

The WisTransPortal Volume, Speed, and Occupancy Application Suite

**SPR 1000-41-32
TOPS 2006-005**

January 2007

To Be Presented at the 2007 TRB Annual Meeting

**Steven T. Parker
Michael J. Runnels
Shan Di
Changxuan Pan
Douglas Dembowski**



**Wisconsin Traffic Operations and Safety (TOPS) Laboratory
University of Wisconsin-Madison
Department of Civil and Environmental Engineering**

DISCLAIMER

This research was funded by the Wisconsin Department of Transportation. The contents of this report reflect the views of the authors who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Wisconsin Department of Transportation, Federal highway Administration or the University of Wisconsin.

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The agencies listed above assume no liability for its contents or use thereof. This report does not constitute a standard, specification, or regulation, and its contents are not intended for construction, bidding, or permit purposes.

The name of any products or manufacturers listed herein does not imply an endorsement of those products or manufacturers. Trade and manufacturers' names appear in this report only because they are considered essential to the object of the document.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle The WisTransPortal Volume, Speed, and Occupancy Application Suite.		5. Report Date January 2007	
		6. Performing Organization Code	
7. Author(s) Parker, Runnels, Di, Pan, Dembowski		8. Performing Organization Report No.	
9. Performing Organization Name and Address University of Wisconsin-Madison Wisconsin Traffic Operations and Safety Laboratory 1415 Engineering Hall, Engineering Drive, Madison WI 53706		10. Work Unit No. (TRAIS)	
		11. Contract or Grant No.	
12. Sponsoring Agency Name and Address Wisconsin Department of Transportation 4802 Sheboygan Avenue Madison, WI 53703		13. Type of Report and Period Covered	
		14. Sponsoring Agency Code	
15. Supplementary Notes			
16. Abstract The Wisconsin Traffic Operations and Safety Laboratory at the University of Wisconsin-Madison has developed an online web-facility, the WisTransPortal V-SPOC Application Suite, for analyzing traffic detector volume, speed, and occupancy data from the Wisconsin Department of Transportation (WisDOT) ITS network. The V-SPOC project was initiated by WisDOT to improve upon an existing application at the Statewide Traffic Operations Center in Milwaukee that supports several engineering and planning functions. The new V-SPOC web-facility includes a complete detector database query selection tool, data visualization and export capabilities, data quality reporting, and integration with other ITS data sets. As part of the V-SPOC project work-plan, the new software was also designed to be portable and freely distributed for use in other states. This paper provides an overview of the V-SPOC Application Suite capabilities and technical design.			
17. Key Word ATMS, ADMS, ITS, Traffic Operations, Data Archiving, Online Query Tools		18. Distribution Statement	
19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price

ABSTRACT

The Wisconsin Traffic Operations and Safety Laboratory at the University of Wisconsin-Madison has developed an online web-facility, the WisTransPortal V-SPOC Application Suite, for analyzing traffic detector volume, speed, and occupancy data from the Wisconsin Department of Transportation (WisDOT) ITS network. The V-SPOC project was initiated by WisDOT to improve upon an existing application at the Statewide Traffic Operations Center in Milwaukee that supports several engineering and planning functions. The new V-SPOC web-facility includes a complete detector database query selection tool, data visualization and export capabilities, data quality reporting, and integration with other ITS data sets. As part of the V-SPOC project work-plan, the new software was also designed to be portable and freely distributed for use in other states. This paper provides an overview of the V-SPOC Application Suite capabilities and technical design.

INTRODUCTION

The Wisconsin Department of Transportation (WisDOT) has invested heavily over the past decade in its Intelligent Transportation Systems (ITS) infrastructure. This includes the emergence of a Statewide Traffic Operations Center (STOC) in Milwaukee and the deployment of Advanced Transportation Management System (ATMS) software at the STOC and other regional offices. In addition, WisDOT has formed a partnership with the Traffic Operations and Safety (TOPS) Laboratory at the University of Wisconsin-Madison to develop a Wisconsin Transportation Data Hub (WisTransPortal). TOPS Lab works closely with the STOC to provide support for ITS data integration, archiving, and applications development on the WisTransPortal infrastructure.

In this paper, we describe the WisTransPortal Volume, Speed, and Occupancy (V-SPOC) Application Suite that is scheduled for release August 15, 2006. The V-SPOC suite consists of a collection of web-based tools to access archived ATMS traffic detector information for purposes of corridor-based performance analysis and freeway management. The web-interface includes a complete detector database query selection tool, data visualization and export capabilities, data quality reporting, and integration with other ITS data sets in the WisTransPortal database.

A number of states have developed traffic detector data archiving and analysis systems as part of their ITS programs. Like V-SPOC, many of those systems were developed fully or in part through state DOT / university partnerships. Examples include ADMS Virginia (University of Virginia), California PeMS (University of California Berkeley / Berkeley Transportation Systems), CATT Lab (University of Maryland), Central Florida Data Warehouse (University of Central Florida), Kentucky ADMS (University of Kentucky), and PORTAL (Portland State University).

The V-SPOC suite was developed to overcome several limitations in an existing desktop application at the WisDOT STOC. Desired enhancements included increased efficiencies in the data retrieval process, access to users outside of WisDOT to easily obtain data, and integration with a monthly performance measure report. In addition, the V-SPOC project work-plan included a requirement to design the new software to be portable and freely distributed for use in other states.

INFORMATION TECHNOLOGY FRAMEWORK

In this section, we provide a general overview of the WisTransPortal information technology platform and the design concept for the V-SPOC project.

WisTransPortal Overview

The WisTransPortal Project was started in June 2003 at the TOPS Lab to support WisDOT in the following areas: ITS and safety data archiving, real-time traffic information data sharing, transportation operations applications development, and transportation research. The Phase I project, which was completed in December 2005, included developing a business plan and technical design for the WisTransPortal system, deploying the primary hardware and software infrastructure, and developing initial data archiving capabilities.

An important component that was started in Phase I is a physical link from the WisTransPortal system to the WisDOT fiber optic backbone (ITSNET). The project plan for

Phase II, which started June 2006, covers tasks to extend the WisTransPortal data sharing capabilities with respect to the ITSNET link and to further develop the database and software components of the system. Specifically, the Phase II project has the following objectives: enhanced support for data archiving capabilities, support for real-time data sharing capabilities, support for web-applications development (online query tools and map presentations), and support for ITS data standards initiatives.

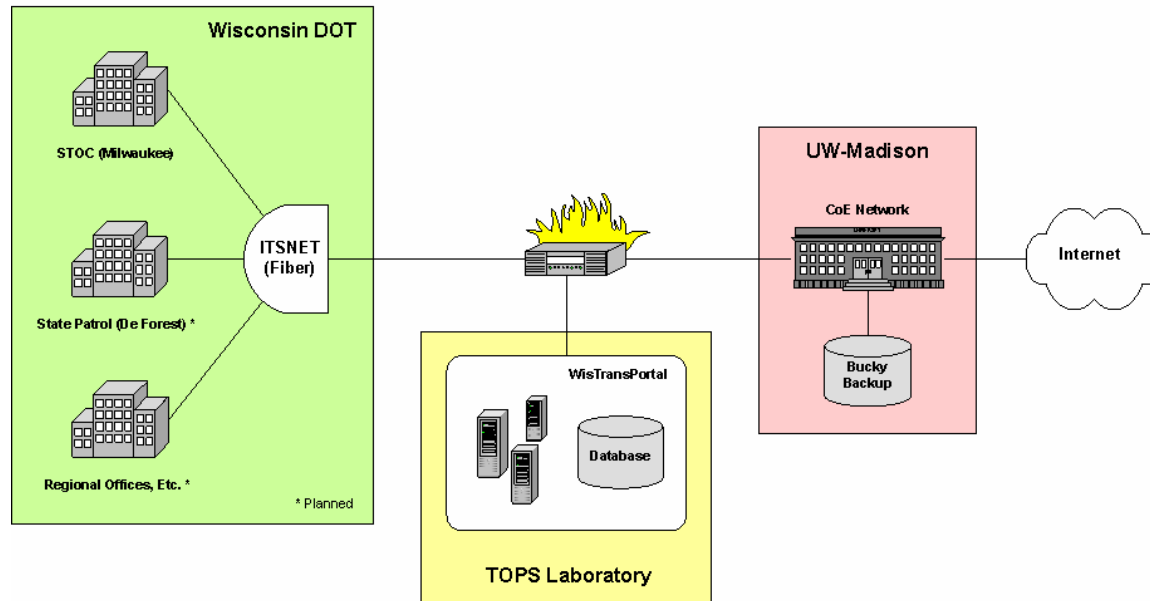


FIGURE 1 High level overview of the WisTransPortal network architecture. The WisTransPortal sits directly on the UW College of Engineering network and connects to WisDOT over the ITSNET fiber backbone or via the internet.

The WisTransPortal software design incorporates two basic functions: ITS data archiving and ITS data dissemination. The data archiving category consists of a collection of automated Java programs that regularly update the WisTransPortal Oracle 10g database with ITS data from WisDOT and other sources. The current data archiving capabilities are focused primarily on replicating and integrating five existing WisDOT and regional data sources: traffic detector data, lane and ramp closure data, traffic incident data, historical crash reports, and road weather information.

The data dissemination category consists of a web-application platform based on Java Enterprise (J2EE) servlet and JSP technologies (1), Open Source Java web-application frameworks such as Apache Struts (2), and Hibernate (3). In addition to V-SPOC, there are several applications being developed or slated for development on the WisTransPortal infrastructure, including:

- Development of a statewide online Lane Closure Permitting System.
- Development of a Wisconsin Safety Analysis Toolbox.
- Support for interagency traffic camera video sharing.
- Support for Wisconsin 511 traveler information.

- Integration with internal WisDOT data systems and applications.

Additional information about the WisTransPortal can be found at the project website:
<http://transportal.cee.wisc.edu/>.

V-SPOC Overview

The V-SPOC project was initiated to replace and improve upon an existing “Data Extractor” application at the STOC. The Data Extractor had been in use for several years by WisDOT traffic engineers and planners, but it was known to have several limitations:

- It was limited to Milwaukee area traffic detectors.
- It had several performance bottlenecks related to its design model.
- The archived detector logs were stored in compressed files on the STOC local network and required specialized programs for accessing the data.
- There was a need to integrate detector data with other operations data such as Milwaukee County incident reports.
- There was a need to add additional flexibility and functionality into the interface; the ability to provide better data quality analysis was particularly important.
- The current interface was felt to be difficult to navigate and was not user friendly.

In order to address these limitations and requirements, the V-SPOC work plan included a full redesign of the application software and database management model by developing the new version on the WisTransPortal infrastructure. In particular, the V-SPOC Application Suite was built to the following specifications:

Accessibility: The V-SPOC Application Suite was designed to be a super thin-client web-based system. Most of the work in V-SPOC is performed on the WisTransPortal, either on the web server itself or in the database. Users can access V-SPOC from virtually any internet connection using standard browsers and take full advantage of its capabilities.

Modularity: The V-SPOC software design is based on Java Enterprise Model-View-Controller (MVC) / Model-2 methodologies. These methodologies promote modular design by decoupling large enterprise level applications into smaller, independent components. V-SPOC was developed, in particular, with Apache Struts which is a well-established open source framework for developing MVC / Model 2 web applications.

Portability: Developing platform independent code was an important design goal not only to promote the use of V-SPOC in other states, but to ease future maintenance by TOPS staff. The modular design of V-SPOC itself is perhaps the most significant factor that ensures portability. In addition, the V-SPOC web-application code adheres to SUN Microsystems Java servlet specification, which provides for portability between web application servers. It also follows best practices in Java programming and database access. V-SPOC incorporates several Open Source Java libraries to handle specific requirements, such as writing to Microsoft Excel and generating graphical output on the web page. By using open source, TOPS Lab has access to all

source code in V-SPOC and has the ability to distribute the complete code base with the V-SPOC system.

High Performance: V-SPOC was designed to be a high performance system. Most of the work is performed in the WisTransPortal Oracle 10g database which is tuned for very large datasets and resides on server class hardware. Web-applications based on Java Servlets are also highly efficient due to the client/server request model built into the web-application architecture. Although network latency is always an issue with internet applications, current benchmarks indicate that the benefits of moving V-SPOC to a server based model have far outweighed the costs in communication time.

V-SPOC APPLICATION DESIGN

The V-SPOC application suite is developed as a set of specialized modules that perform various traffic detector data retrieval and analysis functions. Although each module is designed for a particular task, they are built around a common framework for detector data organization, retrieval and visualization. In this section, we describe the general features of the V-SPOC Application Suite which provides an overview of the capabilities of each module.

General V-SPOC Features

The WisDOT ATMS system monitors nearly 4000 detectors in the combined Milwaukee, Madison, and Wausau metropolitan regions. Detector data is archived in the WisTransPortal detector database from the ATMS at 5 minute intervals. Hence, V-SPOC manages nearly one million detector records per day. In order to facilitate the use of this data, the V-SPOC application suite provides mechanisms to organize detector data into spatial (Corridors, Count Locations and Controllers) and temporal (Time Intervals) groupings.

Detector Selection

Traffic detector location information in V-SPOC is organized into a quasi-hierarchical structure consisting of individual traffic *detectors*, *controllers*, *count locations* and *corridors*. Whereas detectors and controllers represent physical objects, count locations and corridors are logical groupings assigned through V-SPOC. At the highest level, corridors are assigned to one of five WisDOT regions (NW, NC, NE, SW, SE). Generally, users analyze specific traffic detector data in V-SPOC by first selecting a corridor of interest, then selecting a count location or controller belonging to that corridor, and finally selecting from the resulting multiple-selection list of detectors.

Time Selection

The V-SPOC mechanism for choosing the time component is through the selection of multiple *time intervals*. Each time interval can be at most a 24-hour segment and may cross date boundaries (e.g., 11:00 PM 12/31/05 to 1:00 AM 1/1/06). Detector data is retrieved and analyzed based on a set of one or more time intervals. Time intervals are chosen by first selecting the start time and end time to be used for the segments. Next users select the starting

dates for each time interval from a calendar component. Although it is common to choose a set of contiguous 24-hour time intervals (e.g., all day Tuesday, Wednesday, and Thursday for a given week), time intervals may be chosen to capture a particular travel time (e.g., every Monday from 3:00 PM – 6:00 PM during January and February 2006). Common start time/end time presets are provided in a drop-down list. After both detector and time interval selections are made the user continues to the results interface where data analysis takes place.

FIGURE 2 The V-SPOC General Detector Selections screen. The purpose of this page is to fix the spatial (list of detectors) and temporal (list of time intervals) dimensions for the resulting detector database query.

Aggregation

The V-SPOC analysis interface supports a combination of data aggregation modes. At the basic level (no aggregation), detector data is provided at the database row level of 5 minute data values for individual detectors over each time interval. Users may refine this representation by averaging over all detectors, averaging over all time intervals, or averaging over both detectors and time intervals. For example, given three detectors and five time intervals, the “no aggregation” output would result in fifteen time series (the combination of all detectors and time intervals). Aggregating over detectors will result in five plot series, each plot series corresponding to the average volume and/or speed and/or occupancy over all three detectors for each time interval. Aggregating over time intervals will result in three plot series corresponding to the average volume and/or speed and/or occupancy over all five time intervals for each detector. Aggregating over both dimensions will result in one plot series showing the average performance of a roadway determined by the set of three detectors and five time intervals. In addition, detector data may be presented in terms of 5 minutes, 15 minutes, 30 minutes, or 60 minute intervals.

The aggregation level is a primary setting that determines how detector data will be represented in the various analysis and export tools. In particular, the data export function

(described below) provides the plot series data that is shown in the graph. Moreover, the quality assurance report shows the results for all combinations of aggregation selections.

FIGURE 3 The V-SPOC General Detector Results screen. The purpose of this screen is to refine the detector database query parameters with respect to detector and time interval selections, detector variables (volume, speed, and occupancy), aggregation levels, and presentation attributes.

Visualization

Generally, data is prepared for plotting, export, and quality analysis. Specialized engineering functions are described in the next section.

The basic plot window provides a graphic representation of a set of one or more detector data time series. As described above, the time series may represent raw samples or aggregated detector data in one or two dimensions. A set of time series may be plotted side by side in one plot window or over multiple plot windows. In particular, a user may elect to tab through plot windows grouped by detector, time interval, or both detector and time interval. This mechanism for distributing time series over separate plot windows has a natural connection to the aggregation level. Taking the example above, aggregating over detectors results in five time series, one for each time interval. It is possible then to view each aggregated series side by side or spread over five plot windows tabbed by time interval.

Data Export

All detector data is available for download to a comma separated value (CSV) text file which is readily imported into a third party analysis tool such as Microsoft Excel or Access database. The representation and number of time series in the CSV file depends on the selected aggregation level. In addition to the basic download function, specialized features and file formats are described in the next section.

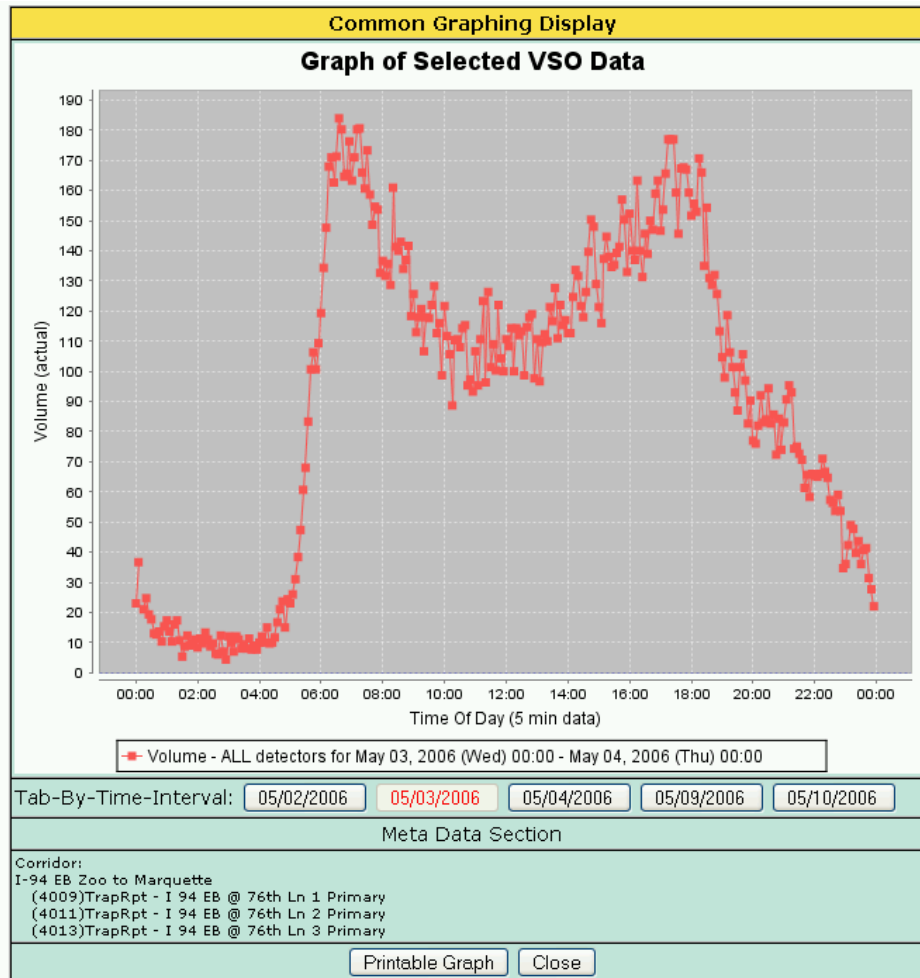


FIGURE 4 A representative V-SPOC plot window. Detector volume time-series are averaged over all detectors and the plot window is set up to tab through individual time intervals.

Other Features

The results page supports several other features that are common to most modules in the V-SPOC Application Suite:

- **Quality Assurance Report** – an important use of the new V-SPOC application is to provide more efficient methods to identify missing or bad data. Data problems arise from a variety of sources, including malfunctioning detectors and intermittent problems with the ATMS archiving functions. The quality assurance report provides the results of a set of data quality tests for the given selection of detectors and time intervals. The tests are reported in terms of each combination of aggregation levels and for individual 5-minute detector records. The quality assurance report provides a summary overview and quantitative record of data quality issues that impact the visualization and higher-level analysis functions. The report is arranged in two parts: the first part presents a series of basic test results for missing and repeating values in the data. The second part presents the results of seven data screening

tests which were developed at the STOC for purposes of analyzing detector data quality at a high level.

- Traffic Events Report – this feature produces a table of lane and ramp closures that are associated with the detector and time interval selections. The purpose of this report is to correlate additional traffic event information with observed patterns in the detector data. It is expected that the events reporting capabilities will be developed over time to include special events and incidents. This function, in particular, leverages the full WisTransPortal data archiving infrastructure for transportation operations and safety data integration.
- Graph and File Options - these features include the ability to select local or global axis scaling for the plot window, view the underlying SQL database query associated with a particular data visualization or export action, and select a method for filling in missing detector data. This last feature, in particular, is currently limited to filling in missing data (those values that are NULL) with ZERO values. However it is expected that over time higher-level imputation techniques such as data smoothing will be implemented. Data quality and imputation is discussed further in the Section on Database Issues.

Specialized V-SPOC Modules

The objective of the V-SPOC Application Suite is to provide an integrated system for transportation operations engineering, analysis, and reporting functions related to traffic detector data. Towards that end, V-SPOC is designed as a set of customized modules that take advantage of a common database backend and web based application platform. The common data selection, visualization, export, and reporting functions were described above. The database features will be discussed in the next Section. We now describe the various specialized modules that are available in V-SPOC. Because of the overall modular design of V-SPOC, it is anticipated that additional modules will be added over time with minimal effort.

General Detector Data Retrieval

The General Detector Data Retrieval module is a general purpose tool for basic selection and analysis of detector data. It essentially provides the common functions for selection, visualization, export, and quality analysis described above. This module is sufficient for most data research and, in particular, supports the common task of retrieving archived detector data for export into third party analysis and reporting tools.

Corridor Analysis

The Corridor Analysis module is used to analyze traffic patterns along a single corridor and to evaluate the performance of the traffic detectors in the system. Unlike the General Detector module, which allows the user to create almost any grouping of individual detectors, the count location is the basic unit of detector measurement in the Corridor Analysis module. The core functionality of Corridor Analysis is built into the Corridor Balancing Tool which is based on the selection of count locations and time intervals. The purpose of this tool is to provide a graphical interface to measure the performance of the traffic detectors by highlighting non-zero net inflow

and outflow volume along a corridor. The graphical tool supports manual calibration of the count location data.

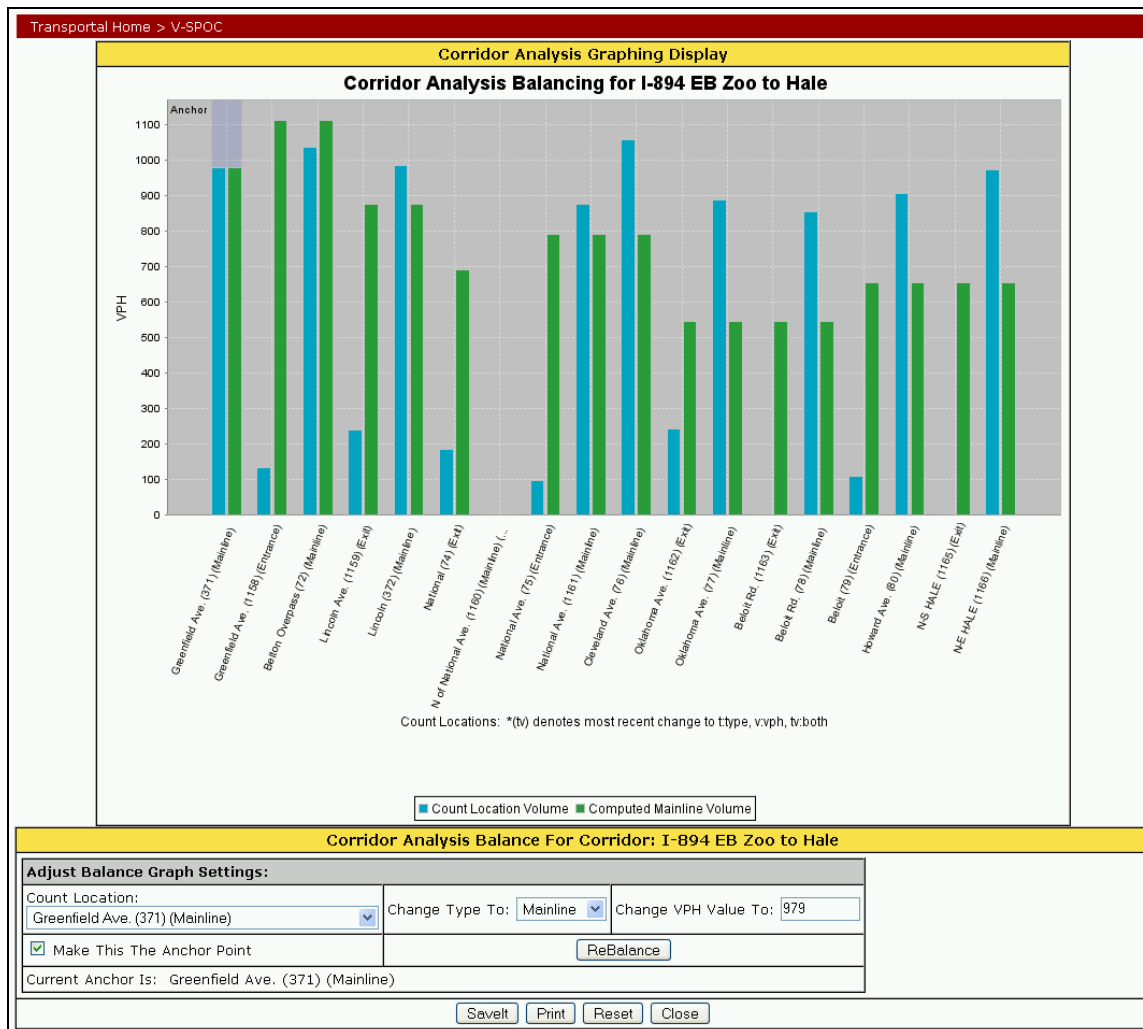


FIGURE 5 The V-SPOC Corridor Balancing Tool. This tool provides a mechanism to evaluate the performance of the ITS network by comparing actual count location volume counts with the accumulated net flow on the network.

Monthly Data Retrieval

The General Detector module is useful for working interactively with archived detector data but is best suited to handle small groups of detectors and time intervals. The Monthly Data Retrieval module provides a batch mode capability to download a complete month of detector data for an entire WisDOT region. Data is distributed across several files that are provided in a ZIP formatted archive. A typical monthly data file for the Milwaukee (SE Region) is about 400 MB and takes around 30 minutes to generate. As such, V-SPOC caches monthly data files on the server for a set amount of time to facilitate multiple downloads of the same data. This module was designed to help process ongoing data requests at the STOC. It will also provide the data

needed to produce a monthly performance measure report efficiently by reducing staff time for data retrieval.

Ramp Meter Retiming

One of the core uses of archived detector data at the STOC is for ramp meter retiming. A separate Ramp Meter Retiming Workbook (Excel application) has been developed at the STOC for that purpose and which integrates with the pre-existing Data Extractor application. Although the underlying software architecture of the two systems is different (V-SPOC is a web-based application built with Java tools whereas the Data Extractor is a desktop tool built with Microsoft Visual basic), the Ramp Meter Retiming Workbook is fully supported in V-SPOC. As in the Corridor Analysis module, the detector results page has a subset of the common data selection, visualization, and reporting capabilities for the purpose of data quality evaluation. The data download feature generates a new Ramp Meter Retiming Workbook by populating a template workbook with ramp detector data from the detector database. The specific retiming functions are done by macros built into the workbook.

The guiding principle in the design of the ramp meter retiming module is to decouple independent software processes. The role of V-SPOC is to provide a mechanism for retrieving archived detector data along with a basic toolbox to analyze and evaluate the resulting data. The ramp meter retiming workbook implements specialized functionality that depends on detector data but is otherwise developed and maintained separately. The link between these applications was developed using the standard application programming interface (API), and agreed upon set of data cells in the workbook.

TRADAS

WisDOT uses the TRADAS software system (5) to fulfill FHWA traffic volume reporting requirements. In addition to the operations-based controllers that make up the majority of the reporting devices on the ATMS system, there is a small subset of permanent planning based controllers in the field. Detector data is downloaded from the planning controllers over a modem every two weeks and sent to WisDOT in a standardized “261” file format for purposes of ingestion into the TRADAS system. The goal of the V-SPOC TRADAS module is to provide a more convenient TRADAS reporting mechanism with the flexibility to add additional detectors and controllers over time.

There is also a growing interest in WisDOT to use operations data for planning purposes and to better integrate the ATMS and TRADAS data archives. One goal of the V-SPOC system is to provide capabilities to analyze the historical performance of the ITS system in order to correct weaknesses in the system and gain experience in the use of archived ITS data for planning and other engineering purposes.

Corridor Meta Data

The association of detectors to corridors and count locations is the underlying model that drives most of the V-SPOC functionality. A general understanding of the corridor configuration and logical ordering of detectors is necessary to effectively use the system. The Corridor Meta Data module allows the user to select corridors in order to produce a list of all detectors for those

corridors broken down by corridor, controller, and count location. It is expected that this module will be expanded over time to include map presentations and other information about the freeway system. (Note: the Corridor Meta Data module now contains links to existing Controller Map PDF files).

Other Modules

Several other modules are still in development or planned for the next round of enhancements. The most important of these is a Corridor Management module that will allow V-SPOC administrators to create, modify, and delete count locations and corridors. (Note: the Corridor Management module, which was in development when this paper was first submitted, is now complete and in production). Other primary tasks that are in the planning stage include built-in support for a Performance Measures Report process under development at the STOC, and an advanced Data Quality module to provide detector data quality analysis and system performance evaluation.

In addition to extending V-SPOC functionality through the development of new modules, it is expected that V-SPOC will integrate with other major WisTransPortal applications. Of particular importance is a planned online Lane Closure Permitting System that will be developed on the WisTransPortal infrastructure. Moreover, as the WisTransPortal data hub itself adds more data sources, such as GIS mapping, traffic video, incident data, and weather data, new opportunities for ITS data integration will become available.

DATA ARCHIVING ISSUES

There are two basic components that make up the V-SPOC system: the application logic consisting of the data processing software and web interface pages described above; and the database backend. The database environment for V-SPOC is the WisTransPortal Oracle Database 10g Enterprise Edition (6). A custom built software application downloads detector data from the STOC every night and imports the data into Oracle. In this section, we highlight several aspects of the detector database and archiving process.

Database Issues

Currently, the WisTransPortal manages a complete archive of 5 minute detector data from the WisDOT ITS system. This includes Milwaukee area detector data (1996-present), Madison area detector data (2004-present), and the Wausau area (2005-present). There are approximately 4000 detectors on the ITS system within these three regions which generate over one million records per day. There are over 2.3 billion detector records in the WisTransPortal detector database to date.

Detector Data Management

The WisDOT ATMS software polls the controllers every 20 seconds to retrieve volume, speed, and occupancy detector data. This data is then aggregated and written to a Microsoft Access detector log every 5 minutes. The detector logs are rolled over every 24 hours. The pre-existing

Data Extractor used an in-house compression program to compact and store the 24 hour detector logs in the file system. Retrieving detector data in the Data Extractor involved decompressing the archived files and pulling out relevant data. Because each file contained a single day of detector data, retrieving data for several days involved decompressing several files of data for all detectors in the system. In contrast, data management in the WisTransPortal is based on a formal relational database.

Oracle 10g includes several improvements over previous editions for managing very large data sets. The WisTransPortal database makes extensive use, in particular, of Oracle partitioned tables and indexes, which allows one to transparently decompose data into smaller pieces to improve query performance and availability. Through the use of partitioned tables, it is possible to provide a uniform (one large table) view of the underlying detector data while localizing query processing and data management. The table and index objects in the detector database are also distributed over separate Oracle tablespaces (database storage structures) to reduce disk contention on large-scale queries.

Data Standards

The design of the WisTransPortal detector database incorporated elements from the Traffic Management Data Dictionary (TMDD) standard (4) whenever possible. In some cases, it was felt that strict adherence to the standard was not practical at the database level. For example, it was desirable to store detector data with greater precision than the standard requires, or to retain data outliers that exceed the range of permissible values for a given data element. The guiding principle, however, is that the WisTransPortal detector database is structured to support TMDD requirements on standard data elements. Hence, detector IDs are stored with up to 32 length character arrays, even though the STOC ATMS detector IDs are stored in integer format. A separate TMDD database view of the detector data was also developed as a proof of concept that the detector database is complete in its ability to process TMDD conforming message sets.

The corridor database, which includes location information and other metadata related to detectors, controllers, count locations, and corridors, is still in development. The project plan for the corridor database is to incorporate LRMS conventions in the location information.

The Data Archiving Process

The Java Archiving Service

The detector data archiving process is formally part of the WisTransPortal archiving system, which is separate from the V-SPOC application suite. Each night, the STOC ATMS system generates a Microsoft Access file containing detector logs for the 24 hour period. A separate Access file is generated at the Madison office. TOPS Lab staff developed a Java program to retrieve the Access files each night and to insert the records into the WisTransPortal database. The Java archiving program is completely automated – it runs as a Windows service on a separate server assigned for data acquisition and writes a log of its activity.

This data archiving program was developed before the ITSNET fiber connection was available. It is expected that program will be enhanced to take advantage of web service capabilities in the ATMS software to retrieve detector data in real-time.

Data Densification

One of the difficulties in working with ATMS detector data is that the raw data is generally incomplete. The ATMS system is a large network of servers, software, and communications infrastructure that ultimately connects the daily detector logging functions with detection devices in the field. Consequently, it is not uncommon for the detector logs to exhibit short-term (under five minutes) and longer-term breaks in data availability.

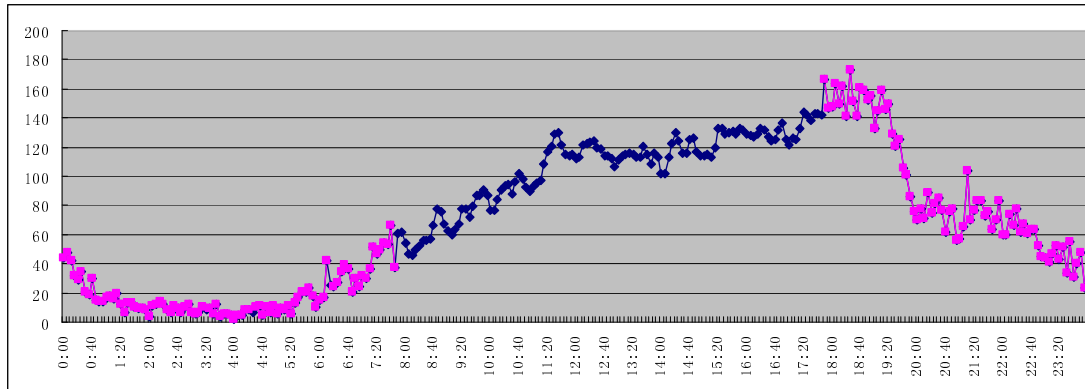
In order to provide accurate reporting and aggregation functions through V-SPOC, it is desirable to work with a dataset that has a uniform sample space. There are many ways to deal with this issue, including developing work-around procedures in the application code or modifying the raw data during the archiving process. The development team felt that this should be handled fundamentally at the database level in order to keep the application code simple and efficient. Moreover, different states might store detector data at varying levels of completeness; hence it was felt that a database solution would be most portable.

For the WisTransPortal system, raw data is archived exactly as it is received. In order to provide a uniform dataset to V-SPOC, a database *view* is generated by joining the raw detector data with template tables consisting of all detector IDs and 5-minute time-stamps. This view provides a dense representation of the detector data by filling in NULL values at timestamps where archived data is missing. Since a view is a just a virtual representation of underlying data, the TOPS development team is investigating the possibility of developing additional views that provide higher level imputation of missing values.

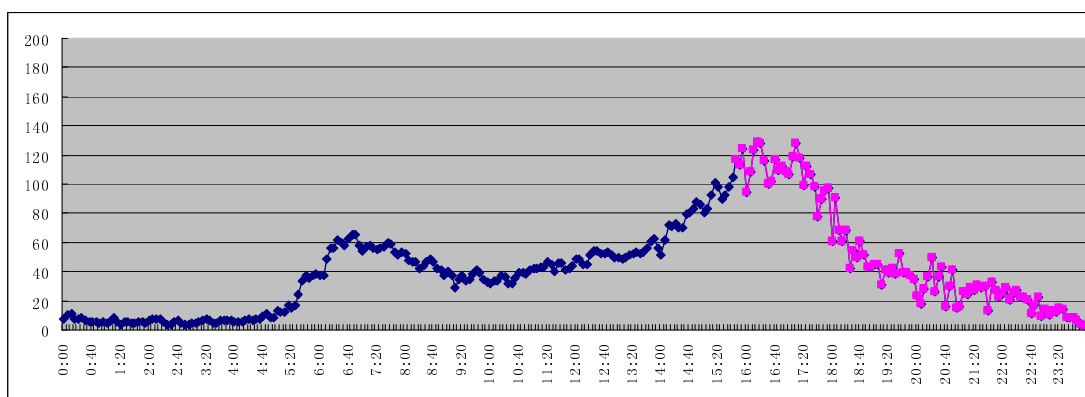
Data Imputation

An imputation program is under development to provide a complete view of daily detector data. At present missing and continuous zero records will be replaced by modeled data. The model generates a table of monthly averaged data aiming to capture daily variation patterns. The averaging is based on distinct temporal categories for each detector. We define the categories as seven weekdays for every 5 minute interval. A sufficient sample size is required for each category based on continuous one or two months of data. Outliers and single samples are discarded in the averaging. Time series interpolation is also applied to the modeled data for short breaks. Experimental results show that the daily variation pattern is steady and obvious for the defined categories. Substantive tests also indicate the modeled values are good substitutes for the missing detector data as they linked up the broken daily data curves very closely, especially for volume and occupancy.

An underlying assumption in V-SPOC is that any detector reports table must be uniform in its representation of time. By making this assumption, the problem of data densification is pushed down to the database level where it is most efficiently solved. Different states may choose to handle that requirement differently. It is also possible to implement several detector tables in different ways without impacting the V-SPOC system.



Detector 4009 02/19/2005



Detector 709 10/12/2005

FIGURE 6: Pink dots are raw detector data which is missing for part of the day. Blue dots are modeled values.

FUTURE WORK

A test version of V-SPOC has been available to engineers at the STOC since June 2006. The final version is scheduled for release August 15, 2006 and is expected to fully replace the existing Data Extractor application for use at the STOC, Regional Offices, and WisDOT Central Office. Because it was developed as a web-based system, it is also expected to provide an online interface for use by consultants and other traffic engineers outside of WisDOT.

A primary requirement of the V-SPOC design was to build for platform independence in order to facilitate the use of V-SPOC in other states. The STOC will provide support through TOPS Lab to assist in out-of-state deployment of the V-SPOC system.

The next phase of development is likely to focus on data integration with other WisTransPortal applications (such as the Lane Closure Permitting System) and data quality and system performance measuring capabilities.

ACKNOWLEDGMENT

The V-SPOC application suite was developed at the Traffic Operations and Safety (TOPS) Laboratory through collaboration with the Wisconsin Department of Transportation (WisDOT). The TOPS project team consisted of: Bin Ran (UW Principle Investigator), Steven Parker (UW Co-PI / Project Manager), Michael Runnels (Systems Programmer), Jiangang Lu (Research Assistant), Changxuan Pan (Research Assistant), Shan Di (Research Assistant), and Zhaozheng Yin (Research Assistant). The STOC team consisted of: Douglas Dembowski (WisDOT Project Manager), John Mishefske (WisDOT Information Specialist), Tim Vik (DAAR Engineering) and Ertan Ornek (DAAR Engineering).

REFERENCES

1. Singh, Stearns, Johnson, et al. (2002). *Designing Enterprise Applications with the J2EE Platform, Second Edition*. Sun Microsystems, Inc. <http://java.sun.com/blueprints/guidelines/>.
2. The Apache Struts Project (2005). *Action Framework User Guide. Version 1.3.0-dev*. <http://struts.apache.org/>.
3. The Hibernate Project (2005). *Hibernate Reference Documentation. Version: 3.1* <http://www.hibernate.org/>.
4. Institute for Traffic Engineers (2004). *Traffic Management Data Dictionary (TMDD), Version 2.1*. <http://www.ite.org/tmdd/>.
5. Chaparral Systems Corporation. *The Traffic Data System (TRADAS)*. <http://www.chapsys.com/tradas.html>.
6. Oracle Corporation. *Oracle Database Documentation Library. 10g Release 1 (10.1)*. <http://www.oracle.com/technology/documentation/index.html>