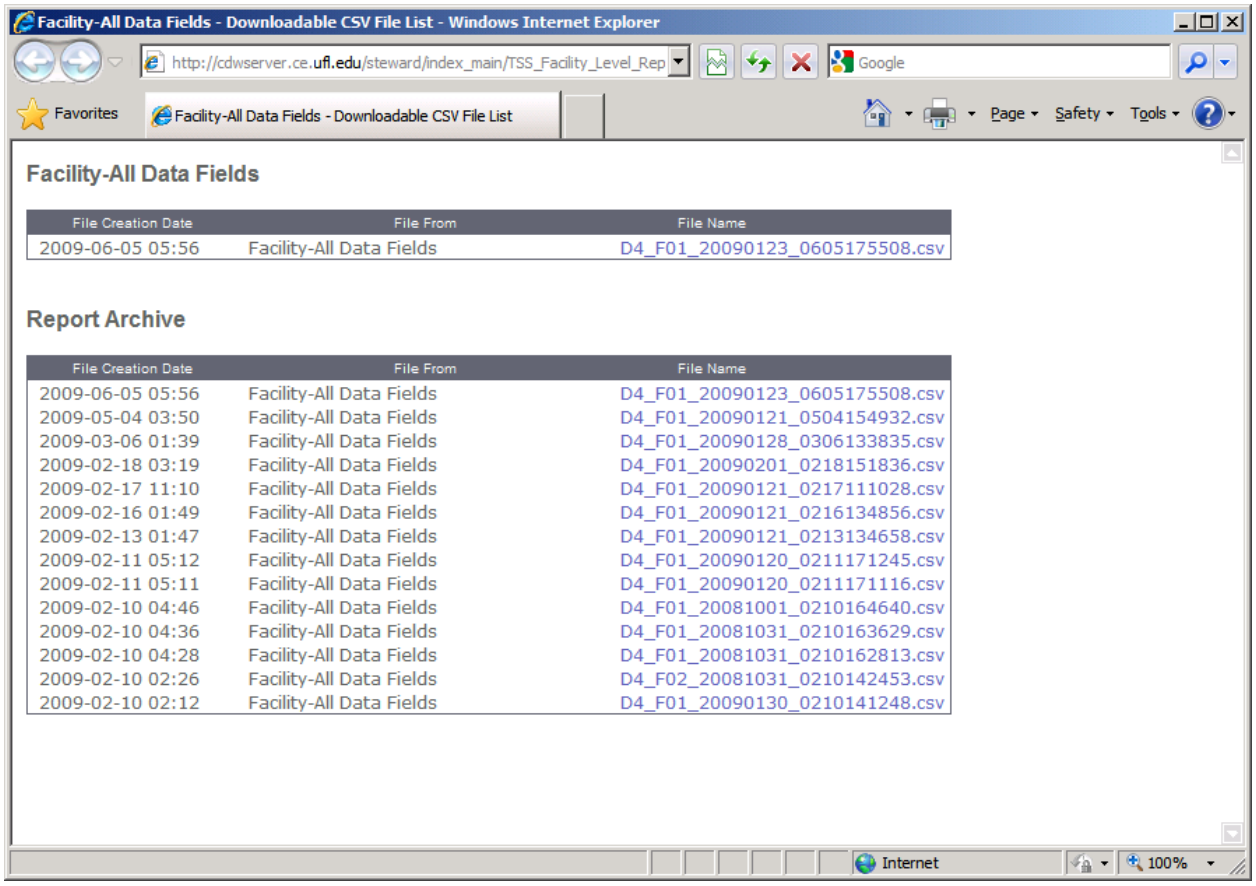


Figure 20: Output results to create the *All Data Fields* report

Step5 This file is also shown in the *Report Archive* list. As shown in Figure 21, click on the report archive in the left pane, and a downloadable CVS file list will be displayed in reverse chronological order. The same file (D4_F01_20090123_0605175508.csv) is displayed at the top of the list. This window will save the time for the user to create the same report repeatedly.

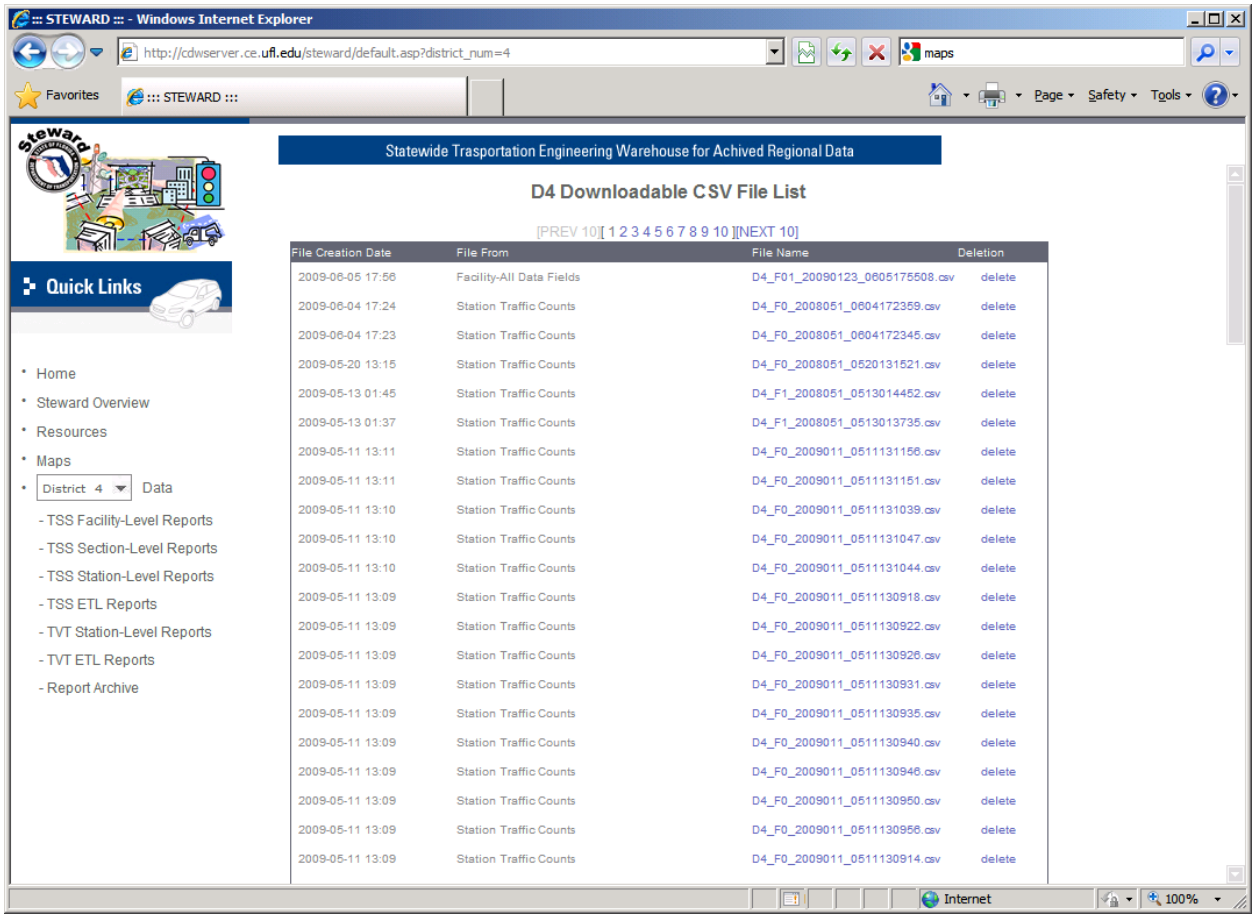


Figure 21: Newly-created report in the report archive

5.2 Example2 Retrieve Travel time reliability report from TSS section-level report

This example demonstrates how to retrieve travel time reliability report from TSS section-level report. The requirements are given as follows:

- Retrieve the travel time reliability report.
- Time period: One month data (Dec. 2008) at evening rush hour (4:00PM to 7:00PM)
- Location: District6 SR 286 SB from *West of NW 27th Ave* to *West of NW 47th Ave*
- Report format: On-screen

Step 1 Open the STEWARD web site (http://cdwserver.ce.ufl.edu/steward/index.html) and select *District 6* from the district pull-down menu (Figure 22).

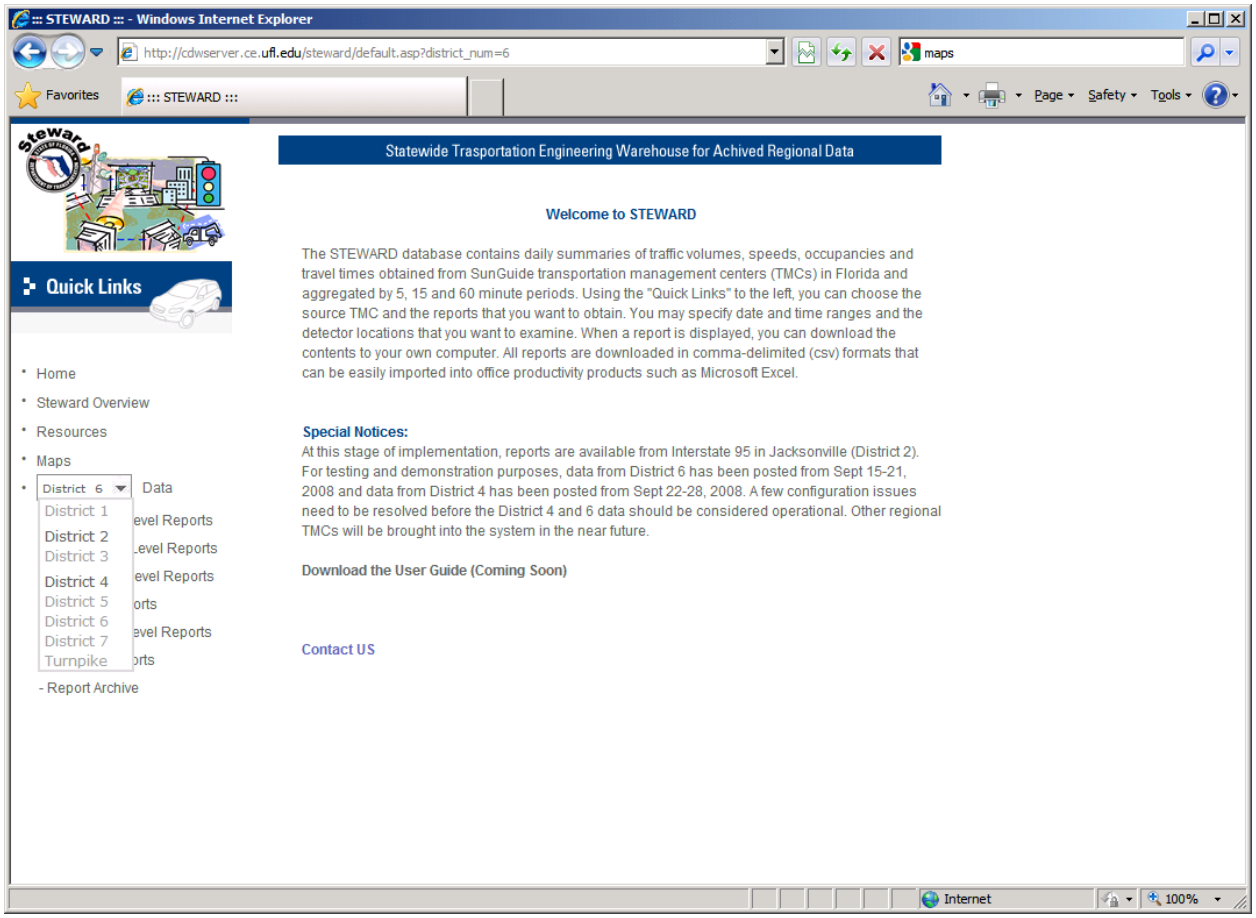


Figure 22: Change the district from the pull-down menu

Step 2 Select *TSS Section-Level Reports* from the left pane. The screen will appear as shown in Figure 23.

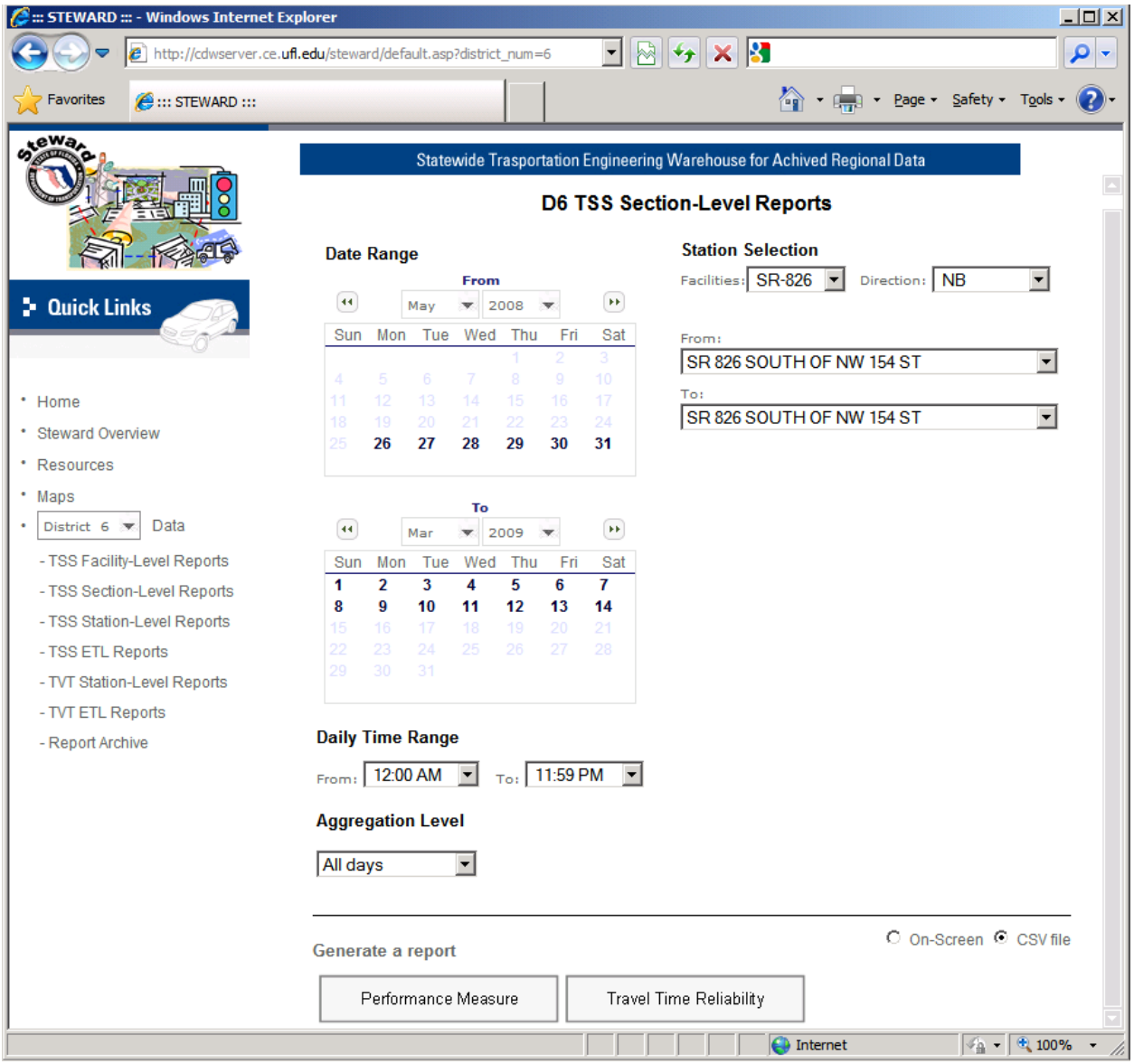


Figure 23 Select the TSS section-level report from the left pane

Step3 Set the input parameters: Data Range, Daily Time Range, Aggregation Level and Station Selection as shown in Figure 24.

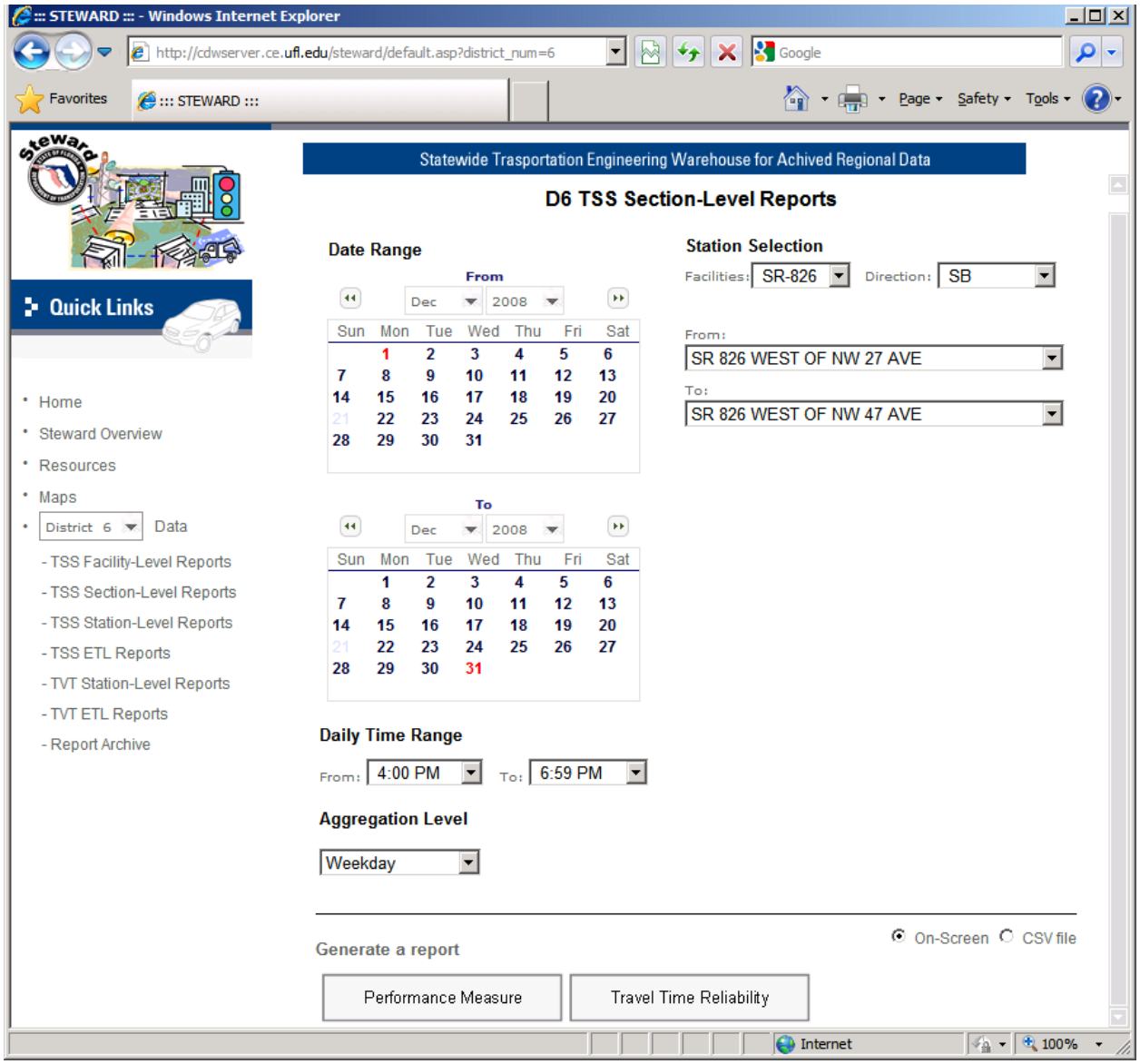
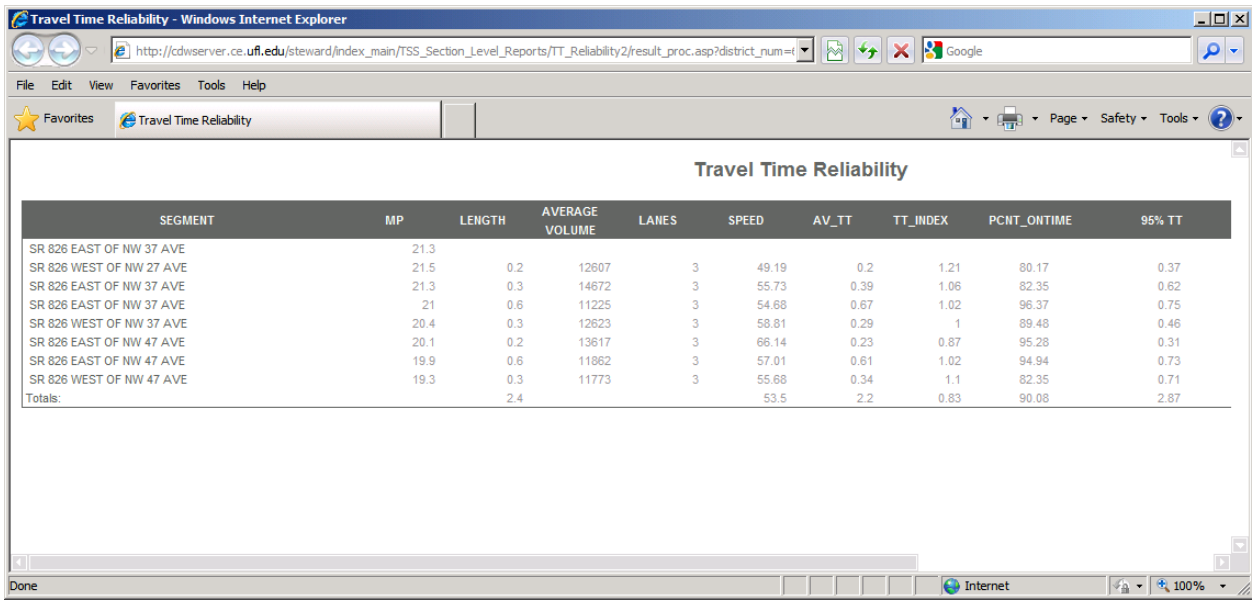


Figure 24: Set the input parameters for data retrieval

Step4 Click the *Travel Time Reliability* button to create the reports. It will take several seconds to create the reports. The output window will be displayed as shown in Figure 25. Users can verify the result from the screen directly.



Travel Time Reliability

SEGMENT	MP	LENGTH	AVERAGE VOLUME	LANES	SPEED	AV_TT	TT_INDEX	PCNT_ONTIME	95% TT
SR 826 EAST OF NW 37 AVE	21.3								
SR 826 WEST OF NW 27 AVE	21.5	0.2	12607	3	49.19	0.2	1.21	80.17	0.37
SR 826 EAST OF NW 37 AVE	21.3	0.3	14672	3	55.73	0.39	1.06	82.35	0.62
SR 826 EAST OF NW 37 AVE	21	0.6	11225	3	54.68	0.67	1.02	96.37	0.75
SR 826 WEST OF NW 37 AVE	20.4	0.3	12623	3	58.81	0.29	1	89.48	0.46
SR 826 EAST OF NW 47 AVE	20.1	0.2	13617	3	66.14	0.23	0.87	95.28	0.31
SR 826 EAST OF NW 47 AVE	19.9	0.6	11862	3	57.01	0.61	1.02	94.94	0.73
SR 826 WEST OF NW 47 AVE	19.3	0.3	11773	3	55.68	0.34	1.1	82.35	0.71
Totals:		2.4			53.5	2.2	0.83	90.08	2.87

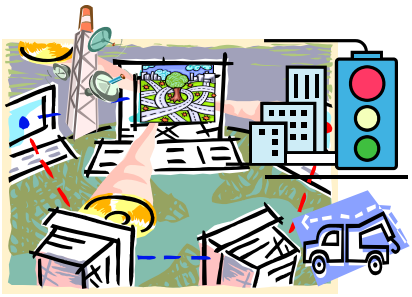
Figure 25: Output results to create travel time reliability report

Appendix 4
STEWARD Software Utility Documentation

This appendix describes several utility programs developed for use with Statewide Transportation Engineering Warehouse for Archived Regional Data (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

Appendix 4a
SunETLUtility



Statewide Transportation Engineering Warehouse for Archived Regional Data (STEWARD)

Florida Department of Transportation
University of Florida
Transportation Research Center

ETL Utility Program Documentation

1 Purpose

The Extract, Transform, and Load (ETL) utility performs the first steps of the ETL process by which daily archive data from the SunGuide Transportation Management Centers (TMCs) are acquired and transformed into the STEWARD database format. It was developed for use by personnel responsible for loading archive data into the STEWARD database. This documentation is provided primarily for those personnel.

2 Concept of Operations

An overview of the ETL Utility data flow is presented in Figure 1. There are two inputs and four outputs associated with this program. The inputs and outputs are described here.

2.1 Raw SunGuide Archive Data

The raw SunGuide archive data are provided daily from the SunGuide TMC. The method of acquisition may vary among districts. The important point is that the raw archive data must exist in a specified folder (see the section on file locations). The SunGuide file naming convention for archive data files is:

TSS-mmddyyyy--1.dat for TSS archive files

Each TSS archive file contains multiple records with the following fields:

- timestamp
- detector_id
- lane_id
- speed
- volume
- occupancy

Each TSS record covers 20 seconds or 30 seconds of operation for one lane, depending on the district.

2.2 Facility Description Spreadsheets

The configuration of the facility data is described in the STEWARD Final Report document. The purpose of the facility description data is to relate the contents of the raw archive data files

to the geometrics and geography of the facility monitored by the TMC. The complete facility description is described in an Excel workbook containing three spreadsheets:

- General facility data
- Detector station data
- Detector lane data

The outputs of the ETL Utility are described here.

2.3 Daily Reports

Diagnostic reports are generated to indicate the health of the detector and communications systems. The diagnostic information produced by the conversion process is presented in some detail in the daily reports. These reports identify the actual stations, ramps and lanes represented in the diagnostic summary. The daily report will be useful in assessing the detector maintenance and service requirements. A sample of the Daily Report is shown in Figure 2.

2.4 History Log (CSV Format)

The diagnostic summary produced by the conversion process is appended to the conversion log each time the conversion is performed. By this process, the conversion log presents a history of all of the daily conversions and their diagnostic summary information.

2.5 Grouped Data Summaries by Lane (CSV format)

The lane-specific records in the 20 or 30 second converted data files are grouped into 5, 15, and 60 minute summaries with the same format and loaded into the STEWARD database. These files are can be downloaded as reports from the Steward report section. One minute aggregation levels are also retained in CD or DVD storage to support subsequent requests from researchers.

2.6 Grouped Data Summaries by Station (CSV Format)

The station-based summaries are produced from the lane-based summaries in 5, 15, and 60 minute aggregation levels and loaded into the STEWARD database. These summaries are much more compact and useful because the performance measures to be reported apply to the detection stations as a whole and not to the individual lanes.

2.7 Quality Control

Data quality rules were developed and implemented inside the ETL Utility. During the ETL process, nine verification rules are applied into the raw SunGuide archive data. The traffic data that passes these verification rules are processed and grouped for the next steps. The failed traffic data are recorded into the converted data file with an error ID but are not used for the data processing or grouping. These test results are also summarized into the daily reports. The details of these test rules are described in the Phase 3 final report.

Table 1 Sample input/output file

Sample input data file	Output files
TSS-03032009--1.dat	• Converted file ConvertedData\D2_TSS-03032009-0.csv • Conversion log file ConvertedData\D2_TSS-03032009-DailyReport.Log • Station-level files GroupData\D2_TSS-03032009-1S.csv GroupData\D2_TSS-03032009-5S.csv GroupData\D2_TSS-03032009-15S.csv GroupData\D2_TSS-03032009-60S.csv • Lane-level files GroupData\D2_TSS-03032009-1L.csv GroupData\D2_TSS-03032009-5L.csv GroupData\D2_TSS-03032009-15L.csv GroupData\D2_TSS-03032009-60L.csv • Conversion log file (updated) TSSConversionHistory-D2.csv

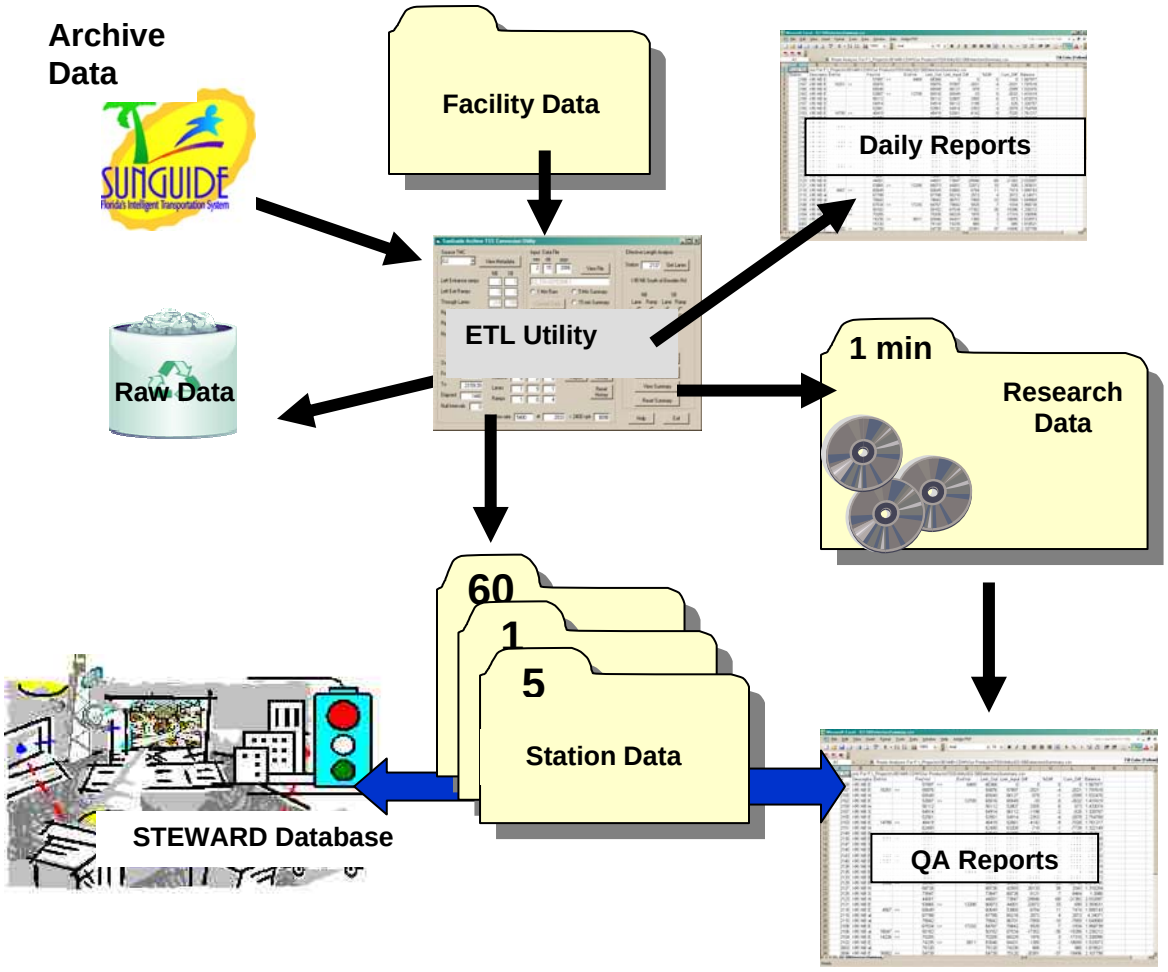


Figure 1: ETL Utility data flow

```
Orphan Station. DS-2010NS (Lane: DS-2010NS-link2-S-lane1) is not in the TSS Map.
Orphan Station. DS-0092BN (Lane: DS-0092BN-link1-ML-lane4) is not in the TSS Map.

OFFLINE STATIONS: [ ] Indicates Ramps with no Detectors
   600122,  I-95 AT SW 3 ST ON RAMP FROM SR 970
   600141,  I-95 AT SW 3 ST OFF RAMP TO SR 970

NEGATIVE OR ZERO DETECTOR SCAN INTERVALS:
   -40  Sec:  SR 826 WEST OF NW 67 AVE, Lane ID 61041131 at 03:55:47
   -40  Sec:  SR 826 EAST OF NW 47 AVE, Lane ID 61021133 at 04:04:47

NULL MINUTES: NO REPORT FROM ANY LANE

NULL STATIONS: NO REPORT FROM ANY LANE
   600012,  I-95 NORTH OF SW 32 RD
   600021,  I-95 NORTH OF SW 32 RD

SCAN INTERVAL DISTRIBUTION
   <=  0      0.001%
 1 -  15      0.003%
 15 -  29     99.905%
...

SCAN INTERVALS (less than (SCAN_INTERVAL-5) seconds and positive values BY HOUR)
 0:00 -  1:00     0.000%
 1:00 -  2:00     0.013%
...

SCAN INTERVALS (larger than (SCAN_INTERVAL+5) seconds BY HOUR)
 0:00 -  1:00     0.044%
 1:00 -  2:00     0.068%
...

MISSED SCANS BY HOUR
 0:00 -  1:00       0
 1:00 -  2:00       0
...

Maximum volume error: 2263
Maximum occupancy error: 34
Maximum speed error: 452
Multi-variate consistency (zero speed with non-zero volume) error: 123
Multi-variate consistency (zero volume with non-zero speed) error: 108
Multi-variate consistency (zero volume, zero speed with non-zero occupancy) error: 3155
Truncated occupancy values of zero error: 91
Maximum estimated density error: 635
All zero or stuck data for 5min: 70895, All zero data:70882
Total number of valid data: 1410526

TRANSFORMATION SUMMARY FOR D6_TSS-01012009-0
From 00:00:03 to 23:59:58 Elapsed Time  1440  Minutes
Total Records Processed 1781384
Total vehicles processed  4696811
Negative/zero Scan Intervals:  10
Null Minutes  0
Orphan  Stations  10
      Lanes     0
      Ramps     0
  Null  Stations  52
      Lanes     5
      Ramps     0
Offline Stations  72
      Lanes     0
      Ramps     0
```

Figure 1 Daily Report Sample

3 Loading the STEWARD Database

The data from the group files produced by the ETL Utility must be loaded into the STEWARD database. The method by which this step is accomplished is documented in Appendix 2 of the Phase 2 Final Report.

4 File Locations

The ETL Utility program expects and/or produces files in the following folders. All of the folders listed except for the program folder must be created as subfolders of the program folder.

- Program Folder
  - Program (SunETLUtility.exe)
  - This document (SunETLUtility.pdf)

- "FacilityData" Folder
  This folder must contain all of the facility data files as Excel workbooks. The facility data requirements are explained in detail in the STEWARD Final Report.
  - District 2 facility data file: TSSMaster-D2.xls
  - District 4 facility data file: TSSMaster-D4.xls
  - District 5 facility data file: TSSMaster-D5.xls
  - District 6 facility data file: TSSMaster-D6.xls
  - District 7 facility data file: TSSMaster-D7.xls
- "ToConvert" Folder
  The Raw SunGuide TSS data files from the district TMCs must be placed in this folder
- "ConvertedData" Folder
  ETL Utility writes three types of files to this folder during the conversion process:
  - Converted data files: Raw input data are converted into Steward data format but there are no time or station aggregation. These files are used for other data aggregations (time and station).
  - Daily conversion reports (TXT format)
  - Conversion history, 1 record per day (CSV)
- "GroupData" Folder
  Sun ETLUtility writes the following files into this folder
  - Lane based grouped data files for 1, 5, 15 and 60 minute intervals (same format as converted files except for time stamp)
  - Station-based grouped data files for 1, 5, 15 and 60 minute intervals, in the format required by the STEWARD database

5 Operation

When the program is run, the main screen appears as shown in Figure 3.

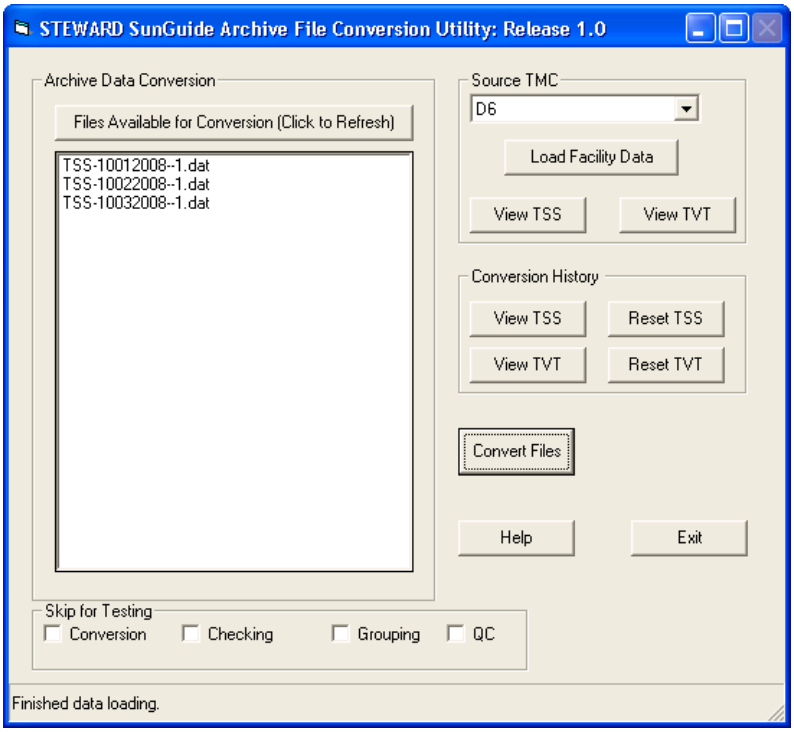


Figure 3: SunETLUtility main screen

5.1 "Archive Data Conversion" Frame

All of the files in the *To Convert* folder are shown in the Archive *Data Conversion* list box. Remember, it is up to the user to ensure that the files to be converted are placed in the "To Convert" folder before conversion and removed to some other location after the conversion has been completed. You may click the box above this list to refresh the list after you have made changes.

5.2 Source TMC Frame

The drop down list in the *Source TMC* frame will display all of the TMCs for which facility data files exist in the *FacilityData* folder. The facility data for the chosen TMC will be applied to the archive data files to carry out the transformation to the STEWARD database format. Command buttons in this frame may be used to load and view the facility configuration data.

5.3 Conversion History Frame

This frame contains command buttons that let you view and reset the conversion history files.

5.4 Converting files

Clicking the "Convert Files" button will initiate the conversion of all of the archive data in the *ToConvert* folder into the STEWARD database format. Conversion history and daily report files will be created as a part of this process.

6 Command Line Execution for the ETLUtility

The ETL Utility supports command line execution to support the loading process in the Oracle Warehouse Builder. Command line parameters specify the district number, input directory and output directory. The input directory indicates the directory that includes the *ToConvert* and *FacilityData* folders for input data files and facility data files. The output directory indicates the directory that includes the *ConvertedData* and *GroupData* for converted and aggregated output data files.

The calling syntax is:
>> SunETLUtility /d"arg1" /i"arg2" /o"arg3"
The arguments are:

- arg1: District name
- arg2: Input directory path and name
- arg3: Output directory path and name

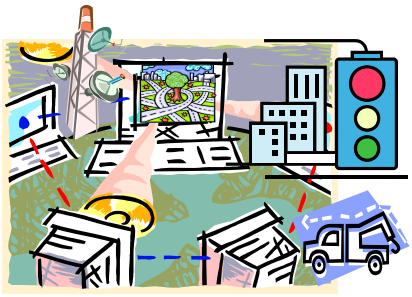
Example:
```
>> SunETLUtility /d"D2" /i"c:\progra~1\trc\SunETLUtility"
/o"c:\progra~1\trc\SunETLUtility"
```

7 Additional Guidance

A couple of points worth remembering:

1. The archive file conversion handles many megabytes of data. This process is time consuming and is best carried out in an unattended mode (e.g., overnight)
2. The SunGuide system does not distinguish between districts in the naming of archive files. On a given date, each district will produce an archive file that cannot be distinguished by its name from other districts. If you are processing files from multiple districts, careful attention must be given to separating the individual files.

Appendix 4b: ITSCounts



Statewide Transportation Engineering Warehouse for Archived Regional Data (STEWARD)

Florida Department of Transportation
University of Florida
Transportation Research Center

Conversion of SunGuide Archive Data to the FDOT Count Formats

ITSCounts Program Release 2.42:
Revision Date Nov 30, 2009

The traffic counts in the SunGuide archives have essentially the same content as the Florida Department of Transportation (FDOT) Statistics Office and District Planning Office traffic count files. There is clearly a potential benefit that could be derived from a mutual exchange of traffic count data between the intelligent transportation system (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. ITSCounts was developed to support traffic count extraction from SunGuide archive data for FDOT counting program use.

Figure 1 shows an example comparison between the daily counts from a Statistics Office station and the upstream (2124) and downstream (2126) traffic sensor subsystem (TSS) stations. 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.

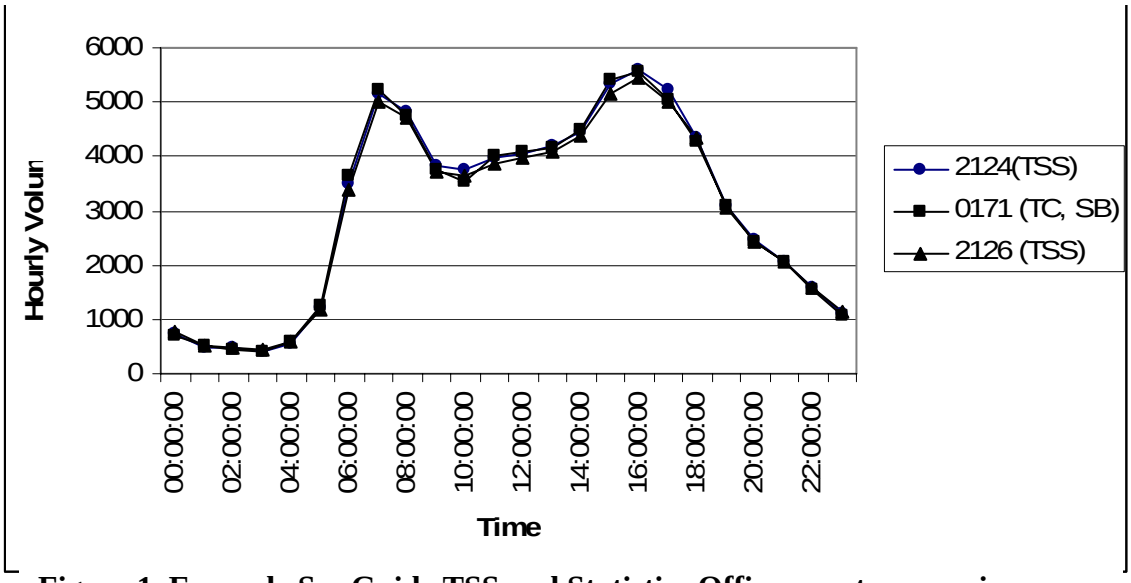


Figure 1. Example SunGuide TSS and Statistics Office count comparison

The Statistics Office Count Format

ITSCounts converts the count data in the SunGuide archive to the Statistics Office and District Office count formats. The Statistics Office format is presented in Table 1, along with the rules used for conversion from the SunGuide archive. Note that each field in this format has a fixed width.

Table 1:Format for FDOT Statistics Office Count Files

TAG	CNT	String constant
CO	County Number	Duval County = 72.
STAT	Station Number	Assigned by Statistics office, e.g., C721234
	Filler	
YY	Survey Year	
MM	Survey Month	
DD	Survey Day	
HR	Survey Hour	End of reporting interval
MN	Survey Minute	End of reporting interval
	Filler	
INT	Survey Interval	60 minutes
	The following values are repeated for each lane up to a total of 8 lanes	
	Filler	
L	Lane Number	TMG standard: right to left. Note that this is inverted from the SunGuide and other ITS standards
	Filler	
IVOL	Interval Volume	As computed

The Statistics Office has also prescribed a standard for counts derived from ITS detectors. The ITS format uses the "ITS" tag in place of the "CNT" tag. This format also includes a validity check for each lane (N = normal, B = bad). As a preliminary criterion for generating ITS files, the detector data are assumed to be normal if the detectors reported volumes for at least 90% of the 20 second SunGuide polling intervals.

District Planning Office Count Format

The district planning offices perform non-continuous counts periodically at selected locations using various types of portable equipment. The SunGuide archive offers a potentially rich source for this type of data. The district count file format differs from the central office format because of differences in the processing software. The district format requires one record for each lane as shown in Table 2:

The TSS lane data contains a field for the count data station. The station number in the STEWARD configuration file begins with a "C" or a "D", followed by the 4 digit number shown in Figure 2. Stations beginning with "C" are processed in the central office format and those beginning with D are processed as district count files. Count files are generated for all lanes with valid data in the station number field. Counts are not generated if the STEWARD configuration field is blank. The prefix letter is stripped from the station number when the count data file is created.

Table 2: District Planning Office Count Format

Field	Length	Comment	Example
County	2 Numeric	FDOT County Number	72
Station	4 Numeric	Supplied by District	0004
Date	8 Numeric	DD/MM/YYYY	01/29/2008
Direction	1 Alpha	Direction of traffic	"E"
Lane	1 Numeric	1=Outside,2=Middle, etc.	1
Survey Type	1 Numeric	2=Vehicle count	2
Survey Program	1 Numeric	9=ITS program	9
Interval	2 Numeric	Minutes	15
Time	4 Numeric	Military time, leading zeroes	0830
Volume	5 Numeric	leading zeros required	000357

Note: Blanks between Fields, Carriage Return/Line Feed at end of each line of data.

Production of Traffic Count Files

An overview of the process by which traffic count files are produced from the SunGuide archive is shown in Figure 2. The numbers on this figure refer to the steps in the process. The steps are described as follows:

1. SunGuide generates a separate archive file for each 24 hour day of operation. The archive reports volume, speed and occupancy for each polling interval. Polling intervals are either 20 sec or 30 sec, depending on the district preference.
2. The extraction, transformation and loading (ETL) utility reformats the raw SunGuide data, adds some facility configuration data and accumulates the results.
3. The results are stored in files accumulated by 5, 15 and 60 minute intervals.
4. The accumulated data records are transferred into the STEWARD database.
5. The STEWARD database is queried manually from an internet site. The data files may be isolated by station, date, etc.
6. The results of the queries are stored as comma-delimited (CSV) files on the local computer. File names are assigned by the user. Be sure to give the file a name that you will recognize later when you select it for processing.
7. The ITSCounts program reads and processes the CSV files.
8. Traffic count files are generated in either the Central Statistics Office or District Planning Office format. The assimilation of these files is performed by FDOT using their own software. The assimilation process is beyond the scope of this document.
9. Additional CSV files are produced for analysis and plotting of the count data using Excel.

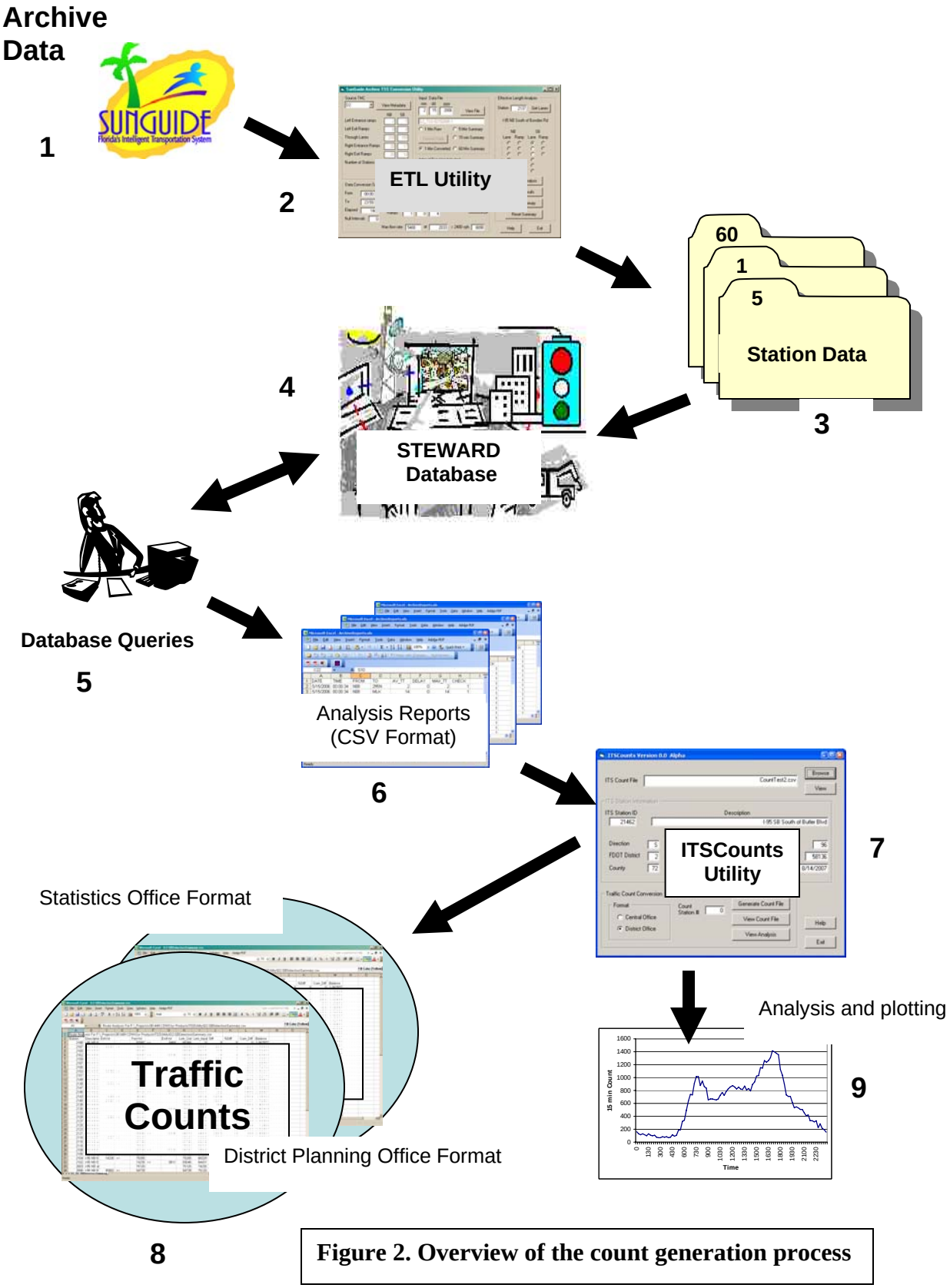


Figure 2. Overview of the count generation process

Assimilation of Traffic Count files by FDOT

The ITSCounts program produces traffic count files as text files in the required FDOT formats. Assimilation of these counts is accomplished by FDOT using their own software. The Statistics Office uses dedicated software that was developed to suit their purposes. The district offices use the Survey Processing Software (SPS) program to incorporate the count files into a statewide database that resides on the FDOT mainframe computer. The details of the assimilation process are beyond the scope of this document.

Installing the ITSCounts Program

ITSCounts is installed on a local computer using the normal Windows installation process. Five sub-folders are created within the program folder during the installation:

- **ITSData**, which contains the CSV files produced from the STEWARD database query in Step 6 above.
- **COCounts,** which contains the traffic counts produced by ITSCounts in the Central Office format.
- **DDCounts,** which contains the directional traffic counts produced by ITSCounts in the District Office format. Each count is associated with an ITS station that covers one direction.
- **StationLists**, which contains one or more Excel files that identify the ITS stations and properties associated with each district count station. The Excel files have been provided for all districts by the Statistics Office.
- **Working,** which contains a parameter file that gives the locations of programs, folders and URLs required by the program. It also contains spreadsheets that identify the ITS stations corresponding to each district traffic count location.

The installer must have read and write access to the program folder. After the program has been installed and set up, the operator must have read and write access permission for all of the folders identified above.

The overall file structure for ITS counts is shown in Figure 3. The installation routines will establish the default folder and program locations. If a different configuration is desired, the locations may be changed from within the program, keeping in mind that the operator must have read and write access to the working folder and all of the data folders.

> Note: You will find that the installation will be much easier if you choose all of the default installation options. The discussion and examples in this document assume that the default options have been chosen.

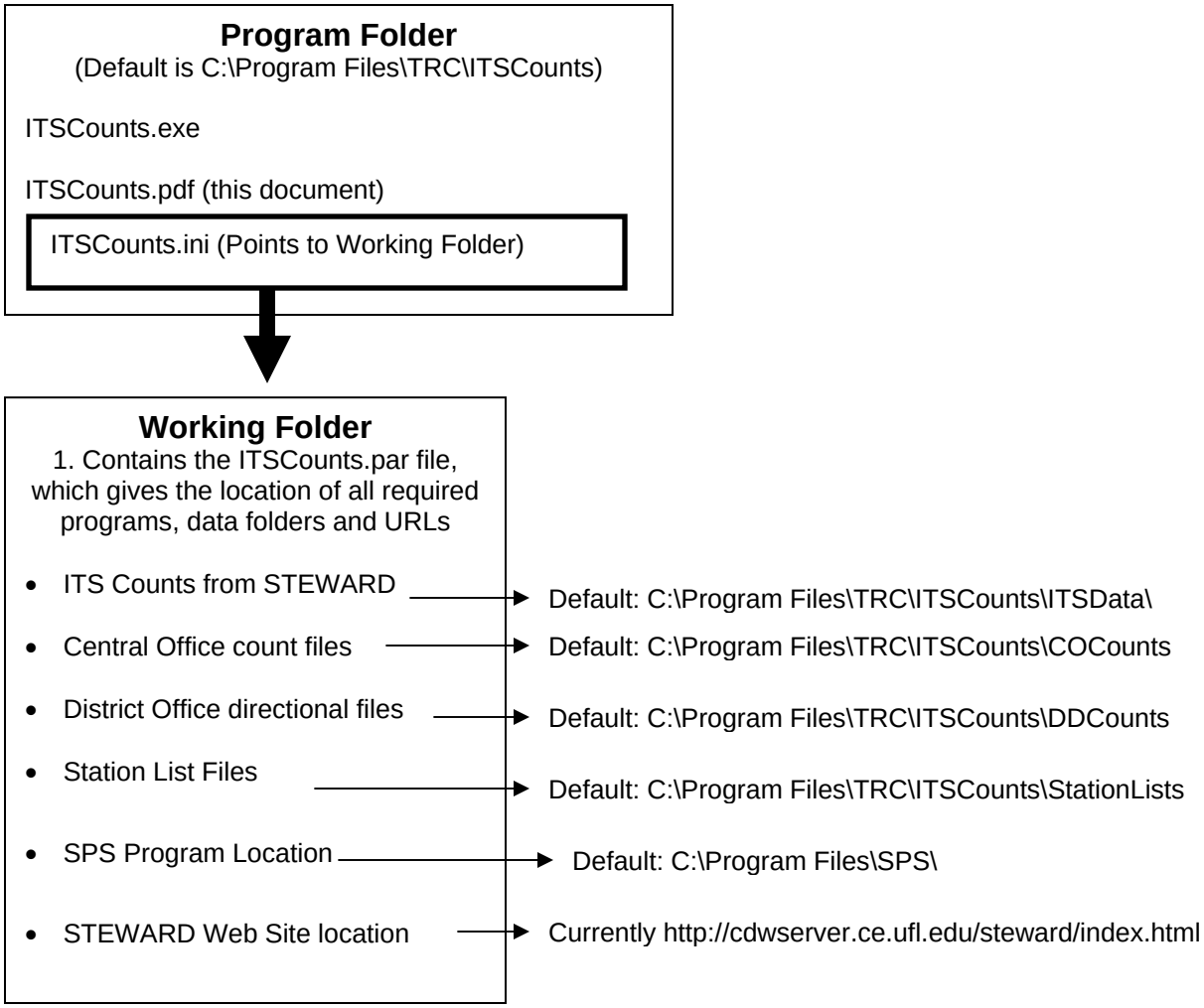
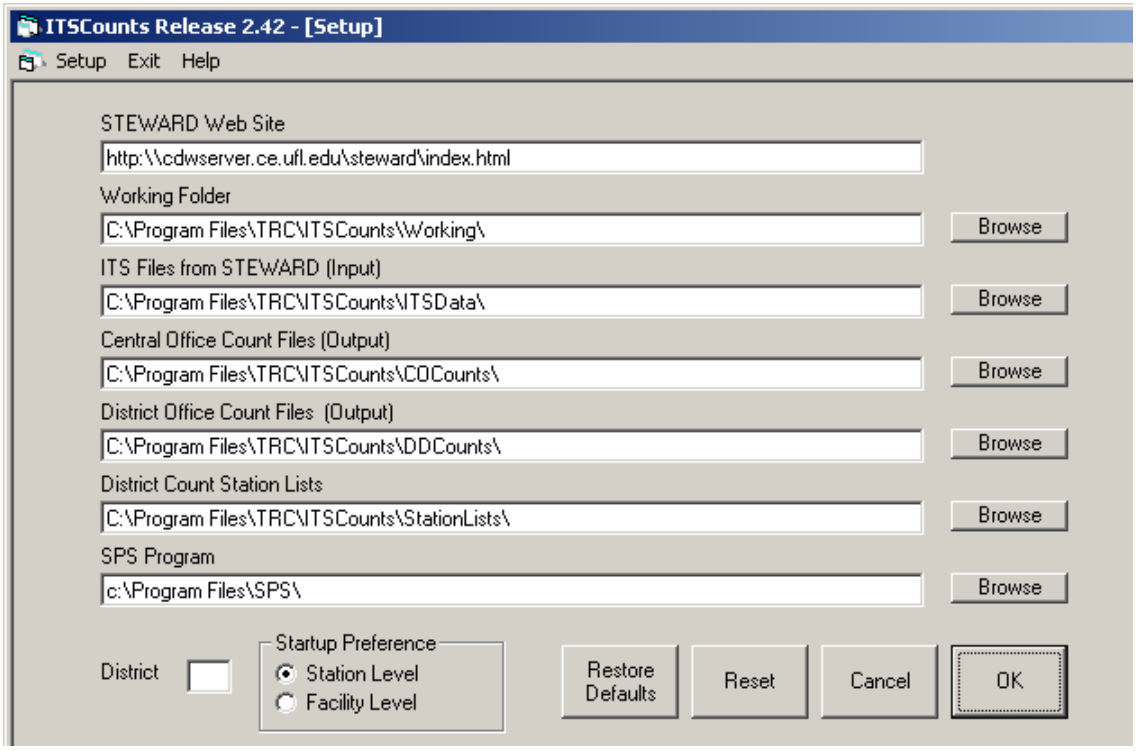


Figure 3. ITSCounts folder structure

First Time Setup for File locations

When you run the program for the first time, you will get a message telling you that the file configuration needs to be set up. Click ***OK*** and you will see the following screen:

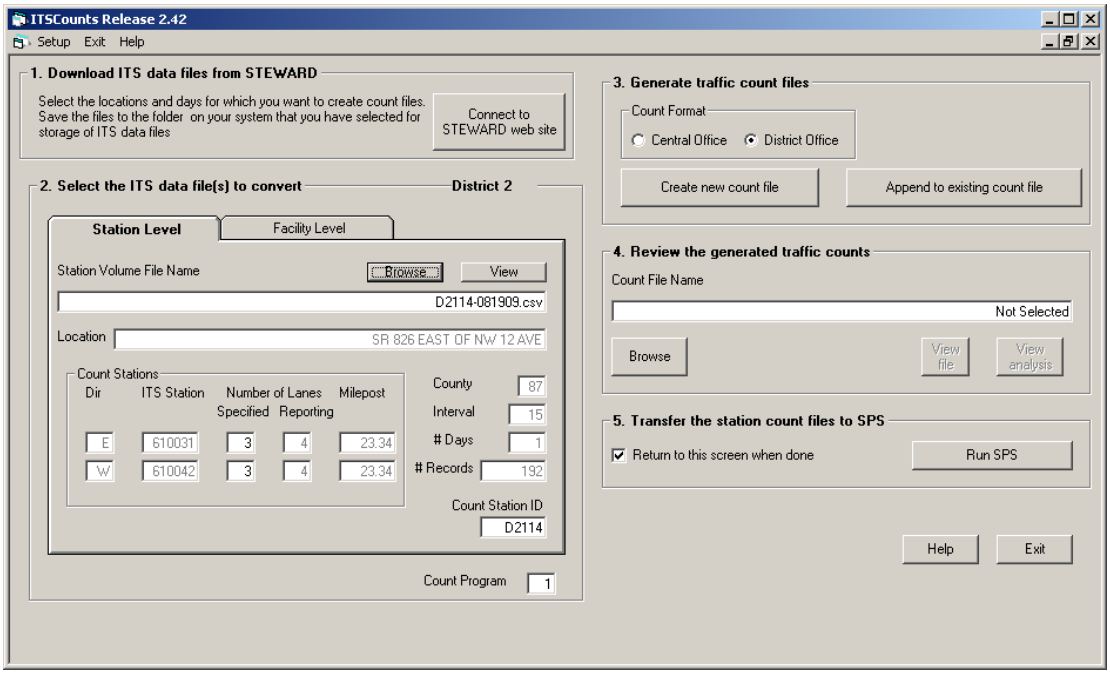


The file locations will contain the default assignments shown in Figure 3 when you first enter this screen. You can browse for each of the items using the ***Browse*** button. The selection will lead you first to the default file locations but you may select any folder that already exists on your system. The step will be very easy if you use the default locations. All of the required folders were created when the program was installed. Click ***OK*** when you are finished.

Don't forget to enter your district number. You can also set a startup preference for station or facility level analysis. The difference will be explained later (See *processing options)*. The level can be changed at any time from the main screen. The startup presence is provided here as a convenience and should be set to the level that you will likely use more often.

Running the ITSCounts Program

The following discussion assumes that you are familiar with basic Windows operations and that you have set up the program using all of the default program and file locations. When you run ITSCounts, the main screen will appear as shown here. Note that there are five group boxes containing GUI elements, each of which is associated with one step in the count file creation.



1. Download ITS data files from the STEWARD web site.
2. Select and convert ITS data files to traffic count files.
3. Generate the traffic count files in the selected FDOT format.
4. Review the generated traffic counts.
5. Transfer the counts to SPS (District counts only).

The ITSCounts program acts like a wizard that guides you through these steps. Each of the steps will be discussed in more detail.

Processing Options for ITS Count Conversion

The STEWARD archive files contain data for a station that covers one direction of a facility. The SPS program requires bidirectional files. There are two options for creating bidirectional district traffic counts from ITS files:

1. Files may be downloaded from the two ITS stations that correspond to a district count station (one for each direction). The ITSCounts program can combine two such files into a single file compatible with SPS.
2. Data for the entire facility may be downloaded to produce multiple SPS files in a batch process.

The two options are illustrated in Figure 4.

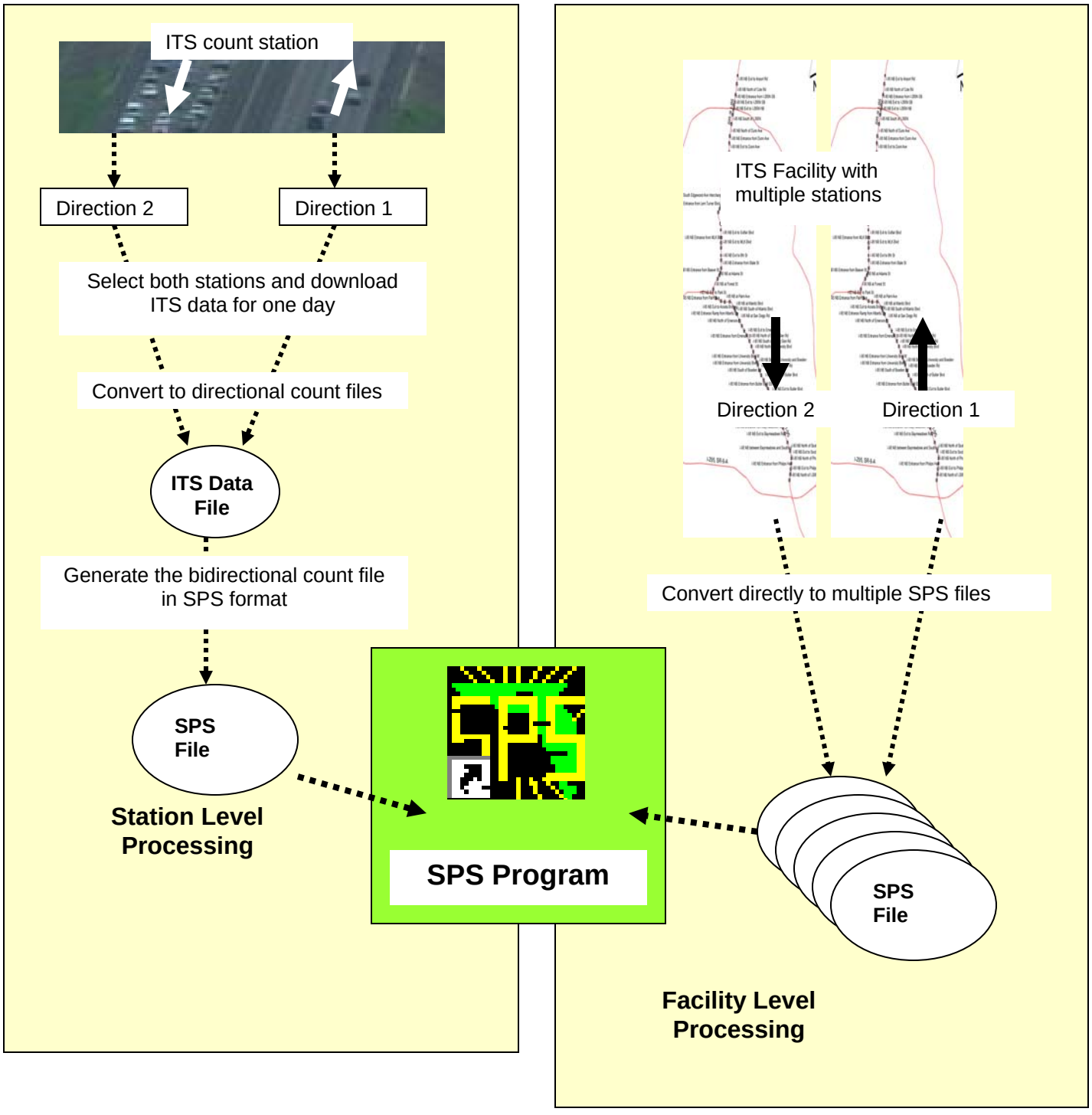


Figure 4. Processing options for district ITS count conversion

The facility level processing is more productive than the station level processing if you are trying to create several SPS count files from the same facility on the same day. On the other hand, if you are interested in a single count station. You will probably find the station level option easier and faster. The station level option also provides more opportunity for review and selection of appropriate days for count extraction.

The Station List File

The *StationList.xls* file is required for facility level processing. This Excel workbook contains three worksheets:

1. The *Freeway worksheet* identifies the ITS stations for both directions of travel associated with each district count station.
2. The *Ramps-1* worksheet identifies the ITS stations associated with each ramp for the "increasing milepost" direction (usually NB or EB).
3. The *Ramps-2* worksheet identifies the ITS stations associated with each ramp for the "decreasing milepost" direction (usually SB or WB).

The first column of each worksheet contains is the "Enable" instruction. Stations are selected for count generation by entering "Y" in this column for each of the stations for which you want to obtain count files. If a station in the list is not selected in this manner, no count file will be generated for it when the station list is processed. As indicated later, you can access this file from the main screen. You can edit the file to change your requests and save the changes under the same file name or a different one, if you prefer. If you save the file under a different name, you must reselect the new station list file to have the changes recognized.

Downloading and Converting ITS files

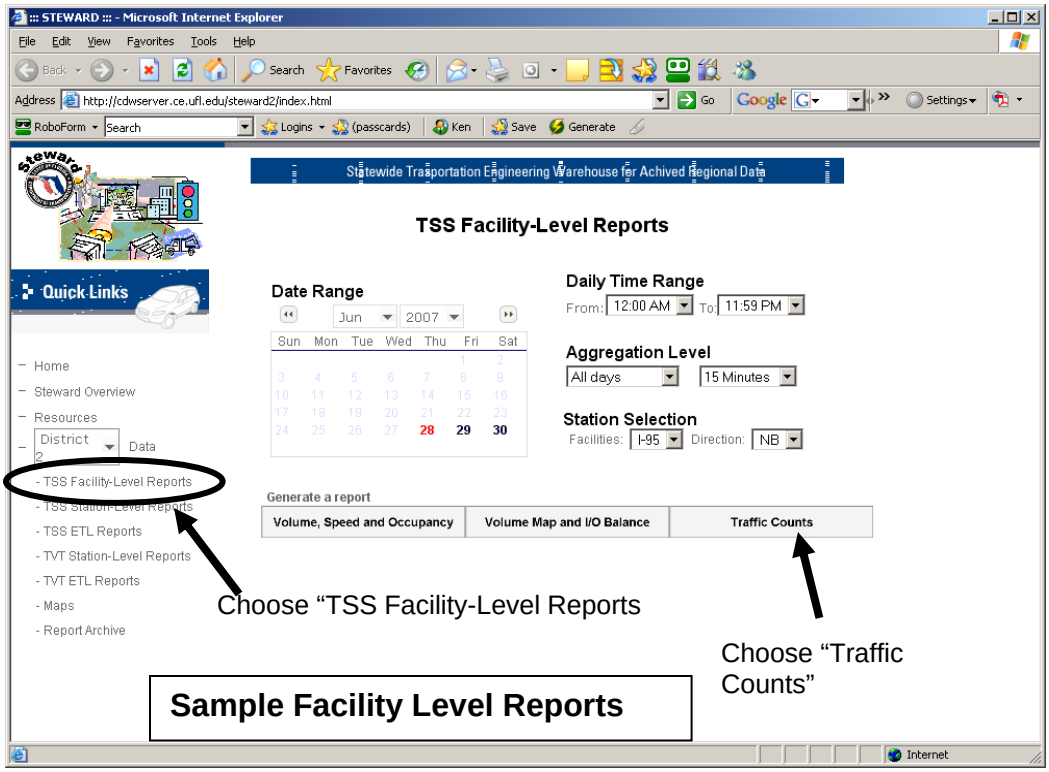
The steps in downloading and converting the ITS data to traffic counts will now be described.

1. Download ITS Data Files from the STEWARD Web Site

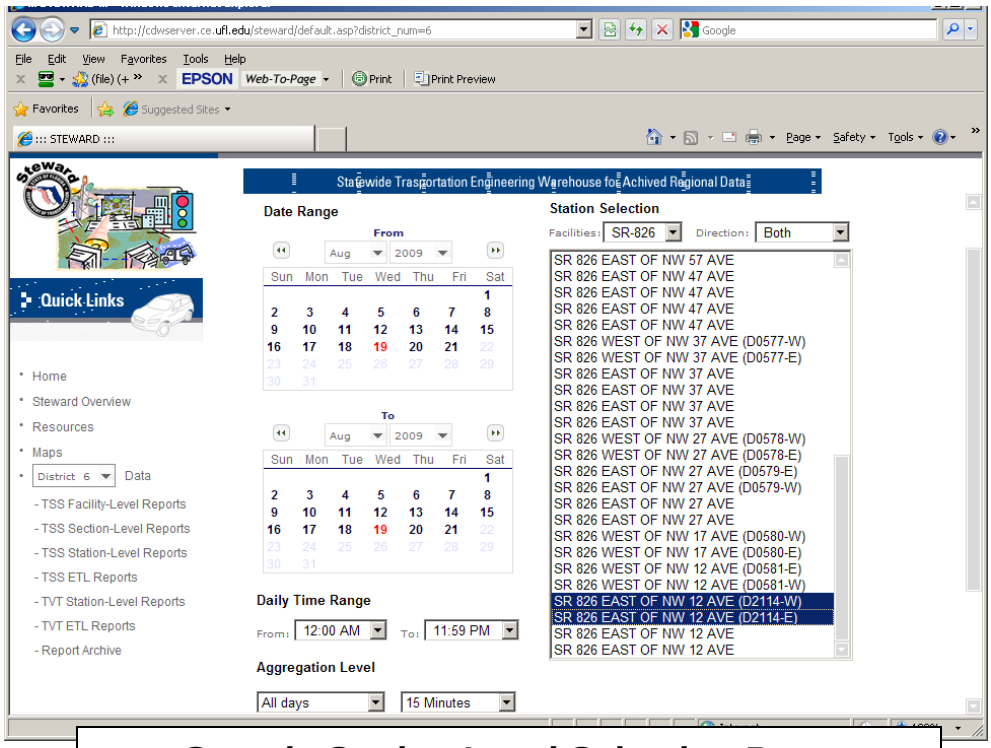
(Note: This action corresponds to Step 5 as shown in Figure 2).

When you click the button in this frame you will be signed on to the STEWARD web site. Station level and facility level downloads are handled differently. The frame at the left of the screen lets you choose what you want to do. Select the district and choose "Facility Level" or "Station Level" TSS files. A different screen will appear for each of these choices. A sample "Facility Level" screen is presented on the next page. A new set of boxes will appear to let you select the following items:

- The date of the count. When you select a date in the "From" calendar, the "To" calendar will be assigned the same date to give you a single day of data.
- The beginning and ending times
- The aggregation level. Normally you want 15 minutes for District Office counts and 1 hour for Central Office Counts.
- The station on the freeway from which you want the counts if you are dealing at the station level.



Sample Facility Level Reports



Sample Station Level Selection Page

The station level page has a slightly different appearance than the facility level page. The main difference is that you have to select the station that you want to deal with. To generate a bidirectional count, you must download the data from both directions at once. This is done by selecting two stations from the list using the ***Ctrl*** key with the mouse.

Click the "Traffic Counts" button. The count data will be sent as a CSV file to the STEWARD server and a list of available files will be displayed. If there is more than one file in the list, the file that you just created should be at the top with a creation time and date close to the current time and date. All files in this list are purged at midnight, so it is necessary for you to download each file to your own computer on the day that it was created.

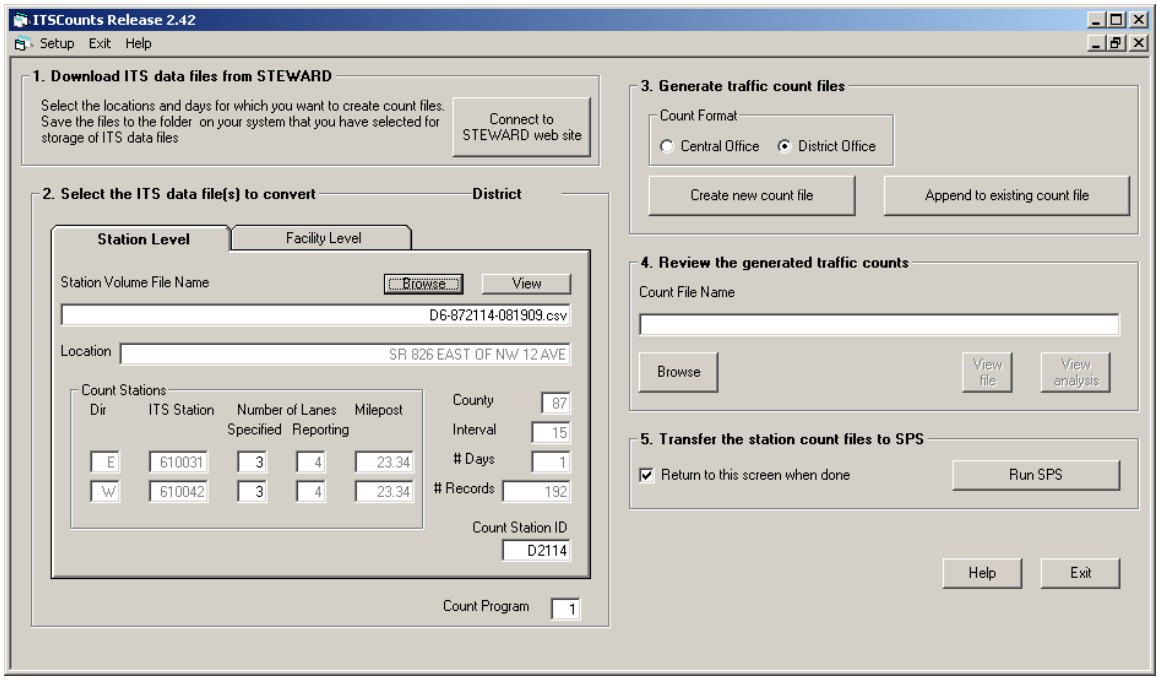
(Note: this action corresponds to Step 6 as shown in Figure 2) Download the file to your computer by clicking the "File Name" entry in the last column. A set of standard Windows dialog boxes will guide you through the process of saving the file. Be sure to save the file to the folder that you have specified in the ITSCounts program for STEWARD data. You probably want to give the file a name that you will be able to remember easily. When you have downloaded all of the files that you need, you should exit the STEWARD web site and return to the ITS program.

2. Select the ITS Data Files to Convert

(Note: This action corresponds to Step 7as shown in Figure 2).

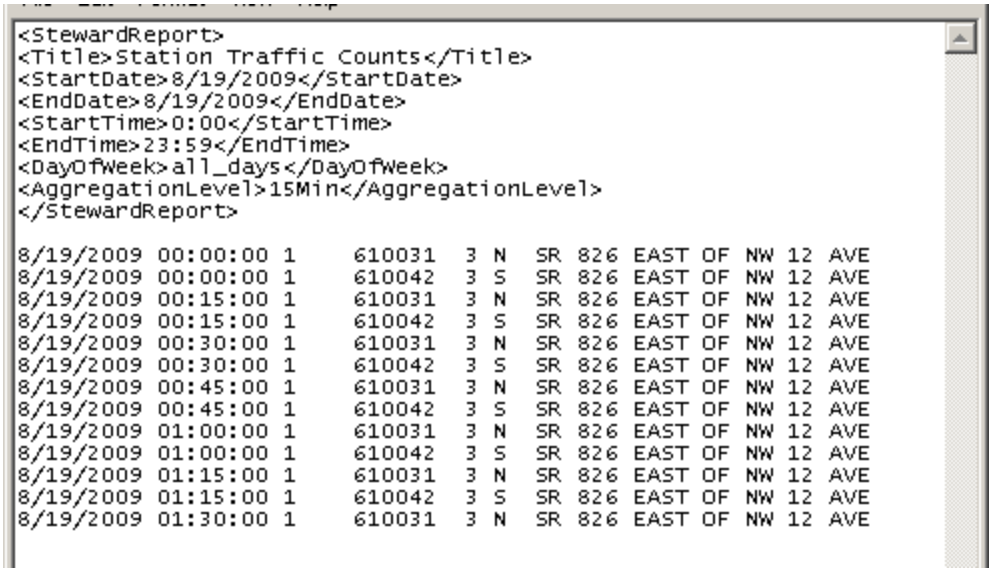
Station Level processing

The group box on the main screen for this step contains two tabs that let you choose between station level and facility level processing. We will look at the station level tab here. The screen is presented below. The ***Browse*** button lets you browse for the file. A typical screen display is shown below.



Note that the count station number shown on this screen was placed in the configuration database for this facility. If archive data are from an ITS station with no corresponding central or district office count station in the database, the number will revert to 9999. A different station number may be assigned here if you want to create a count file with that station number. You may still generate a count and review the results without a station number, but the file will not be recognized by SPS.

You can view the ITS Data file by clicking the ***View File*** button on this screen A partial view of a sample file is shown here. Note the XML header that is used by ITSCounts and other post processing utilities to extract the file properties.



```
<StewardReport>
<Title>Station Traffic Counts</Title>
<StartDate>8/19/2009</StartDate>
<EndDate>8/19/2009</EndDate>
<StartTime>0:00</StartTime>
<EndTime>23:59</EndTime>
<DayOfWeek>all_days</DayOfWeek>
<AggregationLevel>15Min</AggregationLevel>
</StewardReport>

8/19/2009 00:00:00 1    610031  3 N  SR 826 EAST OF NW 12 AVE
8/19/2009 00:00:00 1    610042  3 S  SR 826 EAST OF NW 12 AVE
8/19/2009 00:15:00 1    610031  3 N  SR 826 EAST OF NW 12 AVE
8/19/2009 00:15:00 1    610042  3 S  SR 826 EAST OF NW 12 AVE
8/19/2009 00:30:00 1    610031  3 N  SR 826 EAST OF NW 12 AVE
8/19/2009 00:30:00 1    610042  3 S  SR 826 EAST OF NW 12 AVE
8/19/2009 00:45:00 1    610031  3 N  SR 826 EAST OF NW 12 AVE
8/19/2009 00:45:00 1    610042  3 S  SR 826 EAST OF NW 12 AVE
8/19/2009 01:00:00 1    610031  3 N  SR 826 EAST OF NW 12 AVE
8/19/2009 01:00:00 1    610042  3 S  SR 826 EAST OF NW 12 AVE
8/19/2009 01:15:00 1    610031  3 N  SR 826 EAST OF NW 12 AVE
8/19/2009 01:15:00 1    610042  3 S  SR 826 EAST OF NW 12 AVE
8/19/2009 01:30:00 1    610031  3 N  SR 826 EAST OF NW 12 AVE
```

If you are processing station level data, you should now skip over the remaining discussion on facility level processing and proceed to the next step in which the traffic counts are generated.

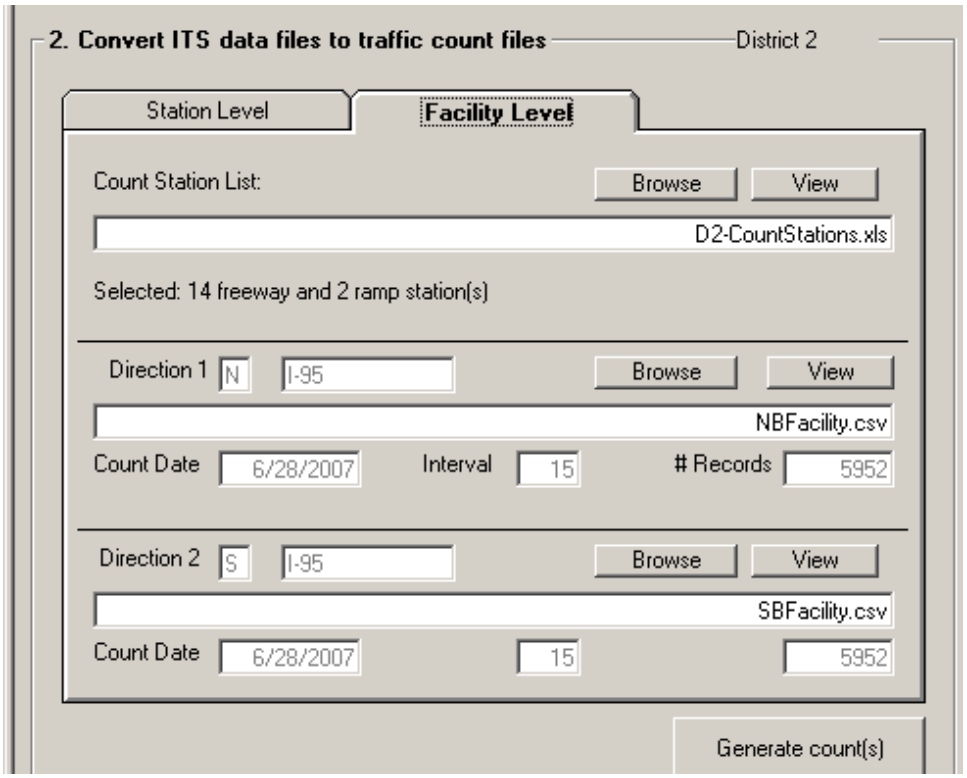
Facility Level Processing

Two sample ITS data files containing facility level District 2 count data were copied to the *ITSdata* folder to illustrate the working of ITSCounts at the facility level:

- NBFacility.csv
- SBFacility.csv

Both of these files were downloaded from the STEWARD web site as described in the previous section of this document. The file names were chosen for easy recognition. These files cover both directions on Interstate 95 between the I-295 north and south interchanges in Jacksonville. Approximately 125 ITS stations are represented in each file. Because of the manner in which STEWARD collects and stores data, separate count files must be downloaded for each direction.

The screen below shows the contents of the facility level tab, using a different example from District 2.



Note that three files must be selected for this operation using their corresponding *Browse* buttons.

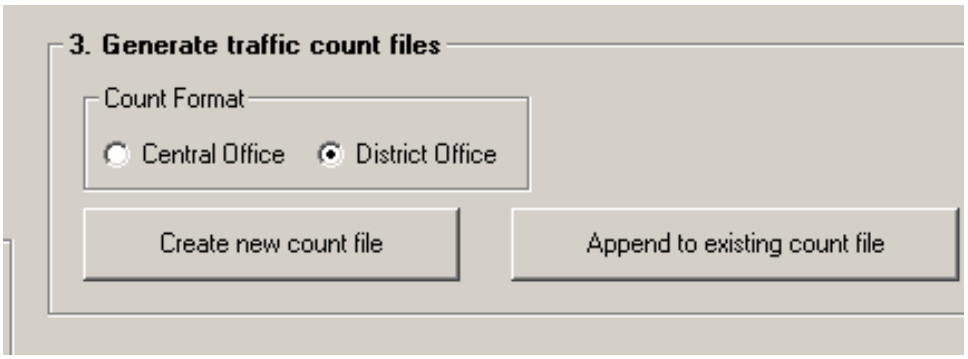
1. The count station list, which was provided by the Statistics Office
2. The downloaded file representing Direction 1 (increasing mileposts)
3. The downloaded file representing Direction 2 (decreasing mileposts)

The count station list (described previously) contains information about the ITS station attributes and identifies the corresponding district count station number. It also lets you decide whether or not you want to generate a count for a particular station. This file is an Excel spreadsheet that can be edited to vary your choice of count files that you want to create. When you click the ***Generate count(s)*** button a bidirectional count file in the format required by SPS will be created for each station that was requested in the count station list. You can edit the selected count station list by clicking its *View* button. Don't forget to save the file when it has been edited. If you save the file under a different name, you must reselect the count station list again to load the new file.

3. Generate the Traffic Count Files (Station Level Only)

(Note: This action corresponds to Step8 as shown in Figure 2 if you are using facility level processing you should skip this step because the files will have been created in the previous step))

This is where the traffic count files are generated. The box looks like this:



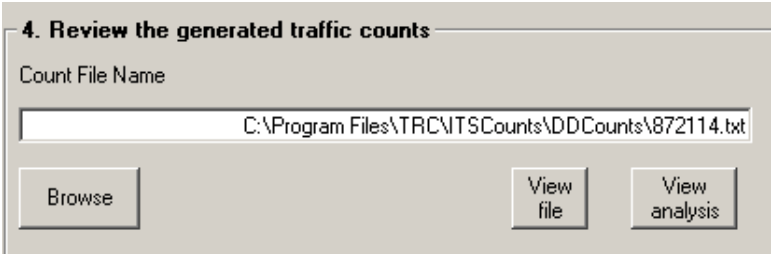
There are two types of count files that you can generate:

1. Central Office counts
2. Directional District Office counts

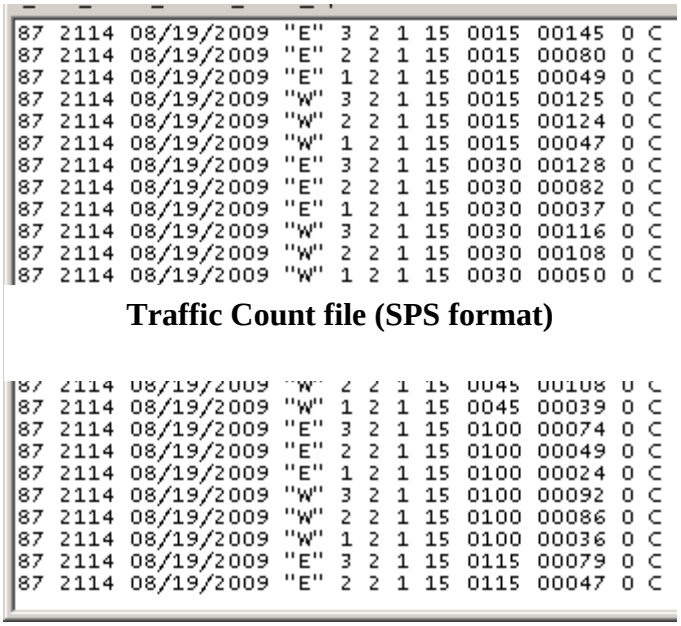
You can now create a new file or append the data to an existing file. Normally you will want to create a new file unless you want to add another day of data to an existing file.

4. Review the Generated Traffic Counts

When a station level traffic count file is generated, its name will appear in the file name box for this frame. If you want to examine a different file, you can use the ***Browse*** button. If you are doing facility level processing, you will probably have created several files and you will have to browse for each file that you want to review. The box looks like this:



You can view the traffic count (text) file with the ***View File*** button. A sample traffic count file is shown on the next page.



```
87 2114 08/19/2009 "E" 3 2 1 15 0015 00145 0 C
87 2114 08/19/2009 "E" 2 2 1 15 0015 00080 0 C
87 2114 08/19/2009 "E" 1 2 1 15 0015 00049 0 C
87 2114 08/19/2009 "W" 3 2 1 15 0015 00125 0 C
87 2114 08/19/2009 "W" 2 2 1 15 0015 00124 0 C
87 2114 08/19/2009 "W" 1 2 1 15 0015 00047 0 C
87 2114 08/19/2009 "E" 3 2 1 15 0030 00128 0 C
87 2114 08/19/2009 "E" 2 2 1 15 0030 00082 0 C
87 2114 08/19/2009 "E" 1 2 1 15 0030 00037 0 C
87 2114 08/19/2009 "W" 3 2 1 15 0030 00116 0 C
87 2114 08/19/2009 "W" 2 2 1 15 0030 00108 0 C
87 2114 08/19/2009 "W" 1 2 1 15 0030 00050 0 C
```

Traffic Count file (SPS format)

```
87 2114 08/19/2009 "W" 2 2 1 15 0045 00108 0 C
87 2114 08/19/2009 "W" 1 2 1 15 0045 00039 0 C
87 2114 08/19/2009 "E" 3 2 1 15 0100 00074 0 C
87 2114 08/19/2009 "E" 2 2 1 15 0100 00049 0 C
87 2114 08/19/2009 "E" 1 2 1 15 0100 00024 0 C
87 2114 08/19/2009 "W" 3 2 1 15 0100 00092 0 C
87 2114 08/19/2009 "W" 2 2 1 15 0100 00086 0 C
87 2114 08/19/2009 "W" 1 2 1 15 0100 00036 0 C
87 2114 08/19/2009 "E" 3 2 1 15 0115 00079 0 C
87 2114 08/19/2009 "E" 2 2 1 15 0115 00047 0 C
```

You can also look at the analysis file by clicking the ***View Analysis*** button. This file is a separate CSV file generated at the same time as the count file. It contains information by time period that can be imported into an Excel spreadsheet for analysis and plotting. Some sample plots are shown on the next couple of pages.

When you browse for a District Office file, only the count file (.TXT) version will be displayed. You can view either the count file or analysis file because the file names are the same except for the extension. Browsing for a Central Office file is more complicated because the Central Office file naming scheme does not follow standard Windows conventions. In this case you must choose either the count file (ending in _CNT) or the analysis file (ending in .CSV). From that point, you will only be able to view the file that you have selected.

Note: Be sure to close all spreadsheet files that you have viewed before generating any new counts. You will get a "permission denied" error if you try to create a spreadsheet that is already open.

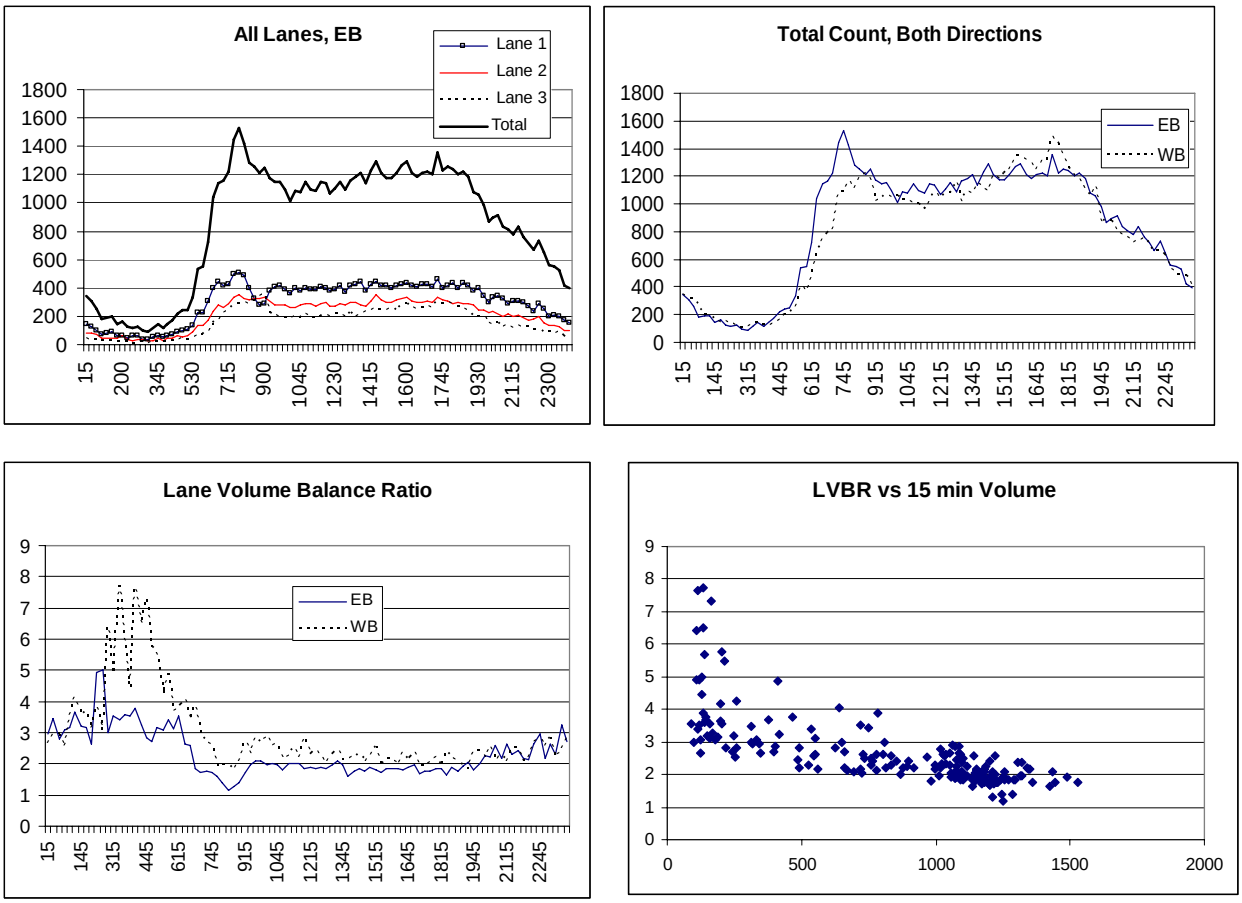
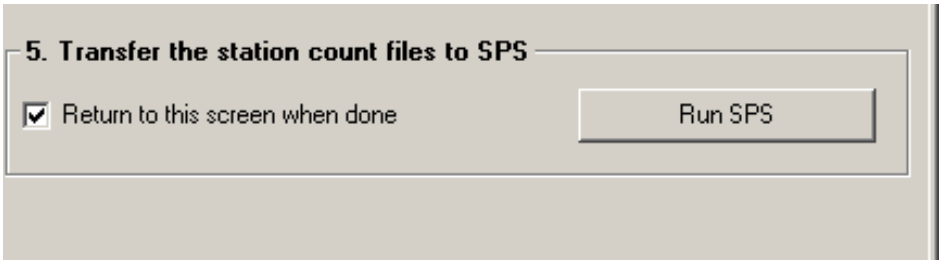


Figure 5. Sample plots from the analysis file

5. Transfer the Counts to SPS (District Counts Only)

To facilitate the assimilation of traffic counts into the SPS database a button is provided in this frame to run the SPS program directly. If the file locations specified in the SPS program correspond to those specified in ITSCounts, the process should go smoothly. Otherwise you might have to do some more searching in SPS. Note that you have the option to return to the ITSCounts program by checking the box in this frame. The details of the SPS operation are beyond the scope of this document.



Changing Program and Folder locations

The ITSCounts installation routine establishes default locations for all files and programs. If you need to specify different locations, you may do so by clicking ***Setup*** on the main menu bar. The *Setup* screen will then be displayed.

You can change all entries except the STEWARD web site by browsing for their locations. Note that files and folders must have already been created before they can be selected. The STEWARD web site is a fixed URL that will not change in the near term.

File and Folder Maintenance

You should be aware that of the files that you create will accumulate in their respective folders. It is a good idea to periodically delete files that are no longer needed, or at least move them to an archive folder. When you delete files from the *Working* folder, you should take care not to delete the ***ITSCounts.par*** file because it is needed by the program. There is no need to delete files from the program folder.

Assessing Count Data Files before Downloading

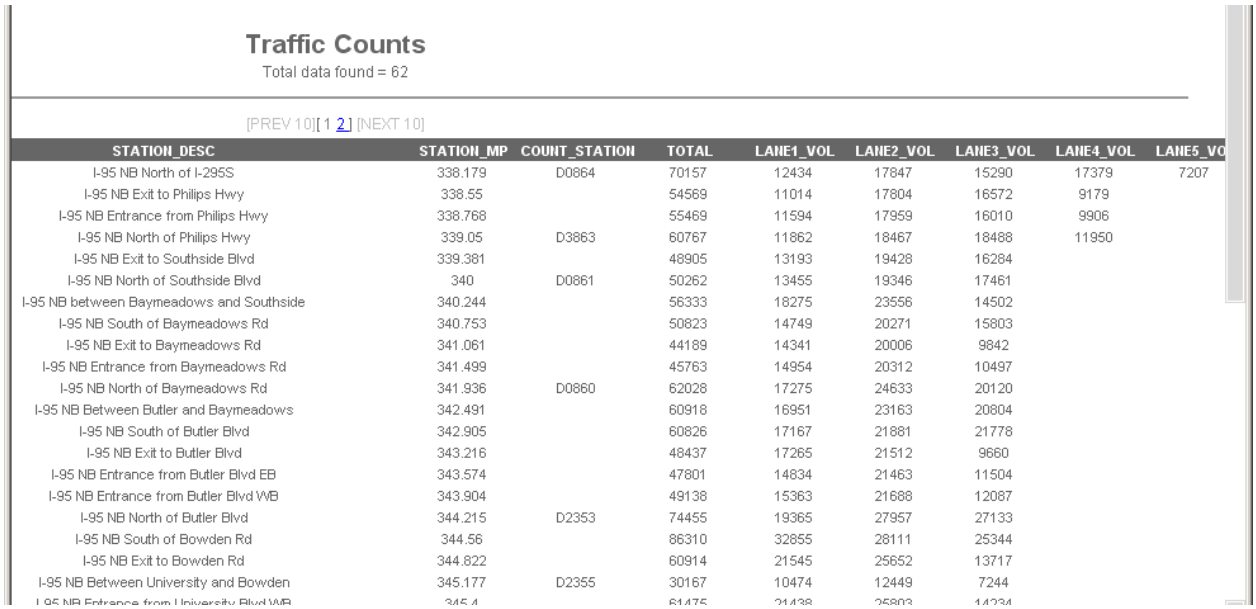
A separate utility program called **SunVol** has been developed to support the selection of "good" days for traffic count extraction. The SunVol program files and documentation are available form the STEWARD website. You can also perform a quick assessment by looking at the ETL reports that are available on the web site.

The SunGuide TSS facilities are still under development at some regional TMCs. At this stage of development some data quality issues have been observed. The main problem seems to be missing counts for either a part of a day or for the full day. There are a couple of things that you can do to scrutinize the count summary data to help you decide which files are likely to produce valid traffic counts for your purposes. When you connect to the STEWARD site, you should choose the "TSS ETL Reports" link from the Quick Links list in the left frame. The first item in the list is the "TSS Conversion History," which presents a summary of the critical attributes and potential problems for each daily archive. The items of most interest to traffic count assessment include:

- Elapsed Minutes: If this figure is less than 1440, then the system was inoperative for a portion of the day.
- Null Minutes: This figure indicates the number of minutes during which no counts were received from any station. It should be zero, or at least very low.
- Null Stations: This figure indicates the number of stations that did not report any data for the day.
- Null Lanes: This figure indicates the number of lanes at active stations that did not report any data for the day. The lanes at null stations are not included in this figure.

You should use the *Conversion History Report* for pre-screening to determine which days are good candidates to examine further. You can get more detailed information by opening the *Daily Report Log* for the days that you have selected. The *Daily Report Log* expands the information in the *Conversion History Report* to tell you which of the locations had null stations or lanes.

The most detailed screening information is found in the facility level traffic count report mentioned earlier. This report is normally created with 15 minute intervals to produce district traffic counts. On the other hand, if you specify the interval as "1 Day" instead of 15 minutes, you will get a very good overview of what happened that day for each station. As the table below shows, the count stations are identified, along with the total volume and the lane-by-lane volumes. So you can check to ensure that the total volume is reasonable and that no lanes were missed.



Traffic Counts

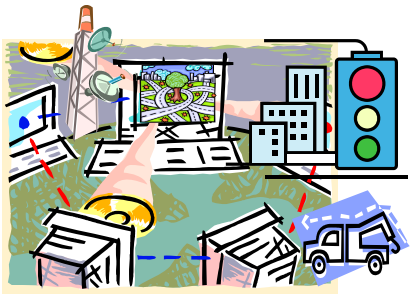
Total data found = 62

[PREV 10][1 2] [NEXT 10]

STATION_DESC	STATION_MP	COUNT_STATION	TOTAL	LANE1_VOL	LANE2_VOL	LANE3_VOL	LANE4_VOL	LANE5_VO
I-95 NB North of I-295S	338.179	D0864	70157	12434	17847	15290	17379	7207
I-95 NB Exit to Philips Hwy	338.55		54569	11014	17804	16572	9179	
I-95 NB Entrance from Philips Hwy	338.768		55469	11594	17959	16010	9906	
I-95 NB North of Philips Hwy	339.05	D3863	60767	11862	18467	18488	11950	
I-95 NB Exit to Southside Blvd	339.381		48905	13193	19428	16284		
I-95 NB North of Southside Blvd	340	D0861	50262	13455	19346	17461		
I-95 NB between Baymeadows and Southside	340.244		56333	18275	23556	14502		
I-95 NB South of Baymeadows Rd	340.753		50823	14749	20271	15803		
I-95 NB Exit to Baymeadows Rd	341.061		44189	14341	20006	9842		
I-95 NB Entrance from Baymeadows Rd	341.499		45763	14954	20312	10497		
I-95 NB North of Baymeadows Rd	341.936	D0860	62028	17275	24633	20120		
I-95 NB Between Butler and Baymeadows	342.491		60918	16951	23163	20804		
I-95 NB South of Butler Blvd	342.905		60826	17167	21881	21778		
I-95 NB Exit to Butler Blvd	343.216		48437	17265	21512	9660		
I-95 NB Entrance from Butler Blvd EB	343.574		47801	14834	21463	11504		
I-95 NB Entrance from Butler Blvd WB	343.904		49138	15363	21688	12087		
I-95 NB North of Butler Blvd	344.215	D2353	74455	19365	27957	27133		
I-95 NB South of Bowden Rd	344.56		86310	32855	28111	25344		
I-95 NB Exit to Bowden Rd	344.822		60914	21545	25652	13717		
I-95 NB Between University and Bowden	345.177	D2355	30167	10474	12449	7244		
I-95 NB Entrance from University Blvd WB	345.4		61475	21438	25803	14234		

Note that the facility counts are directional, so you need to repeat this process for both directions to get an idea of the operation for the whole system.

Appendix 4c:
SunVol Analysis Utility



Statewide Transportation Engineering Warehouse for Archived Regional Data (STEWARD)

Florida Department of Transportation
University of Florida
Transportation Research Center

Analysis of Traffic Volume Counts from SunGuide Archives

Release 2.0: November 15, 2009

This document covers the analysis of SunGuide traffic counts to examine the variability of data from day to day and to identify questionable days. A desktop utility program called SunVol has been developed for this purpose. The actual count extraction is performed by another utility called ITSCounts. The documentation and installation files for SunVol and ITSCounts may be downloaded from the STEWARD website:

http://cdwserver.ce.ufl.edu/steward/index.html

The SunGuide traffic sensor system (TSS) Archive data to be examined by SunVol may be downloaded from the same web site. A variety of reports may be accessed from this site as described in other project documentation, which is also available from the website.

Installing the SunVol Program

SunVol is installed on a local computer using the normal Windows installation process. Three sub-folders are created within the program folder during the installation:

- **ITSData**, which contains the comma delimited (CSV) files produced from the STEWARD database query. Please refer to the STEWARD website user guide, which is available on the website.
- **Outputs,** which contains the analysis results, also in the form of CSV files
- **Working,** which contains a parameter file that gives the locations of programs, folders and URLs required by the program.

The installer must have read and write access to the program folder. After the program has been installed and set up, the operator must have read and write access permission for all three of the above folders.

The overall file structure for SunVol is shown in Figure 1. The installation routines will establish the default folder and program locations. If a different configuration is desired, the locations may be changed from within the program, keeping in mind that the operator must have read and write access to the working folder and all of the data folders.

Note: You will find that the installation will be much easier if you choose all of the default installation options. The discussion and examples in this document assume that the default options have been chosen.

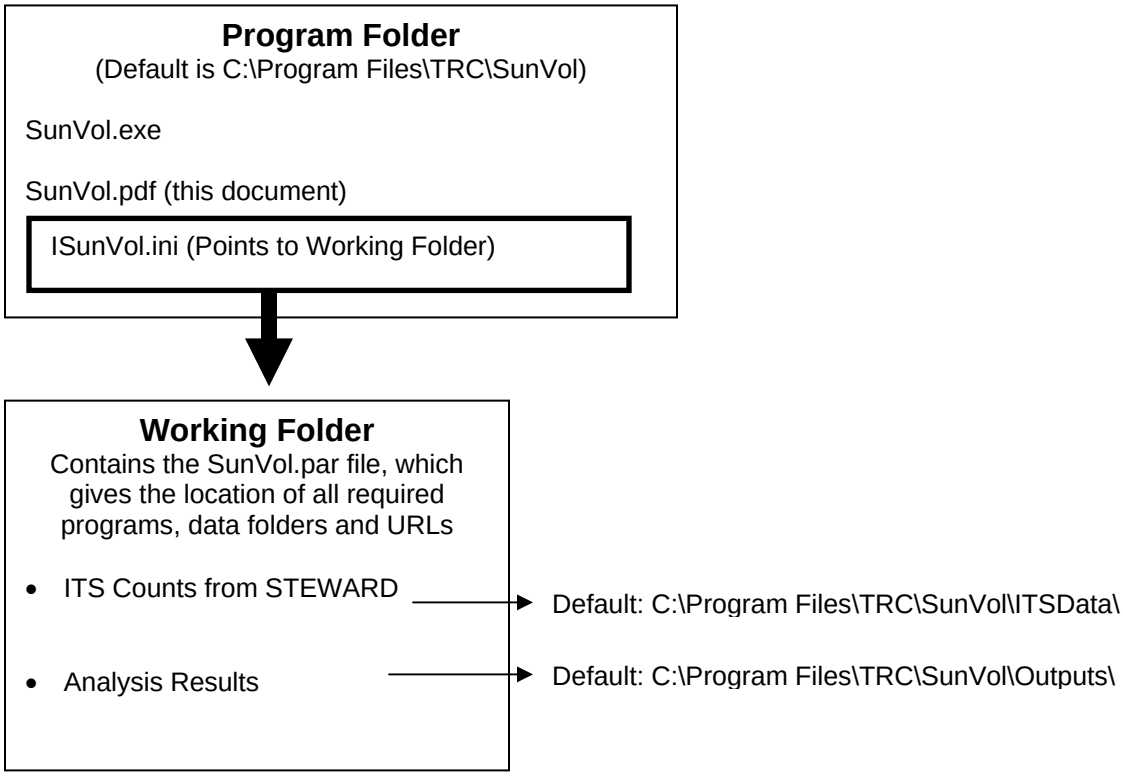
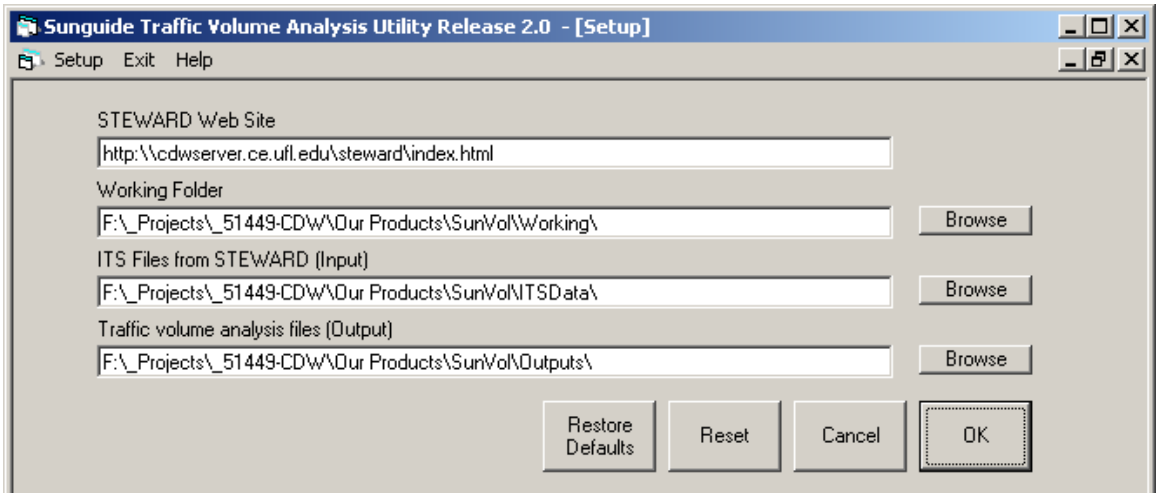


Figure 1: SunVol folder structure

First Time Setup for File locations

When you run the program for the first time, you will get a message telling you that the file configuration needs to be set up. Click ***OK*** and you will see the following screen:

The default file locations will be shown when you first enter this screen. It is recommended that you accept these defaults. The required folders will already have been set up when the program is installed. Advanced users may wish to specify their own file locations.



Running SunVol

The following discussion assumes that you are familiar with basic Windows operations and that you have set up the program using all of the default program and file locations. When you run ITSCounts, the main screen will appear as shown on the next page Note that there are three frames visible at this point, each of which is associated with one step in the file selection and analysis process:

- Download ITS data files from STEWARD
- Select and analyze the ITS data files
- Review the results from a previous analysis

You will first want to download the ITS data to be processed. This is Step 1 on the screen shown below in Figure 1. Click the ***Connect to STEWARD web site*** button.

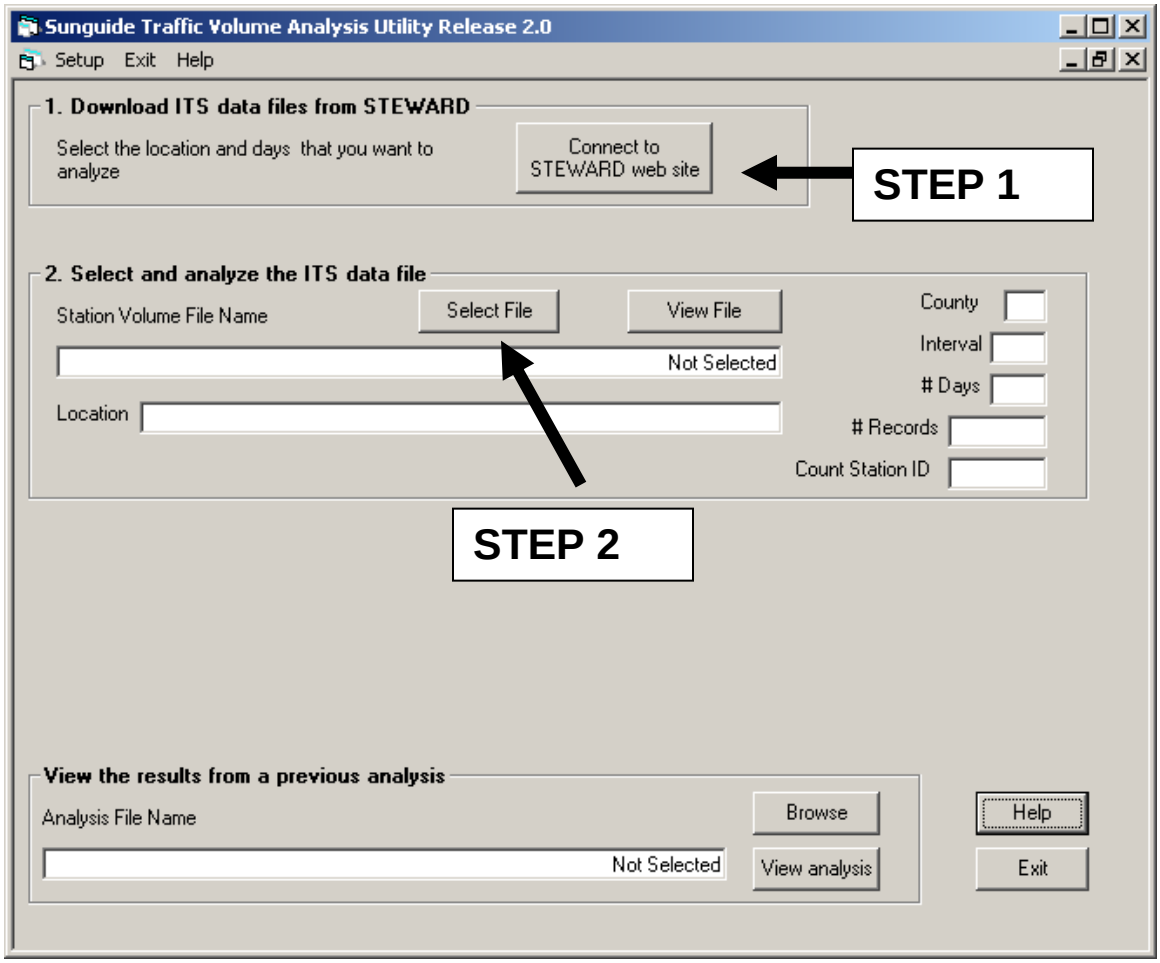


Figure 1. SunVol Main Screen

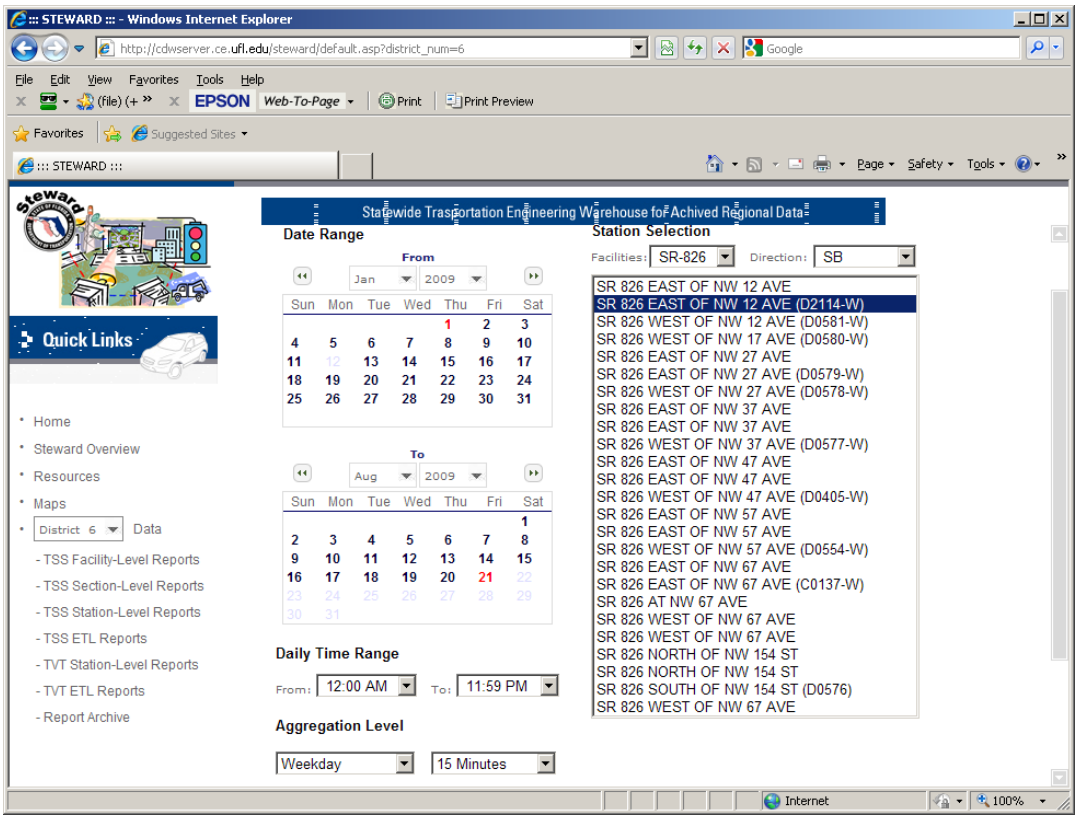
Downloading ITS Data files

After connecting to the STEWARD web site, you must first choose the FDOT district from which you want to obtain data. Other selection criteria include

- Report Level: You want to choose "Station Level Reports"
- Date Range: You may include up to 400 days in the date range
- Time range: The full 24 hour period must be specified. SunVol will give you an error message if you specify a smaller time range.
- Day of Week: There is a variety of choices available (all days, week days, specific days, etc)
- Aggregation Level: SunVol only supports aggregation levels of 15 or 60 minutes.
- ITS Count Station: You must choose a single station to analyze

The examples in this document were developed using "weekday" data from a station on SR836 in District 6. The date range for 2009 that was available at the time ran from Jan 1 to Aug 21.

The STEWARD website selection screen looks like this after the selection criteria have been specified:



When you have made your selections, click the *Traffic Counts* button and follow the instructions to download a CSV file. You should choose a file name that you will recognize later, because you will have to select that file for analysis. Don't forget to keep the CSV extension. The file should be saved in the folder that has been specified for SunVol ITS data. When you exit from the website, you will return to the SunVol main screen.

Selecting and Analyzing the File

Note: This is Step 2 as shown in Figure 1.

The files selection process is carried out using standard Windows dialog boxes with which you should already be familiar. Clicking the select button invokes the SunVol processing on the input station file and produces the output SunVol output (CSV) file in the specified output directory. When you select a file, it will automatically be analyzed and the results will be displayed in Microsoft Excel Window. A spreadsheet covering all of the days will open automatically when your selection has been made. Note: You must be using a Windows-based computer that has Microsoft Excel Version 2003 or later installed. The Excel spreadsheet has two sections, which will be examined individually.

Daily Volumes by Time interval

A partial sample of this section is shown here. The columns present the time interval and the average volume over all of the selected days for that interval (eliminating days with zero volumes). The upper and lower 95% confidence bounds are also displayed, based on the computed standard deviation. The number of periods with zero volume is indicated.

SunGuide Archive Volume Analysis
SR 826 EAST OF NW 67 AVE
Time Range: From 0:00 To 23:59 by 60 Minute Intervals
Date Range: From 09/15/2008 To 10/30/2008: (11) Days
ITS Station ID: 610351 Count Station ID: C0137-E
Mile post: 17.76

DAILY VOLUMES BY TIME INTERVAL

Time	-95%	AvVol	95%	NoCount	StdDev
1:00:00	276	934	1592	0	329
2:00:00	68	526	983	0	229
3:00:00	47	391	736	0	172
4:00:00	200	418	635	0	109
5:00:00	463	606	749	0	71
6:00:00	702	1464	2225	0	381
7:00:00	1293	3490	5686	0	1098

Daily Volume Variation

A sample of the daily volume variation section is shown here.

DAILY VOLUME VARIATION

Date	24HRVol	MaxVol	NoCount
9/15/2008	69190	5083	0
9/16/2008	71785	5633	0
9/17/2008	73301	5237	0
9/18/2008	72956	5356	0
9/19/2008	75557	5209	0

Each day in the range is represented by one row in the table. The columns indicate the 24 hour volume, the maximum count for any interval (either 15 or 60 minutes) and the number of intervals in each day with zero volumes.

A summary description of the data set that you downloaded will be displayed in Box #2 in Figure 2. Note that there are 161 days of data in the file. Note also that a new box has appeared to let you select a single day for more detailed analysis. Before you select a day, you probably want to examine the data for the entire period. You can do this best by plotting some results from the spreadsheet.

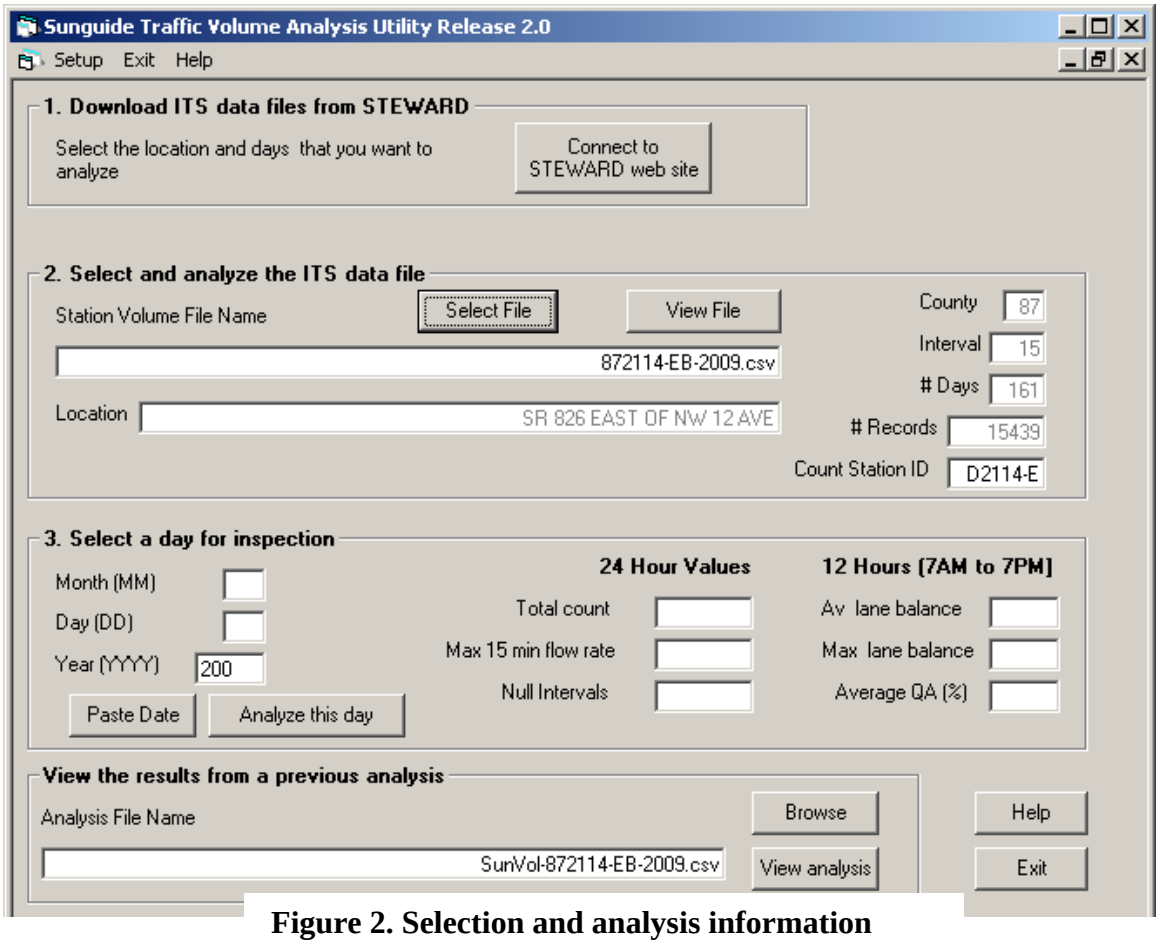


Figure 2. Selection and analysis information

Two plots are shown in Figure 3:

1. Variation of the traffic count data by 15 min period for all days. The average volume and 95% confidence limits are shown.
2. Variation of the 24 Hour counts throughout the analysis period. Since the sample involves only weekday data, the gaps on this plot represent the weekends.

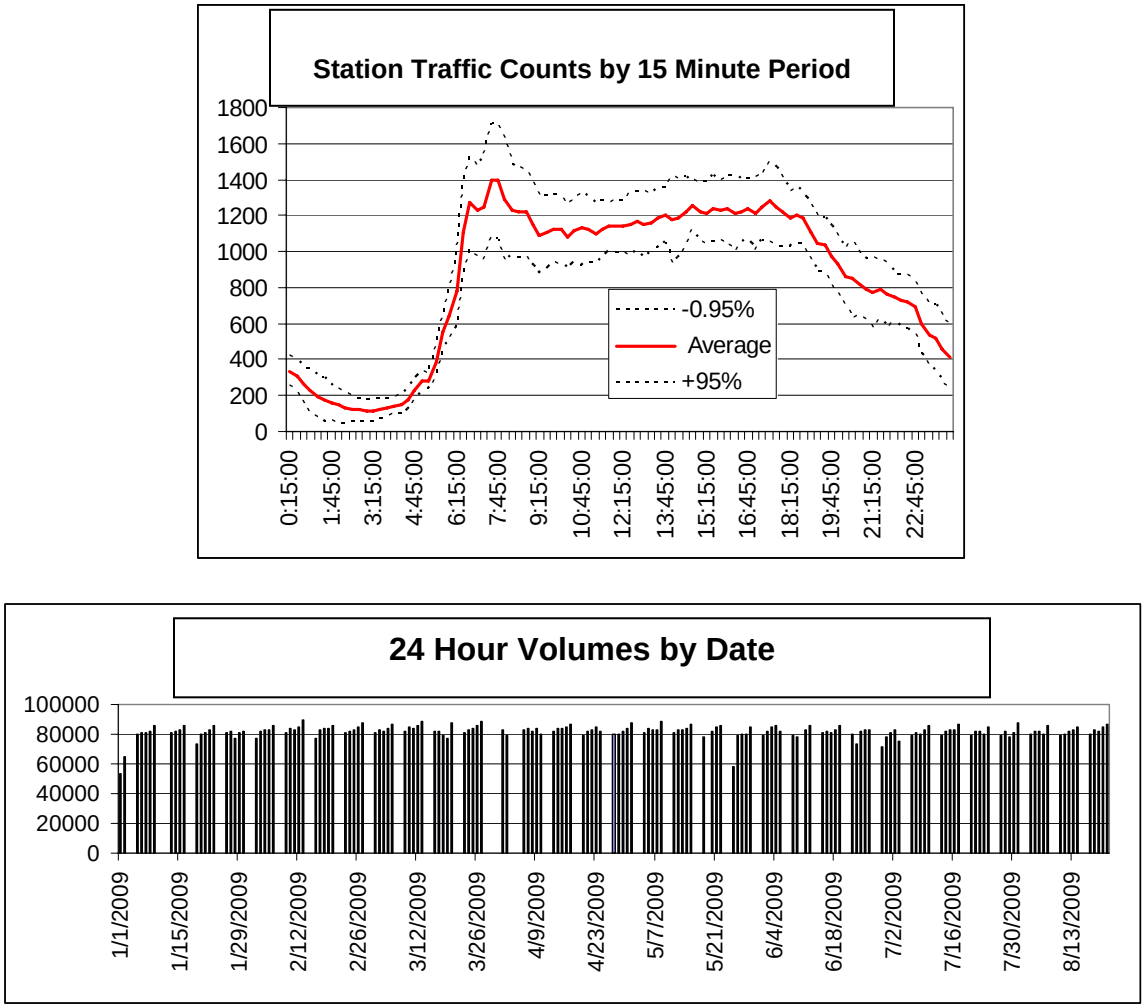


Figure 3. Sample plots from the SunVol Analysis file

The figures plotted here represent eastbound traffic. The westbound plots have a similar appearance.

SunVol was applied to all of the ITS stations in Districts 2, 4, and 6. The graphics shown in Figure 3 were plotted for each station. The results have been posted on the STEWARD website.

Selecting a Day for Count Extraction

It is now necessary to select a single day for creating a traffic count file. Referring to Figure 4, there are two ways to do this:

1. You can enter the desired date in the month, day and year fields in the specified format.
2. You can click the ***"Paste Date"*** button and a dialog box will pop up to let you copy a date field from the spreadsheet and paste it in the box. When you have finished pasting the date, click ***OK*** in the pop up box.

When you have selected a date by either method, click the ***Analyze this day*** box. A spreadsheet will appear with the results and a summary of the analysis will be displayed. The selected day for this analysis is 8-19-09

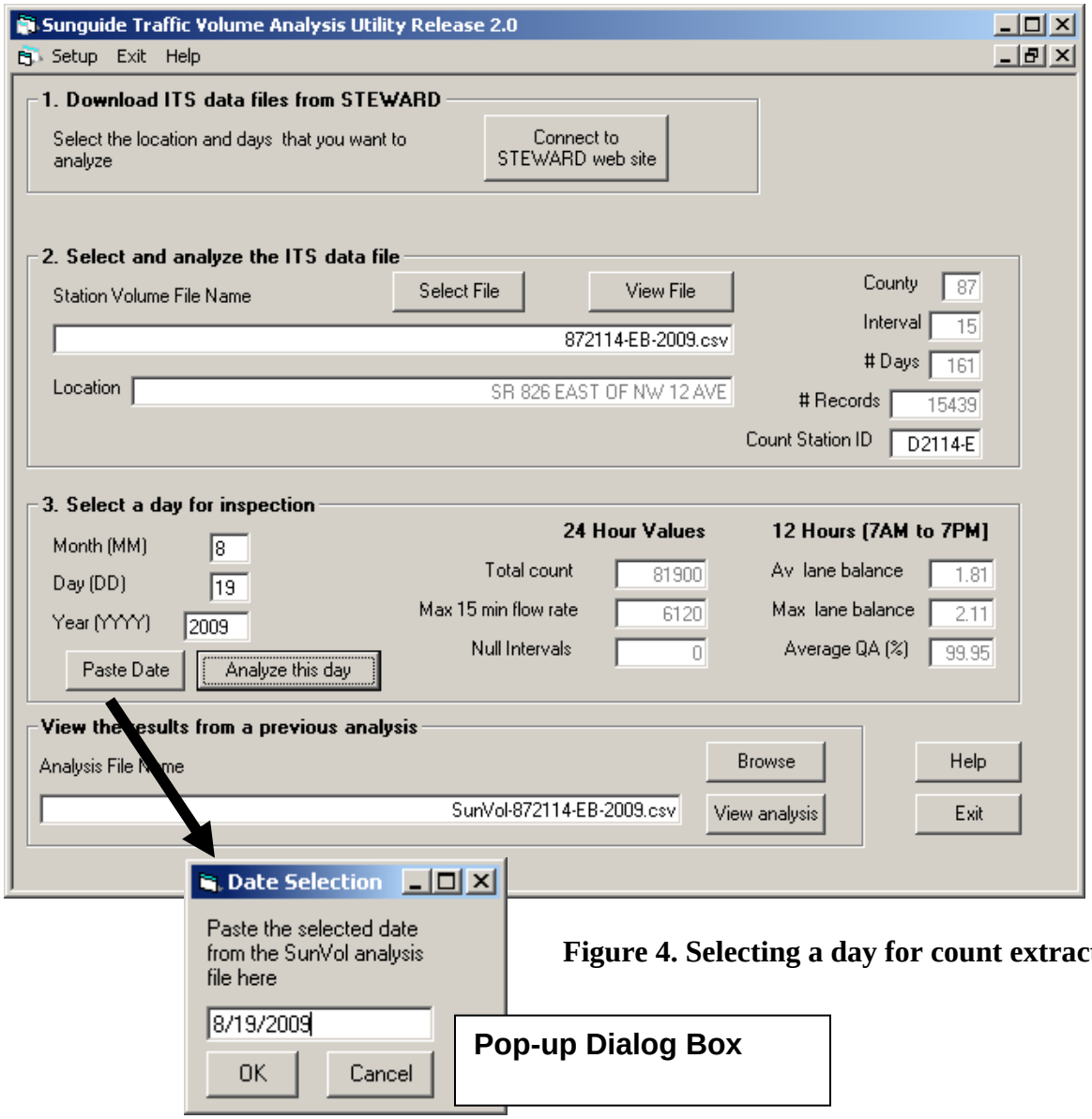


Figure 4. Selecting a day for count extraction

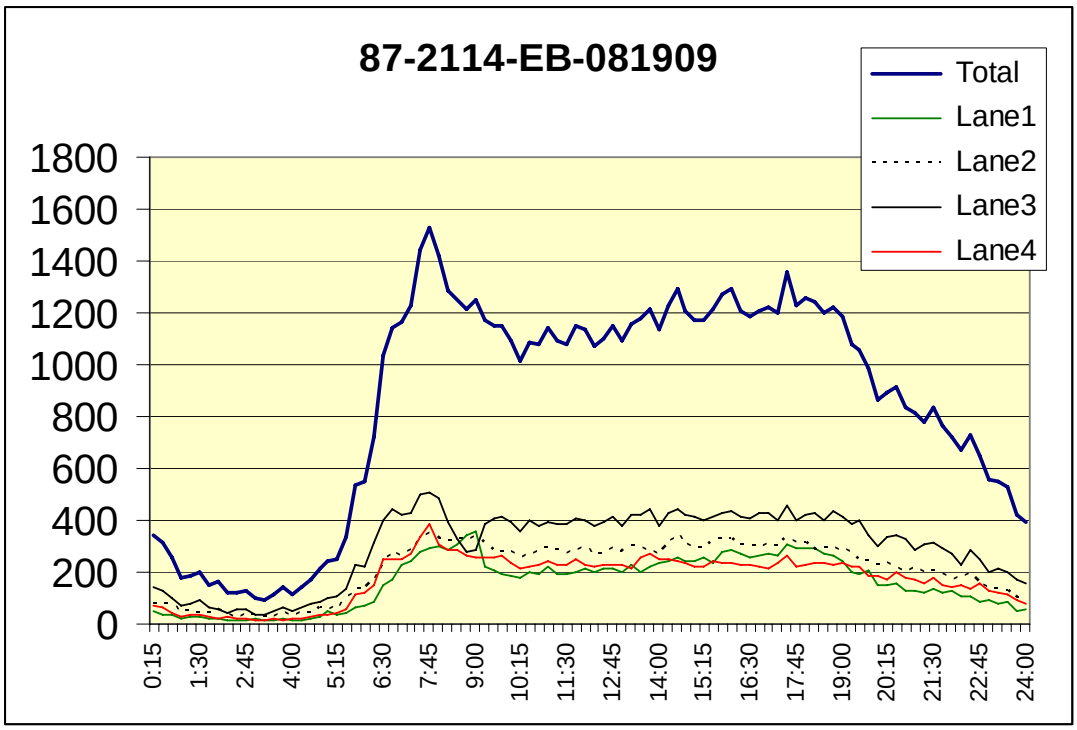
The characteristics that might indicate a problem with the data are displayed in the summary of the analysis in Figure 4. Some fields give 12 hour values because they represent data that can behave erratically during periods of low volume.

- The average lane balance is important because the lane balance indicates the volume ratio of the highest volume lane to the lowest volume lane. An even distribution of traffic among lanes would give a lane balance of 1.0. Typical average values are in the range of 1.5 to 2.0.
- The maximum lane balance indicates the highest 15 minute period. Values much greater than 5.0 suggest that an incident might have disrupted traffic to the point where a different day should be considered for count extraction.
- The quality assurance (QA) is expressed as the percent of polling intervals (usually 20 seconds) that reported valid data. This value should be close to 100%.

The 24 hour values are also of interest:

- The 24 hour volume should be within the expected range based on historical data.
- The maximum 15 minute flow rate should not exceed 2900 veh/hr/lane (i.e., 120% of the 2400 veh/hr/lane capacity suggested in the Highway Capacity Manual).
- The number of null intervals (in this case 15 minute periods with zero volume) should be close to zero. Null intervals generally suggest missing traffic counts.

Results may also be plotted from the daily analysis spreadsheet. An example of a plot showing the lane volumes and total volume over the day by 15 minute period is shown below.



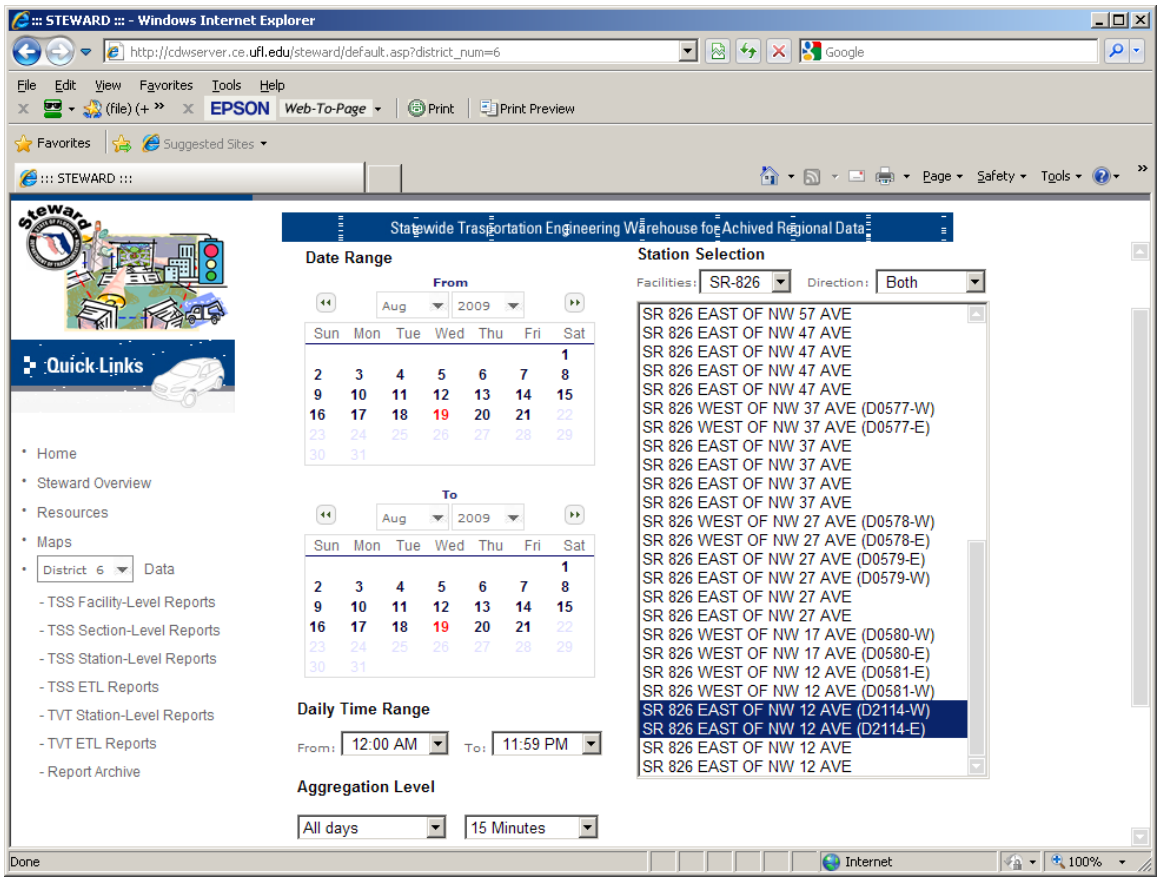
So the selected date of 8-19-09 would seem to be suitable for traffic count extraction.

Traffic Count Extraction

Extraction of traffic counts and conversion to the FDOT counting program formats is done with the ITSCounts utility program described in a separate document that is available on the STEWARD website. For the sake of continuity the example here will be carried through the count extraction stage, assuming that you have a basic familiarity with ITSCounts.

Downloading the ITS Data

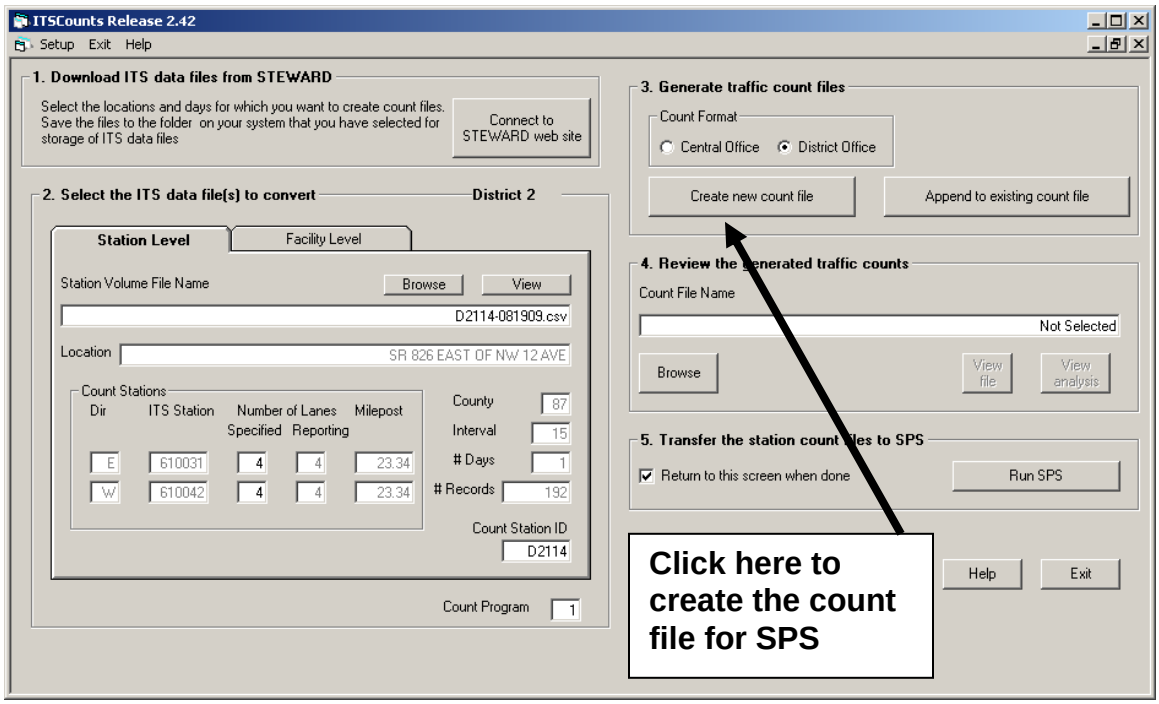
It is necessary now to go back to the STEWARD web site to download the data d for both directions from the selected site for the selected day. To do this, you should first select the facility and request a display of stations for *Both* directions. The screen will appear as shown below.



Data from both directions may be requested using the ***Ctrl*** function with the mouse to highlight multiple stations. The date selection must then be changed to specify a single date; in this case August 19, 2009. The traffic count data for those stations on that day may now be downloaded and saved to the folder that you have specified for ITS data to be processed by ITSCounts.

Generating the Counts

When you have downloaded the data and returned to ITSCounts, you may ***Browse*** for the data file that you have saved. When you select this file, a summary of its important data characteristics will be displayed as shown below.



When you click the ***Create New Count File*** button the count file will be generated in the proper format for further processing by FDOT software.

Reviewing the Counts

When the count is generated, the file name will appear in the Step 4 ***Review the generated traffic counts*** frame. At that point you can choose to look at the count file itself or the analysis spreadsheet. A sample of the count text file itself and a plot showing the hourly count profile in both directions are presented on the next page.

Transferring the Counts to SPS

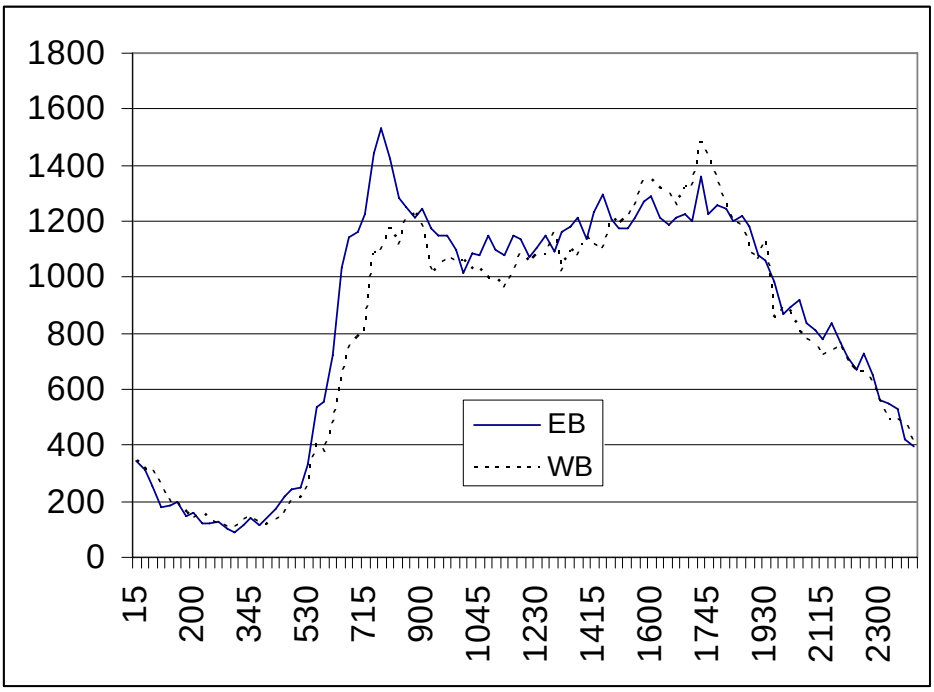
The transfer of counts to the FDOT Survey Processing Software (SPS) is covered in the ITSCounts program documentation. ITSCounts simply provides a link to launch the SPS program. To use SPS, you must be familiar with the operation of the program and have proper access to the statewide traffic count database on the FDOT mainframe computer.

872114.txt - Notepad

File Edit Format View Help

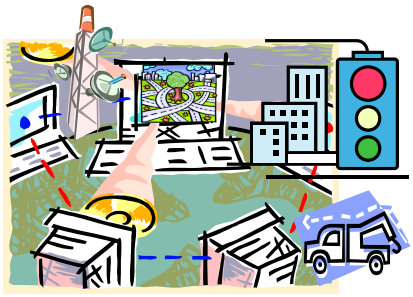
```
87 2114 08/19/2009 "E" 3 2 1 15 0015 00145 0 C
87 2114 08/19/2009 "E" 2 2 1 15 0015 00080 0 C
87 2114 08/19/2009 "E" 1 2 1 15 0015 00049 0 C
87 2114 08/19/2009 "W" 3 2 1 15 0015 00125 0 C
87 2114 08/19/2009 "W" 2 2 1 15 0015 00124 0 C
87 2114 08/19/2009 "W" 1 2 1 15 0015 00047 0 C
87 2114 08/19/2009 "E" 3 2 1 15 0030 00128 0 C
87 2114 08/19/2009 "E" 2 2 1 15 0030 00082 0 C
87 2114 08/19/2009 "E" 1 2 1 15 0030 00037 0 C
87 2114 08/19/2009 "W" 3 2 1 15 0030 00116 0 C
87 2114 08/19/2009 "W" 2 2 1 15 0030 00108 0 C
87 2114 08/19/2009 "W" 1 2 1 15 0030 00050 0 C
87 2114 08/19/2009 "E" 3 2 1 15 0045 00101 0 C
87 2114 08/19/2009 "E" 2 2 1 15 0045 00075 0 C
87 2114 08/19/2009 "E" 1 2 1 15 0045 00036 0 C
87 2114 08/19/2009 "W" 3 2 1 15 0045 00115 0 C
87 2114 08/19/2009 "W" 2 2 1 15 0045 00108 0 C
87 2114 08/19/2009 "W" 1 2 1 15 0045 00039 0 C
87 2114 08/19/2009 "E" 3 2 1 15 0100 00074 0 C
87 2114 08/19/2009 "E" 2 2 1 15 0100 00049 0 C
87 2114 08/19/2009 "E" 1 2 1 15 0100 00024 0 C
87 2114 08/19/2009 "W" 3 2 1 15 0100 00092 0 C
87 2114 08/19/2009 "W" 2 2 1 15 0100 00086 0 C
87 2114 08/19/2009 "W" 1 2 1 15 0100 00036 0 C
87 2114 08/19/2009 "E" 3 2 1 15 0115 00079 0 C
87 2114 08/19/2009 "E" 2 2 1 15 0115 00047 0 C
```

Traffic Count Text file



Hourly Count Profile

Appendix 4d:
HOV/HOT Lane Analysis Utility



Statewide Transportation Engineering Warehouse for Archived Regional Data (STEWARD)

Florida Department of Transportation
University of Florida
Transportation Research Center

SunGuide High Occupancy Toll - Traffic Evaluation Report (HOTTER)

Release 1.0 Beta, December 3, 2009

Introduction

This document describes a utility program for analyzing the operation of managed lanes within SunGuide traffic management systems in Florida. The program uses data downloaded from the STEWARD database on the internet at the following web site:

http://cdwserver.ce.ufl.edu/steward/index.html

The program installation package and the user guide for downloading data are also available on this website.

Managed lanes may be implemented as high occupancy Toll (HOT) or high occupancy vehicle (HOV) lanes. Hotter accommodates both types of lanes.

Data Flow

An overview of the HOTTER data flow is shown in Figure 1. The numbers on this figure refer to the steps in the process. The steps are described as follows:

1. SunGuide generates a separate archive file for each 24 hour day of operation. The archive reports volume, speed and occupancy for each polling interval. Polling intervals are either 20 sec or 30 sec, depending on the district preference.
2. The extraction, transformation and loading (ETL) utility reformats the raw SunGuide data, adds some facility configuration data and accumulates the results.
3. The results are stored in files accumulated by 5, 15 and 60 minute intervals (HOTTER accommodates 15 or 60 minute intervals).
4. The accumulated data records are transferred into the STEWARD database.
5. The STEWARD database is queried manually from an Internet site. The data files may be isolated by facility, date, etc.
6. The results of the queries are stored as comma-delimited (CSV) files on the local computer. The file names are assigned by the user.
7. The HOTTER program reads and processes the CSV files
8. Analysis reports are generated in CSV format for importing into Microsoft Excel worksheets.
9. Results may be plotted if desired from the analysis files.

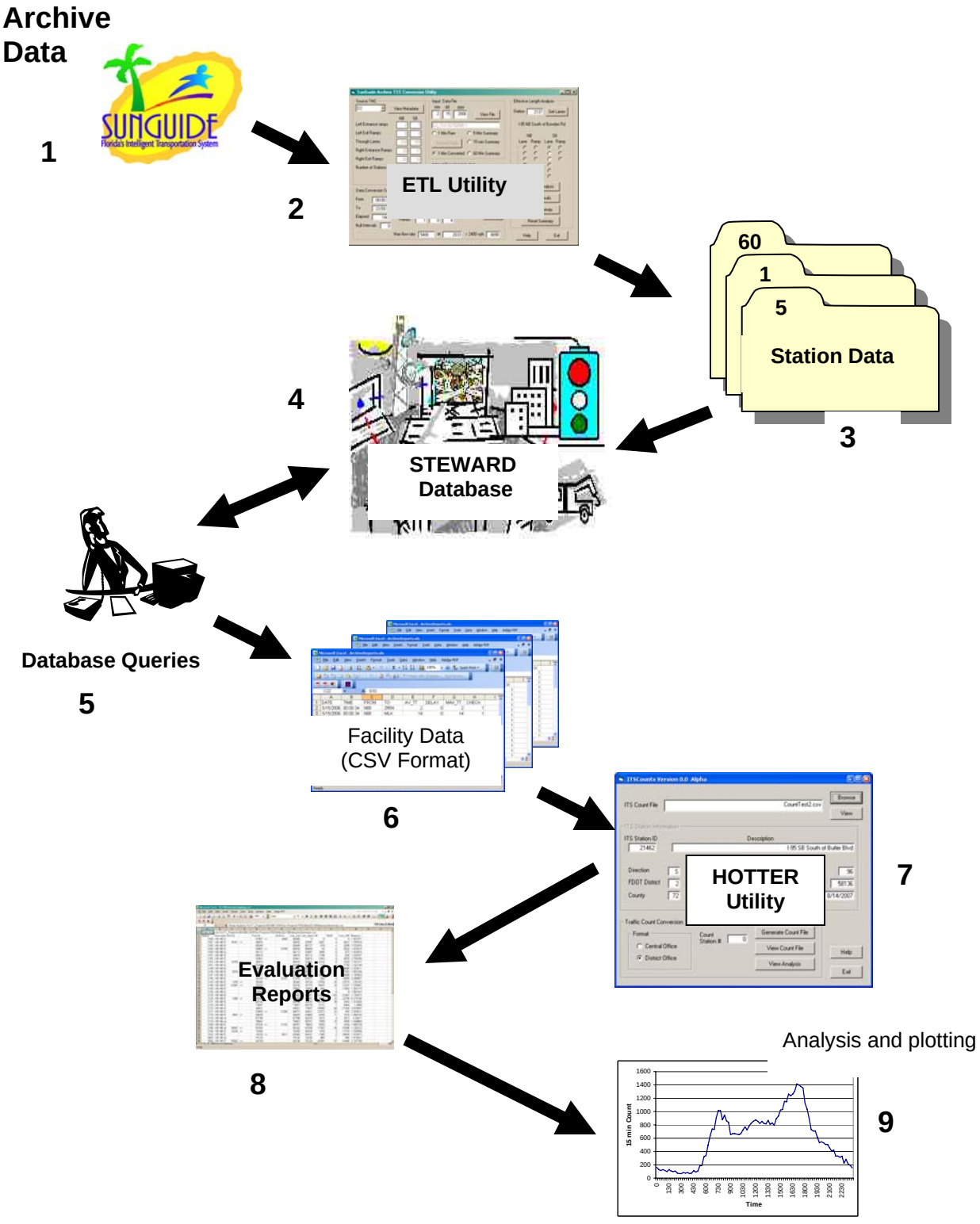


Figure 1: Overview of the HOTTER data flow

Installing the HOTTER Program

HOTTER is installed on a local computer using the normal Windows installation process. Three sub-folders are created within the program folder during the installation:

- **ITSData**, which contains the CSV files produced from the STEWARD database query in Step 6 above
- **Outputs,** which contains the analysis reports and summaries produced by HOTTER
- **Working,** which contains a parameter file that gives the locations of programs, folders and URLsrequired by the program. The installer must have read and write accessibility to the program folder. After the program has been installed and set up, the operator must have read and write accessibility permission for all of these folders.

The overall file structure for HOTTER is shown in Figure 2. The installation routines will establish the default folder and program locations. If a different configuration is desired, the locations may be changed from within the program, keeping in mind that the operator must have read and write access to the working folder and all of the data folders.

Note: You will find that the installation will be much easier if you choose all of the default installation options. The discussion and examples in this document assume that the default options have been chosen.

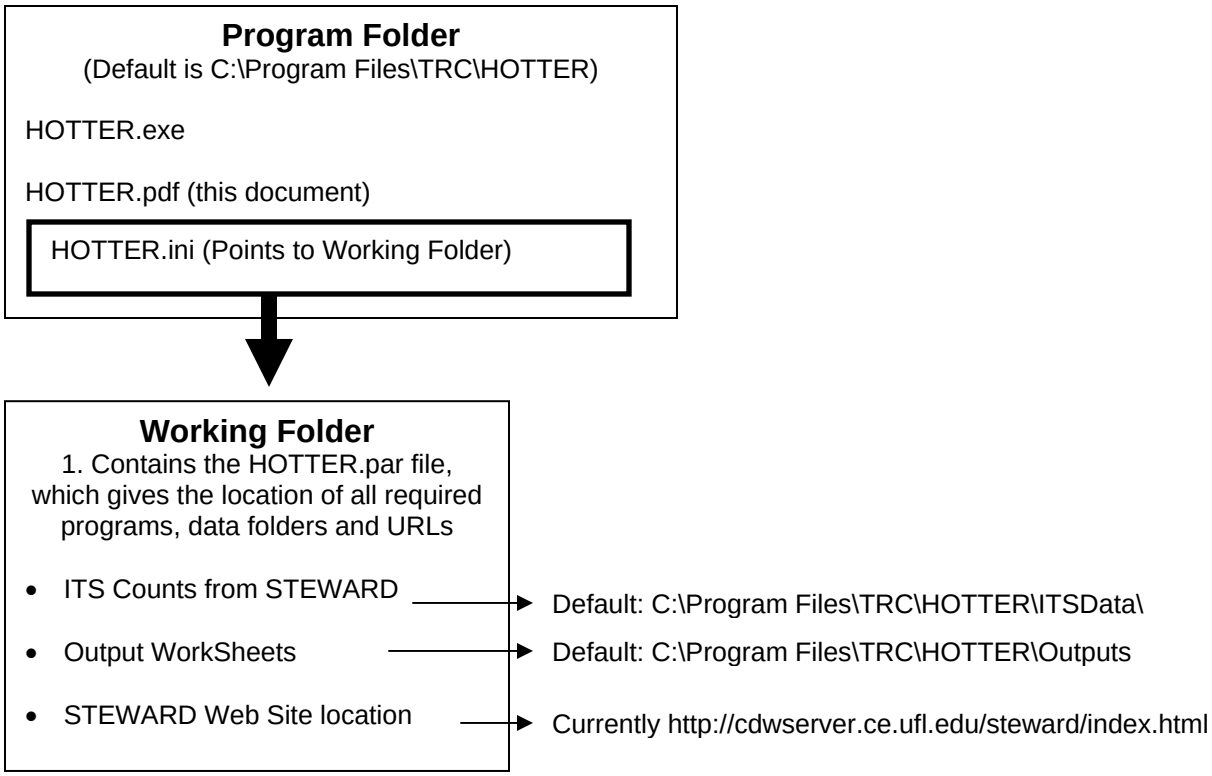
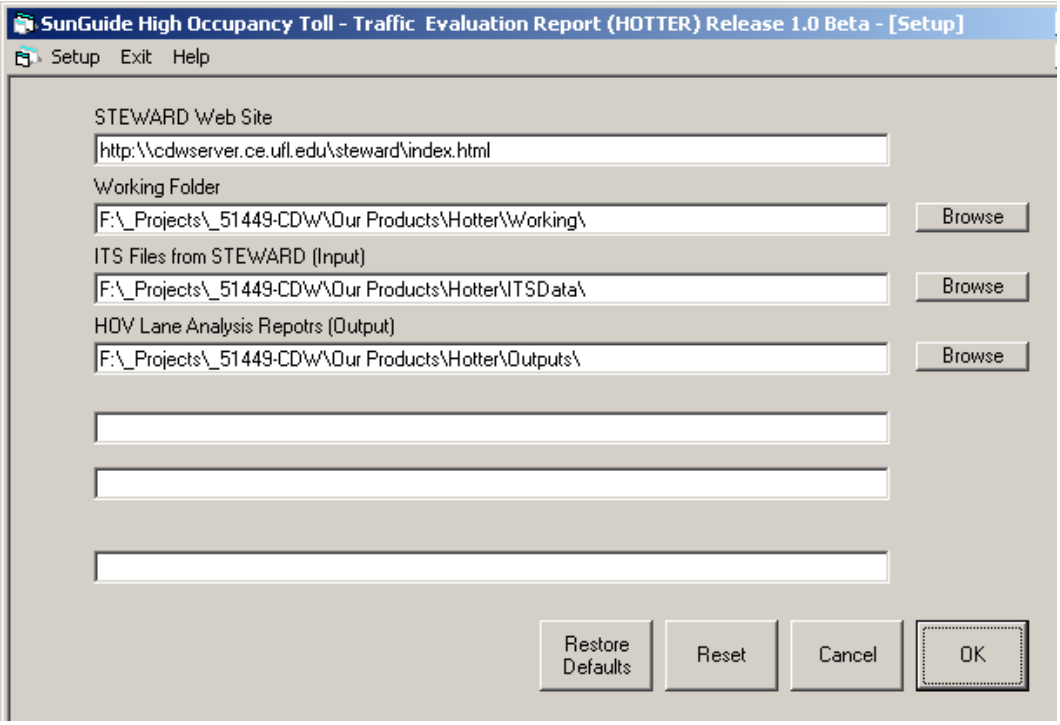


Figure 2: HOTTER folder structure

First Time Setup for File locations

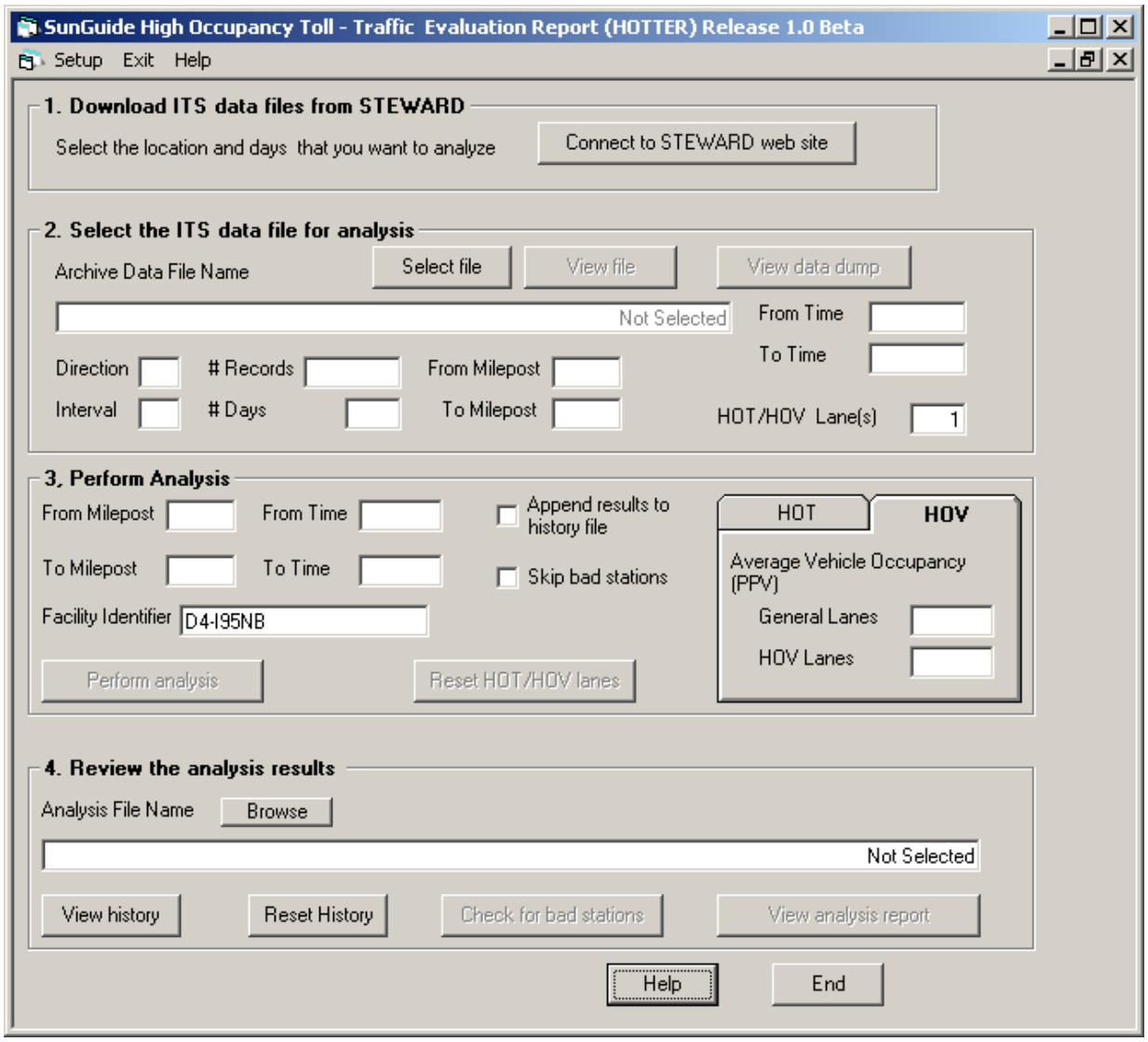
When you run the program for the first time, you will get a message telling you that the file configuration needs to be set up. Click ***OK*** and you will see the following screen:



The default file locations specified in Figure 2 will be inserted when you first enter this screen. You can browse for each of the items using the ***Browse*** button. The selection will lead you first to the default file locations but you may select any folder that already exists on your system. The step will be very easy if you use the default locations. All of the required folders were created when the program was installed. Click ***OK*** when you are finished.

Running the HOTTER Program

The following discussion assumes that you are familiar with basic Windows operations and that you have set up the program using all of the default program and file locations. When you run HOTTER, the main screen will appear as shown here. Note that there are four frames, each of which is associated with one step in the analysis process. The frames are numbered in the order of the analysis steps.

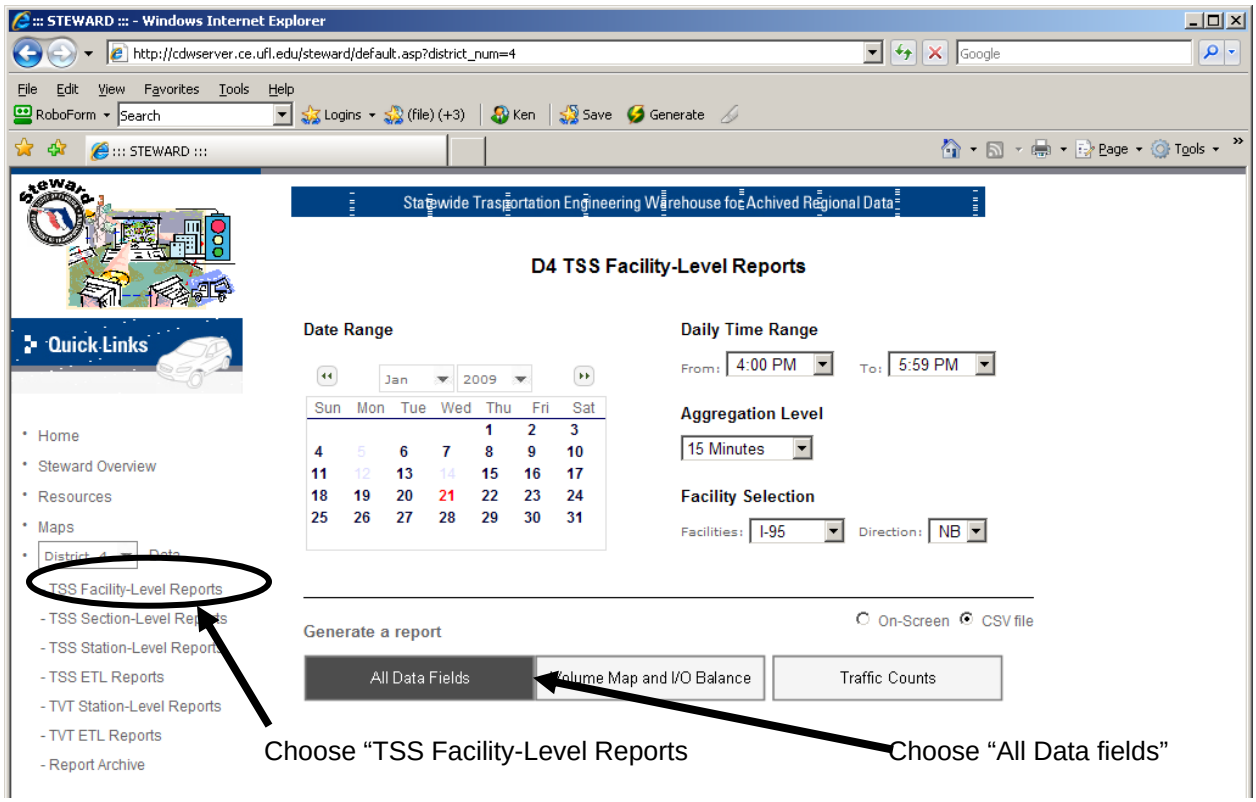


1. Download ITS data file from the STEWARD web site. It is important to give the file a name that you will recognize later.
2. Select the ITS data file for analysis by browsing for the file name that you just downloaded.
3. Perform the analysis.
4. Review the analysis results.

The HOTTER Program acts like a wizard that guides you through these steps. Each of the steps will now be discussed in more detail.

Step 1: Download ITS Data Files from the STEWARD Web Site

When you click the command button in this frame you will be signed on to the STEWARD web site. The frame at the left of the screen lets you choose what you want to do. HOTTER requires *facility level reports*. A sample facility level screen is presented on the next page.



Sample Facility Level Reports Page

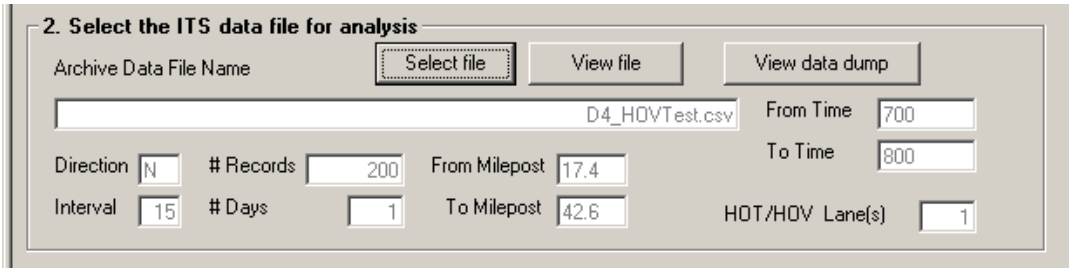
For HOTTER analysis, you must download the *facility level* data using the *all data fields* option. The sample screen above shows the following selections:

- **Date**: Jan 21, 2009 (Facility level data downloads are limited to a single day.)
- **Time Range**: 4:00 PM to 6:00 PM. (The screen actually shows 5:59 PM because this represents the beginning of the last minute in the time range.)
- **Aggregation Level**: 15 minutes (Either 15 or 60 min aggregations may be analyzed.)
- **Facility**: I-95, NB (The two directions must be analyzed separately.)
- **File Format** (On-screen or CSV file): Choose the CSV format to download the file.

Now, Click the ***All Data Fields*** button and follow the instructions to download the file to your own computer. The file location should be specified as the folder in which HOTTER will look for STEWARD data files. The default location is *c:\Program Files\TRC\HOTTER\ITSData.*

Step 2: Select the File for Analysis

The file is selected by clicking the ***Select File*** button in Frame 2. When the file is read, a summary of the contents will be displayed. The file selected in Step 1 above should produce the following display in this frame:

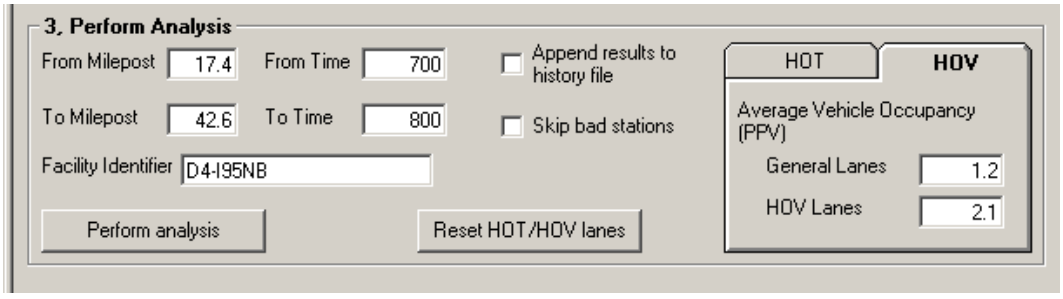


Some managed lane facilities have more than one lane designated. To specify the lanes, simply enter all of their numbers in the ***Hot/HOV lane(s)*** box in any order. For example, if lanes 1 and 2 are designated as managed lanes, then you should enter either "12" or "21."

Clicking the ***View File*** button at this point will display the file contents (i.e., raw data) in an Excel worksheet. Clicking the ***View data dump*** button will display a table of intermediate results that will be used later in the analysis step. The data dump is primarily a program development feature.

Step 3: Perform Analysis

When the file is selected, some of the fields on the analysis frame will also be filled with default data. In the absence of empirical data, the default values for passenger occupancy were assumed to be 1.2 PPV and 2.1 PPV in the HOV lanes and the general lanes, respectively. Since you might not want to analyze all of the stations for the full time range, you may now specify the station range (by milepost) and the time range. The default ranges include the full station and time range. These ranges should be edited if you want something different.



Some other parameters must be specified before the analysis takes place:

1. *Facility Identifier:* The name you put here identifies the facility for creating two other files to be described later. The current default is "D4-I95NB," which corresponds to the example in this document. This name may be edited for other facilities.
2. *Append results to history:* HOTTER maintains a history file with a single record summarizing the results from each day. The name of the history file is derived from the

specified facility identifier. A record will be appended to the history file if you check this box.

3. *Skip Bad Stations:* Occasionally a freeway section will have bad detectors at one or more stations that would invalidate the analysis. These stations are identified in the next step. The analysis may be rerun, skipping the indicated stations.
4. *Reset HOT/HOV lanes:* Allows you to change the managed lane configuration. This action will require reselection of the ITS data file in Step 2.
5. *Managed Lane Parameters:* Separate tabs on this screen distinguish between HOT and HOV lanes. The HOV parameters, persons per vehicle (PPV) occupancies, are shown in the figure above. The HOT lane parameter is the cost to the motorist per mile of travel in the HOT lane.

4: Review the Analysis Results

When the analysis is run, three files are produced:

1. The *Analysis Results File*: This file contains the following information:
   - HOT Lane Volume
   - HOT Lane Speed
   - General Lane Volume
   - General Lane Speed
   - Speed Difference (HOT lane speed – General lane speed)
   - Speed Ratio (HOT lane speed / General lane speed)

   The information is summarized separately by time period and station.

2. The *History File:* Each day's analysis produces a single record with the average values for the five items listed above. The daily record is appended to the history file. So by performing the analysis over a number of days to identify trends. The history file is also a CSV file that can be opened as a worksheet for plotting results.

3. *The Station List File:* This is a text file that contains a list of all stations with the speeds and volumes presented for each station. The list may be inspected to identify stations that should be eliminated from the analysis. Stations to be eliminated in future analyses are identified by inserting "X" as the last character in the line that identifies the station and saving the file. The analysis should then be rerun with the ***Skip Bad Stations*** box checked.

Other command buttons in this frame include:

- ***View History***, which will display the history file representing each day of analysis
- ***Reset History***, which will erase the history file contents and establish the history file column headings. You should save the history file under a different name to preserve the contents. You should also reset the history each time you establish a new facility for analysis to create the column headings for that facility.
- ***Check for Bad Stations,*** which will display the list of stations in the facility, indicating which ones are currently marked as bad.

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.

Example Results

The performance measures have been incorporated into this 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: 25145
- Vehicle-hours spent in the managed lanes: 372
- Average speed in the managed lanes (mph): 67.5
- Vehicle-miles traveled in the general lanes: 143069
- Vehicle-hours spent in the general lanes: 2374
- Average speed in the general lanes (mph): 60.3
- Vehicle-miles traveled in all lanes: 168214
- Vehicle-hours spent in all lanes: 2746
- 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: 52803
- Passenger miles traveled in the general lanes: 171682
- Total passenger miles traveled: 224486
- Passenger hours spent in the managed lanes: 781
- Passenger hours spent in the general lanes: 2848
- Total passenger hours spent: 3630
- 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.