PB2003-100574

FINAL
CONTRACT REPORT

INCORPORATING ADVANCED SIGNAL CONTROL SYSTEMS INTO AN ARCHIVED DATA USER SERVICE PROGRAM

BRIAN L. SMITH, Ph.D.
Assistant Professor
Department of Civil Engineering
University of Virginia

MATTHEW CHRISTIAN GRIMES
Graduate Research Assistant

V · I · R · G · I · N · I · A
VTRC
TRANSPORTATION RESEARCH COUNCIL

VIRGINIA TRANSPORTATION RESEARCH COUNCIL

Standard Title Page - Report on Federally Funded Project

1. Report No. FHWA/VTRC 03-CR7	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Incorporating Advanced Signal Control Systems into an Archived Data User Service Program		5. Report Date October 2002
		6. Performing Organization Code
7. Author(s) Brian L. Smith and Matthew C. Grimes		8. Performing Organization Report No. VTRC 03-CR7
9. Performing Organization and Address Virginia Transportation Research Council 530 Edgemont Road Charlottesville, VA 22903		10. Work Unit No. (TRAIS)
		11. Contract or Grant No.
12. Sponsoring Agencies' Name and Address Virginia Department of Transportation 1401 E. Broad Street Richmond, VA 23219	FHWA 400 North 8th Street, Room 750 Richmond, VA 23239	13. Type of Report and Period Covered Final Report: October 2000 - October 2002
		14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

The deployment of Intelligent Transportation Systems that incorporate some type of data collection or traffic surveillance capabilities has been rapidly increasing over the past decade. More recently it has been recognized that these data collection systems can be used as additional sources of mobility data, augmenting traditional data sources such as relatively sparsely distributed permanent count stations and supplementary 48-hour volume counts. Most of the research conducted to date has focused on data archiving systems that have freeway system data collection equipment as their primary data source. However, little is known about the feasibility of using advanced signal control systems (ASCSs) as sources of mobility data. A possible cause for the relative inattention to the use of advanced signal system equipment is that using an intersection as a source for road segment volume counts is contrary to conventional traffic data collection principles, which stipulate that road segment volume counts be taken outside the influence area of intersections.

The purpose of this research was to determine if data collected from an advanced arterial signal control system could be used to generate information that would be useful for transportation engineering analyses other than signal optimization and control. This research also looked at some of the technical challenges and limitations to using data collected by an arterial signal control system and presents an analysis of the validity of the data. The premise investigated is that data from signal control system surveillance equipment can be used to calculate daily volume counts for a roadway segment. This premise was tested by screening and aggregating data from signal system data collection equipment and comparing it with data from traffic monitoring equipment located in close proximity to the intersections being analyzed.

The results of this research support the conclusion that reasonable volume estimates can be generated from system detectors located upstream of the intersection stop bars if the system detectors are deployed on all major approach through lanes. This research also demonstrated that the utility of ITS data requires more than simply the deployment of ASCS or other ITS data collection equipment In order for an ITS archived data management system to be successfully implemented, the data product needs of the end users must be considered in the design and deployment of the traffic monitoring and control system, as well as the data management system.

17 Key Words		18. Distribution Statement	
Intelligent Transportation Systems (ITS) Traffic Counts Advanced Signal Control Systems Archived Data Management System		No restrictions. This document is available to the public through NTIS, Springfield, VA 22161.	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 29	22. Price

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

FINAL CONTRACT REPORT

INCORPORATING ADVANCED SIGNAL CONTROL SYSTEMS INTO AN ARCHIVED DATA USER SERVICE PROGRAM

Brian L. Smith, Ph.D.
Assistant Professor
Department of Civil Engineering
University of Virginia

Matthew Christian Grimes
Graduate Research Assistant

Contract Research Sponsored by
Virginia Transportation Research Council

Virginia Transportation Research Council
(A Cooperative Organization Sponsored Jointly by the
Virginia Department of Transportation and
the University of Virginia)

Charlottesville, Virginia

October 2002
VTRC 03-CR7

NOTICE

The project that is the subject of this report was done under contract for the Virginia Department of Transportation, Virginia Transportation Research Council. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Virginia Department of Transportation, the Commonwealth Transportation Board, or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

Each contract report is peer reviewed and accepted for publication by Research Council staff with expertise in related technical areas. Final editing and proofreading of the report are performed by the contractor.

ABSTRACT

The deployment of Intelligent Transportation Systems that incorporate some type of data collection or traffic surveillance capabilities has been rapidly increasing over the past decade. More recently it has been recognized that these data collection systems can be used as additional sources of mobility data, augmenting traditional data sources such as relatively sparsely distributed permanent count stations and supplementary 48-hour volume counts. Most of the research conducted to date has focused on data archiving systems that have freeway system data collection equipment as their primary data source. However, little is known about the feasibility of using advanced signal control systems (ASCSs) as sources of mobility data. A possible cause for the relative inattention to the use of advanced signal system equipment is that using an intersection as a source for road segment volume counts is contrary to conventional traffic data collection principles, which stipulate that road segment volume counts be taken outside the influence area of intersections.

The purpose of this research was to determine if data collected from an advanced arterial signal control system could be used to generate information that would be useful for transportation engineering analyses other than signal optimization and control. This research also looked at some of the technical challenges and limitations to using data collected by an arterial signal control system and present an analysis of the validity of the data. The premise investigated is that data from signal control system surveillance equipment can be used to calculate daily volume counts for a roadway segment. This premise was tested by screening and aggregating data from signal system data collection equipment and comparing it with data from traffic monitoring equipment located in close proximity to the intersections being analyzed.

The results of this research support the conclusion that reasonable volume estimates can be generated from system detectors located upstream of the intersection stop bars if the system detectors are deployed on all major approach through lanes. This research also demonstrated that the utility of ITS data requires more than simply the deployment of ASCS or other ITS data collection equipment In order for an ITS archived data management system to be successfully implemented, the data product needs of the end users must be considered in the design and deployment of the traffic monitoring and control system, as well as the data management system.

FINAL CONTRACT REPORT

INCORPORATING ADVANCED SIGNAL CONTROL SYSTEMS INTO AN ARCHIVED DATA USER SERVICE PROGRAM

Brian L. Smith, Ph.D.
Assistant Professor
Department of Civil Engineering
University of Virginia

Matthew Christian Grimes
Graduate Research Assistant

INTRODUCTION

Over the past two decades, a large number of Advanced Traffic Management Systems (ATMS) and Advanced Traveler Information Systems (ATIS) have been developed and deployed throughout the United States and abroad. An estimated $209 billion will be invested in Intelligent Transportation Systems (ITS) programs between now and the year 2011 (ITSA, 2001). The ATMS and ATIS systems have been deployed at the local, statewide, and multi-state corridor levels and have been in use long enough to demonstrate their effectiveness at managing traffic and providing travelers with information about traffic conditions (Mitretek, 2000). ATMS and ATIS systems, typically use a variety of equipment such as cameras, automatic vehicle detection systems and inductive loop detectors to collect the data used for system monitoring, incident management, and traveler information. This information is usually relayed to a Traffic Management Center (TMC) where transportation agency personnel monitor and manage the transportation system.

In general, TMC staffs are focused on their core mission of transportation system efficiency and lack the resources and expertise to manage the data they collect such that they can be accessed and used effectively for purposes other than real time operations (Winick, 2000). Often TMCs either do not archive data or store data only for a short period of time. Surveys by Texas Transportation Institute researchers indicate that 20 percent of TMCs around the nation were not archiving data at all, and there was no consistency or data archiving standards among the others (Turner, 1997).

There are ongoing efforts to address this issue, and to this end the U.S. Department of Transportation Joint Program Office for Intelligent Transportation Systems (JPO) has developed a strategic plan to broaden the use of archived data collected by ITS equipment. The objectives of this plan can be summarized as follows (USDOT, 2001): to integrate ITS derived data with other data sources; make it more accessible to end users; and perhaps most important, to transform the raw data into information useful to practitioners in many areas of the transportation profession.

Most of the efforts to date toward the goal of integrating ITS data into the broader traffic data environment have been focused on data collected by freeway management systems. However, advanced signal control systems that use traffic surveillance equipment are being deployed with increasing frequency. As with the freeway ATMSs, the data from signal control systems can be archived and used for purposes other than signal control; however, the integration of signal control systems and data management systems is rarely implemented. The need for more research on coordination between signal control systems and archive data management systems was addressed in the *ITS Benefits: Data Needs Update 2000* report published by the JPO reference. The report presents the areas where additional information about the benefits of ITS services is desired and was derived from an ITS America workshop on data needs. Of the 19 research needs identified by the data needs task force, Information Management – Data Archiving was ranked second, with incident management being only a slightly higher priority (Mitretek, 2001).

ITS deployment studies conducted by the JPO indicate that a significant percentage of transportation agencies have deployed advanced signal control systems with data collection capabilities, and the use of such technology is projected to increase (USDOT, 2001). Typically these systems collect data such as traffic volumes and speeds using inductive loop detectors. In cases where the data are stored for a period of time, seldom are the data screened for validity or aggregated and transformed into information such as daily traffic counts, annual average daily traffic (AADT), or other measures. This would be advantageous in that departments of transportation (DOTs) traditionally deploy and maintain a separate infrastructure, usually referred to as traffic monitoring systems, to collect AADT and other basic measures. Despite the relative lack of attention to date, the increasing deployment of advanced signal control systems suggests that a significant potential exists for advanced signal control systems to be used as a source of traffic information. This research assesses this potential through a review of published literature concerning the use of ITS data as a resource for transportation engineering information and a case study of the advanced signal control system deployed by the Virginia Department of Transportation (VDOT) in the Northern Virginia region.

PURPOSE

The purpose of this research was to determine if data collected from an advanced arterial signal control system could be used to generate information that would be useful for transportation engineering analyses other than signal optimization and control. The premise is that data from signal control system surveillance equipment can be used to calculate daily volume counts for a roadway segment if the surveillance equipment is deployed in an appropriate configuration. This premise is tested by screening and aggregating data from signal system data collection equipment and comparing it with data from traditional traffic monitoring equipment located in close proximity to the intersections being analyzed.

Thus, the specific goal of this research was to determine if traffic surveillance equipment deployed in a signal control system could be used to estimate daily traffic volume counts. This research looked at some of the technical challenges and limitations of using data collected by an arterial signal control system and present an analysis of the validity of the data. The exploration of these issues involving intelligent signal systems as a data resource will help to address one of the top research needs identified in the report published by the JPO reference.

SCOPE

The signal control system used in this analysis is deployed in the VDOT Northern Virginia District, which includes the counties of Arlington, Fairfax, Loudoun, and Prince William and is referred to as the Northern Virginia Smart Traffic Signal System (NVSTSS). Arlington County, unlike the other counties in the Northern Virginia District, is responsible for their street system. All data collected by "system" detectors of the NVSTSS (system detectors are installed upstream of typical intersection queues to collect basic system demand and performance data) are stored in an archival database at the University of Virginia and the Virginia Transportation Research Council's Smart Travel Laboratory (STL). The data analyzed in the NVSTSS archival database were detector volume, speed, and occupancy, and the daily volume counts from the intersections were compared to daily volume counts collected from traffic monitoring equipment used by VDOT's Traffic Monitoring System (TMS). The TMS uses a network of permanently installed surveillance equipment (permanent count stations) and regular temporary counts collected with portable equipment to produce estimates of traffic volumes on all roads maintained by VDOT.

METHODOLOGY

The following methodology guided the research effort.

1. *Literature review.* A review of the literature was conducted to provide a foundation for the research. The literature review focused on traditional traffic data collection programs, the ITS archived data user service, and advanced data management practices.

2. *NVSTSS data preparation.* This task involved preparing the system detector data in the NVSTSS archival database for comparison with daily volume counts collected by the TMS permanent count stations. The design (usually referred to as the "schema") of the NVSTSS archival database was initially intended only to support long-term storage of the data, and, as such, extracting useful information from it required complex, SQL queries that often took hours to run. The schema was altered by screening the system detector data for validity and screening the sites to determine which ones had system detectors in an appropriate configuration for daily volume analysis and then storing the results of this data screening in additional fields and tables. Tables were also created to store aggregate information, in preparation for Task 3.

Detector data validity. The data files that the STL receives from the NVSTSS contain 15-minute sums of volume, 15-minute averages of speed, and 15-minute averages of loop occupancy for each loop detector in the STSS. These data files can contain invalid records due to equipment or communications failures, which are sometimes manifested by values of negative 1. More commonly, however, the invalid records are not so easily identified, and a series of screening tests must be applied to test the validity of the data.

The data screening tests used in this task are based on either threshold values, traffic flow theory, or a combination thereof, and are applied with SQL commands to the database. The formulae for the screening tests are as follows (Hauser, 2001):

Prescreening tests:
1.Volume AND Occupancy AND Speed >= 0
2.Volume < 3100 AND Occupancy < 100
3.Volume >= Occupancy

Feasible Volumes:

4.IF Occupancy = 0 OR 1	THEN Volume < 580
5.IF 1 < Occupancy <= 15	THEN 1 < Volume < 1400
6.IF 15 < Occupancy < 25	THEN 180 < Volume < 2000
7.IF Occupancy >= 25	THEN Volume > 500

The first test screens out records that have negative values; typically negative values indicate some kind of equipment or communications error. The second test ensures that volumes and occupancies will not exceed maximum thresholds of 3100 vehicles per hour and 100%, respectively. Tests four through seven screen out records for which the combination of volume and occupancy is inconsistent with traffic flow theory relationships of volume and density. Essentially these tests say that the volumes recorded must be consistent with the density measured by the occupancy recording.

Detector configuration. The second step in data preparation task was to identify the sites that were appropriate for analysis. There are nearly 1000 intersections in the NVSTSS, and the STL archival database receives data from 438 of them. Currently, only data from system detectors are transferred to and archived by the STL, which is why the STL does not receive data from all of the intersections controlled by STSS. Among the sites that have system detectors, the number and location of the system detectors installed vary considerably. Some sites have system detectors at only some of the through lanes on one or two of the intersection approaches. Other intersections have system detectors on all the approaches, and a few even have system detectors in left-turn lanes.

In order for an intersection to be useful for the traffic volume studies described in this research, the system detectors must be installed on every through lane on the route or street for which the volume count is desired. For example, if there are four through lanes on the major approach, there must be four system detectors installed on the major street approaches in order for all major approach through traffic to be counted. If there are system detectors on the minor approaches, the traffic on the minor street approaches can also be counted. Figure 1 illustrates an intersection with system detectors on all approaches. It is rarely the case that all major and minor approach through lanes of an intersection have system detectors installed.

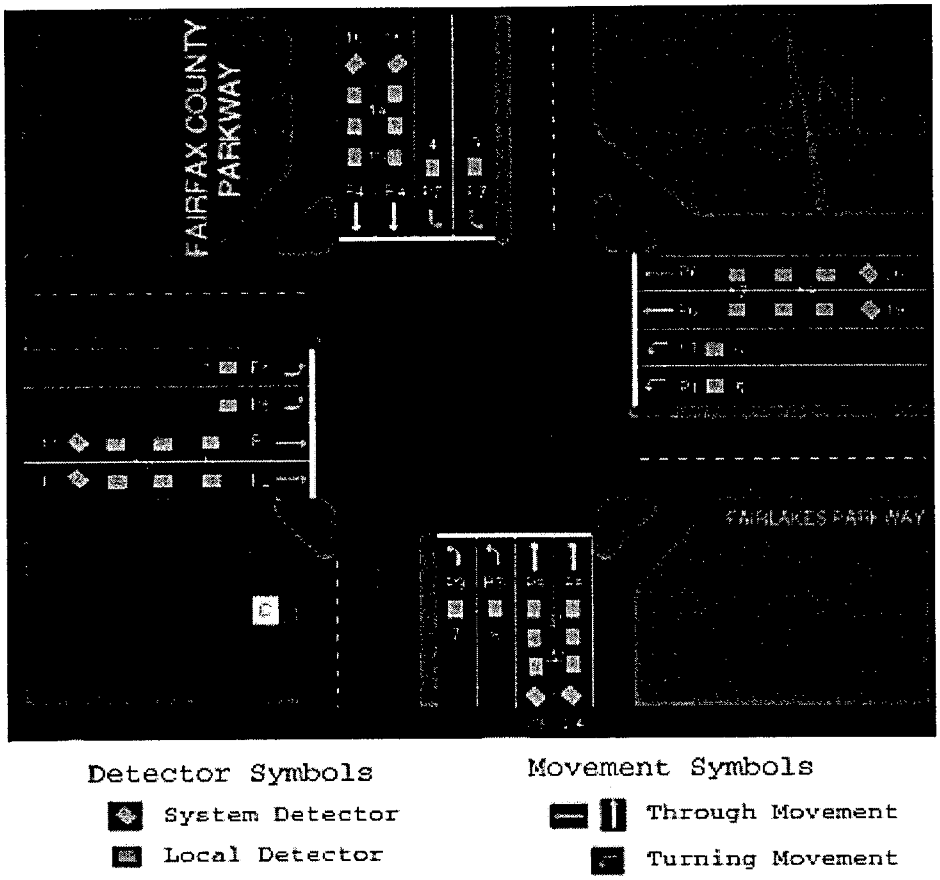

Figure 1: An intersection with system detectors on all approach through lanes.

Figure 2 illustrates a more common configuration, where only the major street approaches have system detectors. The percentage of intersections with all main street approaches instrumented is still relatively low, and most intersections have some combination of major and minor lanes with system detectors, as illustrated in Figure 3. In order to simplify the queries used in Task 3, data aggregation, a new table called "good sites" was created which contains the site identification numbers of all the intersections that have a system detector installed on each major approach through movement.

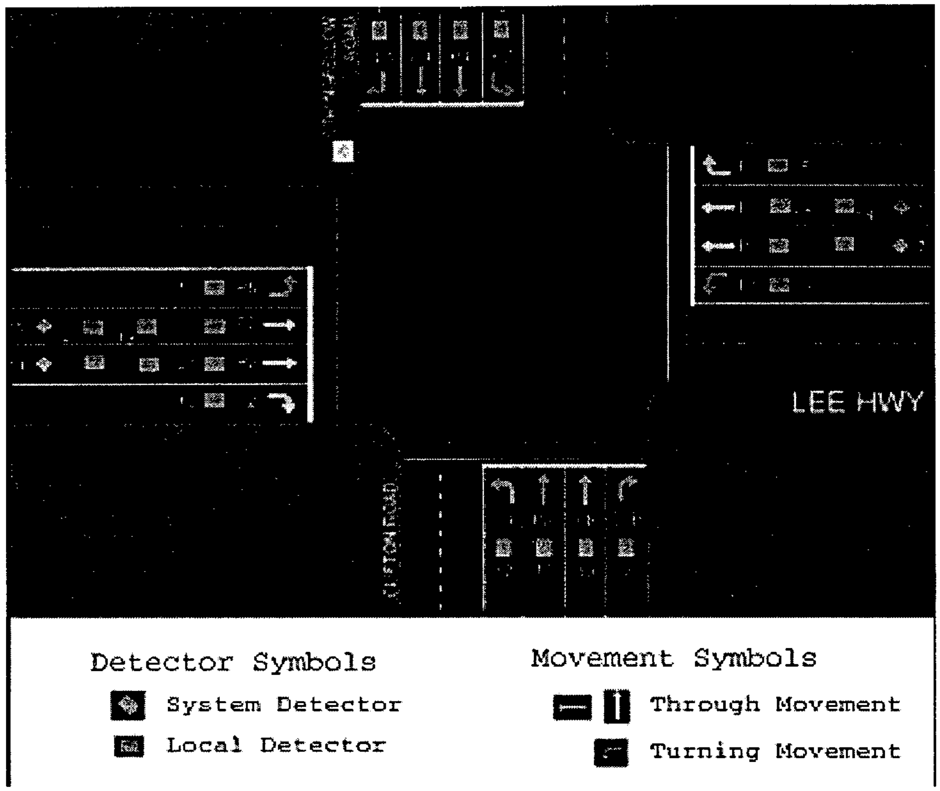

Figure 2: An intersection with system detectors only on the major approach through lanes.

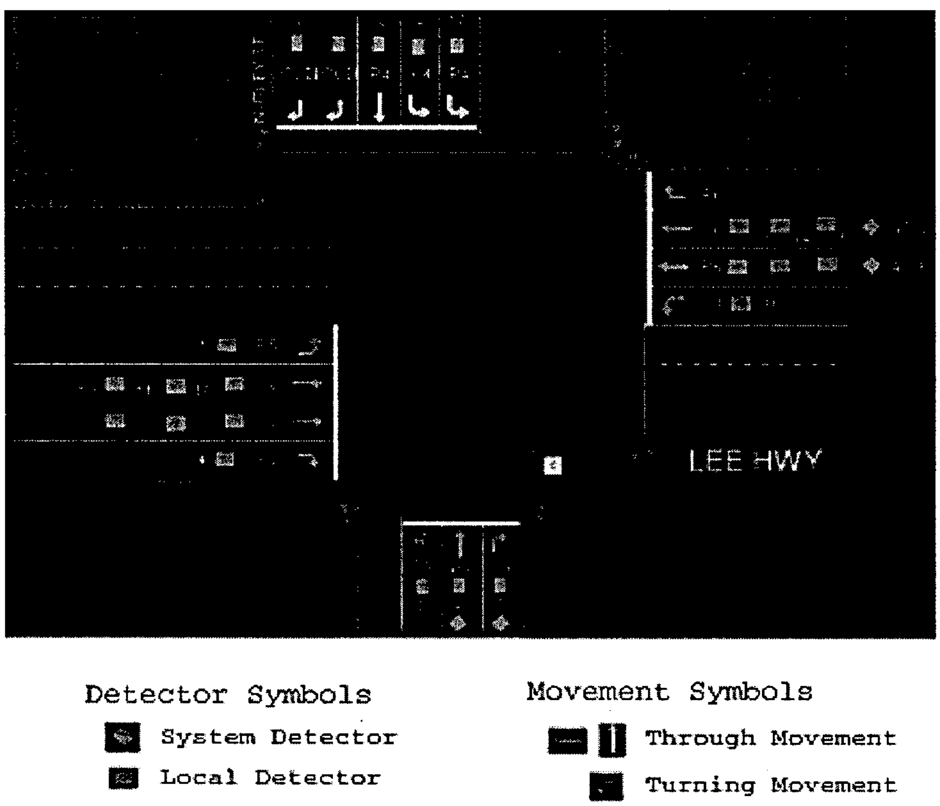

Figure 3: A typical system detector configuration.

3. *Data aggregation.* Daily volume counts are the basis for AADT counts and many transportation performance measures. This task produced daily volume counts using archived data from the NVSTSS. In order for daily major approach volume counts at an intersection to be accurate, all of the collection intervals for all of the major approach system detectors must contain valid records. Since the collection interval is 15 minutes, every major approach through movement system detector must have 96 valid records in the "detector data" table. If some of the intervals contain invalid data, or the detector does not report any data, the volume count will be erroneously low. The aggregate table "site daily vol.", which contains daily volume counts for the major approaches of the intersections in the table "good sites," was created with the following procedure:

- The major approach through movement system detectors at the intersection were identified.
- For major approach through movement detector for each day, the number of valid records was determined.
- If all detectors reported 96 valid records for a day, the volumes for each time period for each detector were summed.
- The total volume, the day it was recorded, and the siteid was added to the "site_daily_vol" table.

4. *Acquisition of baseline data.* After the system detector data were aggregated into daily volume counts for each intersection, a baseline data source was needed with which to compare the signal system daily volume counts. VDOT's network of permanent count stations, the TMS described in a previous section, was the ideal choice for this level of analysis because data were already collected and archived. Having the archived data available meant that approximately 400 daily volume counts could be compared on a day-for-day basis between the TMS and NVSTSS. Furthermore, because data from the TMS are used to calculate the officially published traffic volumes on VDOT highways, the TMS data were considered to be the "correct" volume against which to compare the daily volumes calculated from the signal system data.

5. *Location selection and evaluation.* The final step in preparing the data for comparison of daily volumes was to select locations from both data sets where valid comparisons could be made. In addition, the road and intersection geometry of these locations was examined in more detail.

Site Identification. The locations of all the NVSTSS controlled intersections and the TMS permanent count stations were obtained from VDOT staff. The data were analyzed with Geographic Information System (GIS) software to identify the permanent count stations that are located less than 1 mile from, and on the same route as, an intersection that has an appropriate system detector configuration.

Geometric evaluation. The threshold distance of 1 mile was used based on an inspection of the data available, and with the expectation that as the distance between the intersection and permanent count station increased, the difference in daily volumes observed would also increase. Another factor that can contribute to differences in daily volume between an intersection and its corresponding permanent count station is the configuration of the system detectors with respect to the intersection geometry. The intersection diagrams from MIST and aerial photography were used to assess how well the data from the system detectors should represent the total volume on the roadway segment.

6. *Daily volume comparison*. The final step in this research was to compare the daily volumes calculated from the NVSTSS system detectors with the daily volumes recorded by the TMS permanent count stations. The TMS and NVSTSS volume counts constitute two observations of the same random variable, the traffic volume on a roadway segment. As stated previously, the TMS data, which are the source of officially published volume counts, were used as the baseline; thus the percent error, D_n, between the two data sets was as follows:

$$D_n = \frac{(X_n - Y_n)}{X_n}$$

where: X_n = TMS volume for day n
Y_n = signal volume for day n

A 99 percent confidence interval was then calculated for the mean of each set of D_n values. The formula for the confidence interval, *CI*, for the mean difference between observations variable as follows:

$$CI = d \pm \frac{(t_{\alpha/2,n-1})(s_d)}{\sqrt{n}}$$

where:

d = mean percent error
$t_{\alpha/2,\, n-1}$ = t distribution statistic for confidence level of $1-\alpha$
s_d = standard deviation of percent error
n = number of samples

The true mean error between observations will lie within the limits of the confidence interval computed from the formula. In other words, one can be 99% confident that the mean difference in daily volumes between the intersection and permanent count station will be within the limits of the confidence interval.

RESULTS

Literature Review

The literature review provided a foundation for the project. Findings from the literature review are summarized here.

Traffic Data Collection Programs

Traditional TMSs, such as the program managed by VDOT's Traffic Engineering Division, have been focused on the collection of data to support non real-time analyses of transportation systems. Examples of such analyses include pavement performance monitoring, safety analyses, evaluations of structure and bridge sufficiency, and long-term highway system performance monitoring for use in transportation planning, land use and economic decision making, and some aspects of environmental quality analyses. These analyses are focused more on annual growth trends over many years, or average seasonal, or day of week variations, rather than real time measures of mobility. The data elements collected have traditionally been traffic volume counts, speeds, vehicle classification counts, and the weight of various types of vehicles on the highway. The vehicle weights and vehicle classification counts are very important in pavement performance monitoring, and the classification counts are needed for air quality, noise, and emissions studies.

The TMS programs usually rely on inductive loop detectors embedded in the pavement. The speed, vehicle classification, and weight data collection requires additional equipment such as additional loops, axle sensors, and/or piezos. Various combinations of these devices that can be installed to produce the requisite data are discussed in the *AASHTO Guidelines for Traffic Data Collection Program.* (AASHTO, 1992). In addition to using different equipment configurations, TMSs typically have deployed permanent data collection stations, often at intervals of tens of miles, and used portable data collection devices such as pneumatic tubes to collected data on road segments between the permanent count stations. The permanent count stations collect and report data continuously. The portable data collection equipment is typically used to collect data over a 48-hour or similar interval, and then the data are retrieved and the data collection equipment is installed at another location. In VDOT, these counts are usually the responsibility of districts. A given location may be counted in this manner every 2 or 3 years, depending on the functional classification of the roadway. The data obtained from the portable equipment are transformed from 48-hour volume counts to AADT estimates by applying seasonal and other adjustment factors that are calculated from data collected by the permanent count stations (Chapparal Corp., 2001).

The *AASHTO Guidelines for Traffic Data Programs* (AASHTO, 1992) provide guidance on the entire traffic data collection process, including assessing data collection needs, equipment deployment, and analysis activities such as editing, summarizing, reporting, and retaining data. The *Guidelines* specify, for example, how and where permanent and portable data collection equipment should be deployed: on straight segments of roadway, away from the influence of intersections and entrances that might affect the stable flow of traffic on the segment of road under observation. The *Guidelines* also define the procedures for aggregating raw traffic data into useful statistics and annual estimates such as AADT. These procedures, attempt to minimize the effect of missing or erroneous data, which according to the *Guidelines,* should never be imputed if the data are to be used to provide information products to the end users of an archived data management system.

Archived Data User Service: National ITS Architecture

The National ITS Architecture is an ITS program development framework intended to guide the implementation of ITS. The National Architecture is subdivided into the following principle components (USDOT, 2001):

- *User Services*, which represent what the system will do from the perspective of the user. A user might be the public or a system operator. In version 3.0 of the National Architecture, there are 31 user services, which are organized into seven user service groups, called "bundles" : Travel and Traffic Management, Public Transportation Management, Electronic Payment, Commercial Vehicle Operations, Emergency Management, Advanced Vehicle Safety Systems, and Information Management.
- The *Logical Architecture*, which is best described as a tool that assists in organizing complex entities and relationships. It focuses on the functional processes and information flows of a system. Developing a logical architecture helps identify the system functions and information flows and guides development of functional requirements for new systems and improvements.
- The *Physical Architecture*, which is the physical (versus functional) view of a system. The physical architecture provides agencies with a physical representation (though not a detailed design) of how the system should provide the functionality defined by the user services. It is divided into subsystems that represent the major physical components of ITS infrastructure, such as roadside data collection equipment, in vehicle sensors and traffic monitoring centers.

The user service that directly addresses the needs of traffic data collection programs is the Archived Data User Service (ADUS). Other user services, such as Advanced Traffic Monitoring Systems (ATMS) and Advanced Signal Control Systems (ASCS) typically collect vast amounts of data. The potential uses of these data, beyond real-time operation and system monitoring, prompted the inclusion of ADUS into the National ITS architecture in order to address information management. ADUS specifies the functions that an ITS archived data management system should provide. The ITS architecture identifies 16 user groups of the ADUS in the broad functional categories of planning, operations, safety, and research.

Typically the functionality specified by the ADUS is provided by the implementation of an archived transportation data management system (ADMS). Version 3.0 of the National ITS Architecture includes Data Management as a subsystem of the Physical Architecture and describes its function as follows (USDOT, 2001):

The Archived Data Management Subsystem collects, archives, manages, and distributes data generated from ITS sources for use in transportation administration, policy evaluation, safety, planning, performance monitoring, program assessment, operations, and research applications. The data received is formatted, tagged with attributes that define the data source, conditions under which it was collected, data transformations, and other information (i.e. meta data) necessary to interpret the data. The subsystem can fuse ITS generated data with data from non-ITS sources and other archives to generate information products utilizing data from multiple functional areas, modes, and jurisdictions. The subsystem prepares data products that can serve as inputs to Federal, State, and local data reporting systems. This subsystem may be implemented in many different ways. It may reside within an operational center and provide focused access to a particular agency's data archives. Alternatively, it may operate as a distinct center that collects data from multiple agencies and sources and provides a general data warehouse service for a region.

Advanced Data Management

Simply collecting traffic data does not fully support transportation analyses. The raw data must be transformed into information that is "integral to the transportation practice" (USDOT, 2001), and this information must be readily accessible to end-users. In order to meet these goals, the data archives that store the raw data and derive information from it must be carefully designed. In order to empower end users to access information when and how they most need it, traditional data archives, i.e., databases that are designed for capture and storage of data, need to be redesigned or modified. The simplest way to do this is to screen, aggregate, and otherwise transform the data into information needed by users. Once this transformation is accomplished, the new information could be stored in additional tables. The advantage to storing the information in new tables is that the screening and aggregation procedures, which can be quite complex and time-consuming, need be performed only once, and users can then access the information directly, without having to perform calculations on the raw data. User empowerment could be further enhanced by developing predefined queries and reports, or ultimately developing a Graphical User Interface (GUI) to the data tables that allow users to access information without having to learn database programming or structured query language (SQL) (Corey, 1998).

A more robust means of delivering vital information derived from raw data to users is to migrate the database to a data warehouse. A data warehouse is an information management system that is specifically designed for decision support, as opposed to data capture or transaction processing (Corey, 1998). Data in a data warehouse may be organized or segregated into data marts, which are specialized subsets of data that support decision makers in a particular user group.

Data warehouses often depart from traditional database designs in that user-defined information needs take precedence over minimizing data storage space and data redundancy. Data warehouses schemas tend to have fewer tables and relationships and typically would only store the raw data and the descriptive data used to aggregate the raw data. To perform aggregation and other analyses, data warehouses make use of On Line Analytical Processing (OLAP) tools that help end users quickly and easily access the information they need. In effect, the OLAP tools eliminate the need to aggregate the raw data and store it in new tables, as described previously (Corey, 1998).

NVSTSS Data Preparation

The design (or schema) of the NVSTSS archival database was modified to support this research. The schema had to be modified because it was initially designed for the primary purpose of data capture and storage. The large amounts of disaggregate data in the database made it difficult to generate aggregate measures such as daily volume counts, let alone traffic statistics and performance measures based on daily volume counts. The revised schema presented is perhaps the most expedient way to facilitate the production of information useful to stakeholders because it involves only creating new tables to store information and running SQL queries once or at scheduled intervals to populate them. The revised NVSTSS schema is shown in Figure 4.

While the database architecture described here can provide end users with information they need, the process of acquiring ITS data, and distributing information derived from it could be further improved by implementing a formal data warehouse. In this type of data management system, "staging areas" for receiving and processing raw data from the TMCs would be separated from the data marts of mobility information. On Line Analytical Processing (OLAP) tools would enable users to quickly and easily analyze the raw data and derive any important information they needed, potentially through a web browser interface.

Rather than running the series of screening tests every time a user needs valid data from the database, a "status" field was added to the data table, "detector data." This field has a value of "1" for every 15 minute record that passes the screening tests described in the methodology. Fields describing the lane and system detector configuration at an intersection were added to the table "site info" to facilitate identification of intersections from which valid daily volume counts can be calculated. The fields Main Lanes and Minor Lanes indicate the number of through lanes on the main and minor approaches of the intersection, respectively. The fields Main Sysdet and Minor_Sysdet indicate the number of main and minor approach though movement system detectors at the intersection, respectively. Using these new fields, the table "good sites" was created, which lists the intersections that have a system detector on each major approach through lane.

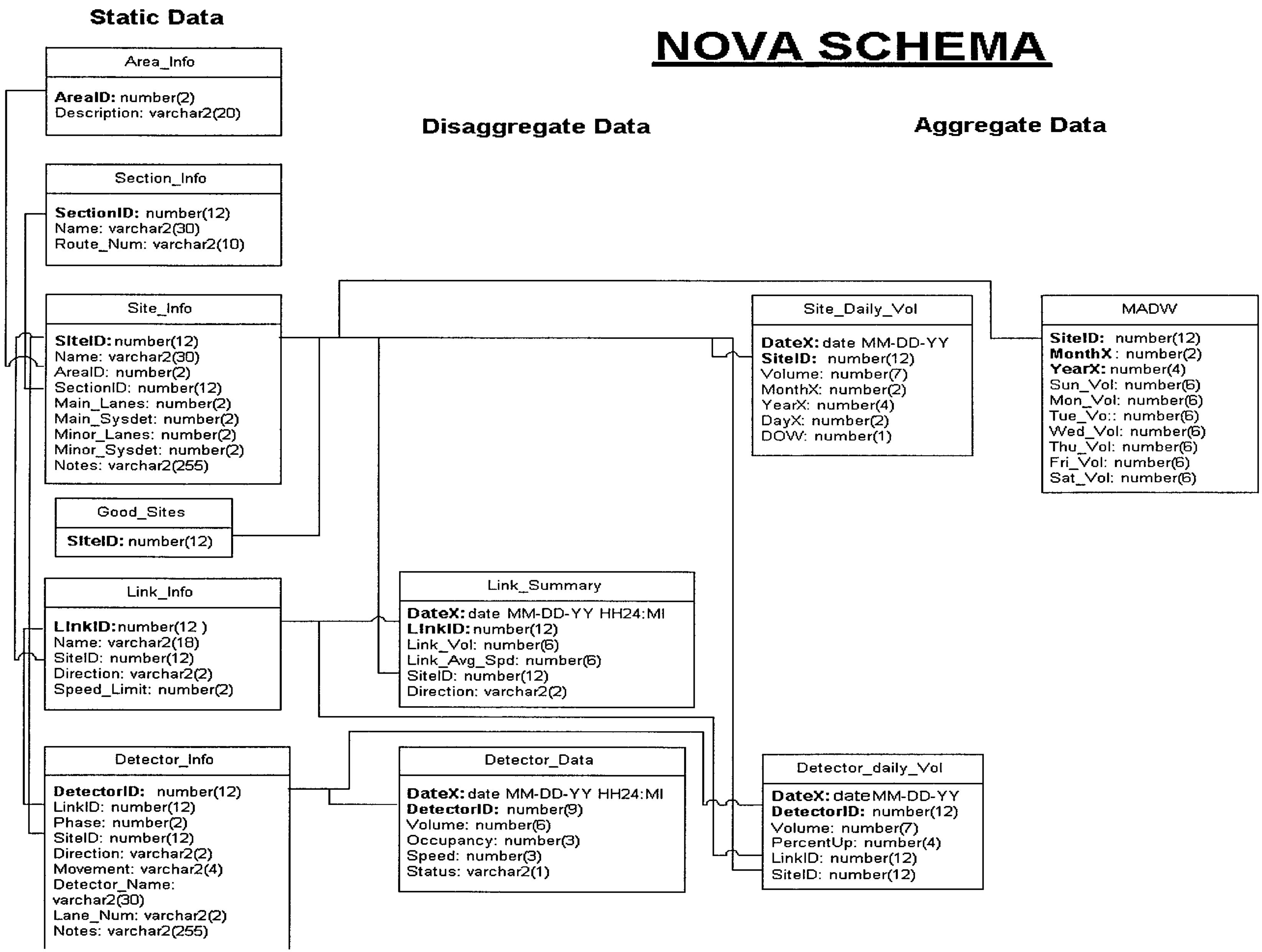

Figure 4: Current NOVA schema.

Other tables were added to store aggregate data and information calculated from the detector data. The tables listed below were added to the schema to facilitate data analysis and distribution. Discussions with VDOT stakeholders indicated that many users are interested in a few common aggregate measures such as AADT and average volumes for a given day of week in any given month. It is more computationally efficient to create tables to store this aggregate data, than to have users issue the queries that perform the aggregation each time they are interested in a traffic statistic such as the AADT at a location.

- Detector_Daily_Vol: This table contains a daily sum of the detector volumes for the detectors listed in table Good Sites. PercentUP is a decimal value of how many valid records are recorded by the detector each day. A valid record is one that has Detector_Data.Status = 1.
- Link_Summary: This table contains link volume and average speed data, one record per link every 15 minutes. Link_Vol is volume in vehicles. Link_Avg_Spd is in mph. These data are calculated for sites in the table Good_Sites, from the data in the table Detector_Data where Detector_Data.Status = 1.
- MADW: This table contains the Monthly Average Day of Week volumes. The data are calculated for the sites listed in the table Good_Sites. Sun_Vol - Sat_Vol are equal to the average value of the sum of the volumes in vehicles, for the respective day in a given month. Assuming at least one valid day (Sunday - Saturday) for each month is available, there are 84 records per site per year. A valid day is one where all main line general purpose system detectors report valid values for the entire day (Detector_Data.Status = 1).
- Site_Daily_Vol: This table contains a daily sum of the main line volumes for the sites listed in the table Good_Sites, from the data in the table Detector_Data where Detector_Data.Status = 1. Monthx, Yearx, Dayx, and DOW are numeric values for month, year and day, and day of week and are included to eliminate the need to perform time-consuming data type conversions when calculating temporal statistics.

Detector Data Validity

The main data table of the archived NVSTSS database, "detector_info" is populated with 15-minute aggregates of system detector volume, speed, and loop occupancy. The database contains data for 216 discontinuous calendar days, from a total of 428 intersections in Northern Virginia. Of the 26,340,524 records in the "detector_info" table, 1,565,533, or 6 percent, fail the screening tests described in the previous chapter. These invalid records are retained in the database for archival purposes and potential future use but are not used in the analyses presented in this report.

Detector Configuration

In order for a daily volume count to be calculated from the intersection system detectors, a system detector must be installed on each through lane of the route for which the volume count is desired; only 162 of the 428 intersections from which the STL receives data have system detectors on each major approach through lane.

Data Aggregation

Daily volume data were aggregated to the "site_daily_vol" table. This table contains data from 162 intersections and 122 discontinuous calendar days. The "site_daily_vol" table has so few intersections and days compared to the "detector data" table because of the screening procedures described. Only intersections that have system detectors on all major street through lanes were included, and furthermore, if any of the detectors at that intersection reported at least one bad record, the volume for that day was not added to the table.

Acquisition of Baseline Data

There are fewer than 15 VDOT-maintained arterial permanent count stations in the Northern Virginia region. GIS analyses reveal that 6 of these are within 1 mile of and on the same route as an intersection where the arrangement of system detectors is such that daily volumes can be estimated. The valid daily volume counts at each of these locations during the year 2000 were downloaded from VDOT's TMS database. These daily volume counts were loaded into a database that also contained daily volume counts from the STL database.

Location Selection and Evaluation

As mentioned, there are six permanent count stations located within 1 mile of and on the same route as a NVSTSS intersection with the proper system detector configuration. One of these six permanent count stations is equidistant from two intersections, which yields a total of seven possible study locations.

All of the intersections analyzed are located in Fairfax County. Routes 1 and 50 are U.S. highways, while Routes 7, 28, and 7100 are part of Virginia's primary road system. The "Distance" and "Number Access Points" columns in Table 1 below, were calculated using GIS maps obtained from VDOT's Northern Virginia District GIS staff. "Number Access Points" indicates the number of access points between the intersection studied and the corresponding permanent count station.

Major Approach Route #	Intersection Name	Distance (mi)	Number Access Points
28	Sully Rd.Eleanor Lawrence Park entrance	0.27	2
7100	Fairfax Co. Pkwy & John Kingman Rd.	0.18	0
1	Richmond Hwy. & Woodlawn Blvd.	0.36	1
7	Leesburg Pike & a commercial entrance	0.65	10
7100	Fairfax Co. Pkwy & Lee Jackson Hwy.	0.36	2
50	Lee Jackson Hwy. & Stringfellow Rd.	0.3	0
7100	Fairfax Co. Pkwy & Fair Lakes Blvd.	0.4	0

Table 1: List of Intersections That Meet the Selection Criteria.

One location pair, the intersection of Lee-Jackson Memorial Highway and Stringfellow Road, could not be analyzed because all the volumes recorded in the TMS database were invalid on the days that the NVSTSS database contained records. The intersection of the Fairfax County Parkway and Fair Lakes Parkway could not be used in the analysis either, because there was only 1 day of valid data in the NVSTSS database for that location.

It is to be expected that the differences in daily volume counts between the intersection and its corresponding permanent count station will increase with the distance and number of access points between them. Errors due to the presence of traffic sources and sinks between the intersection and count station could be positive or negative, depending on whether the net change in volume results in an increase or decrease in the traffic volume at the two locations. Therefore, differences in volume due to traffic sources and sinks are not really errors but rather represent actual differences in traffic volume at the two locations. The configuration of the intersection system detectors is another principal source of error. At some of the intersections studied, the system detector layout is such that large numbers of vehicles will not be counted, since they enter the roadway segment from turning lanes on a minor approach. At these locations, the error should be positive, indicating that the volume calculated from the intersection system detectors is less than the volume counted by the nearby permanent count station. With these concerns in mind, each of the five locations used in this research are evaluated below:

Site 1 - Route 28: Sully Road and Eleanor Lawrence Park

Figures 5 and 6 show that this intersection has only one minor approach, the entrance to a local park. The minor approach volumes are likely to be quite low compared to the major approach volumes; it is therefore unlikely that the configuration of the system detectors will contribute significantly to any error observed. Errors recorded at this location are expected to be the result of actual differences in traffic volumes between this intersection and the count station.

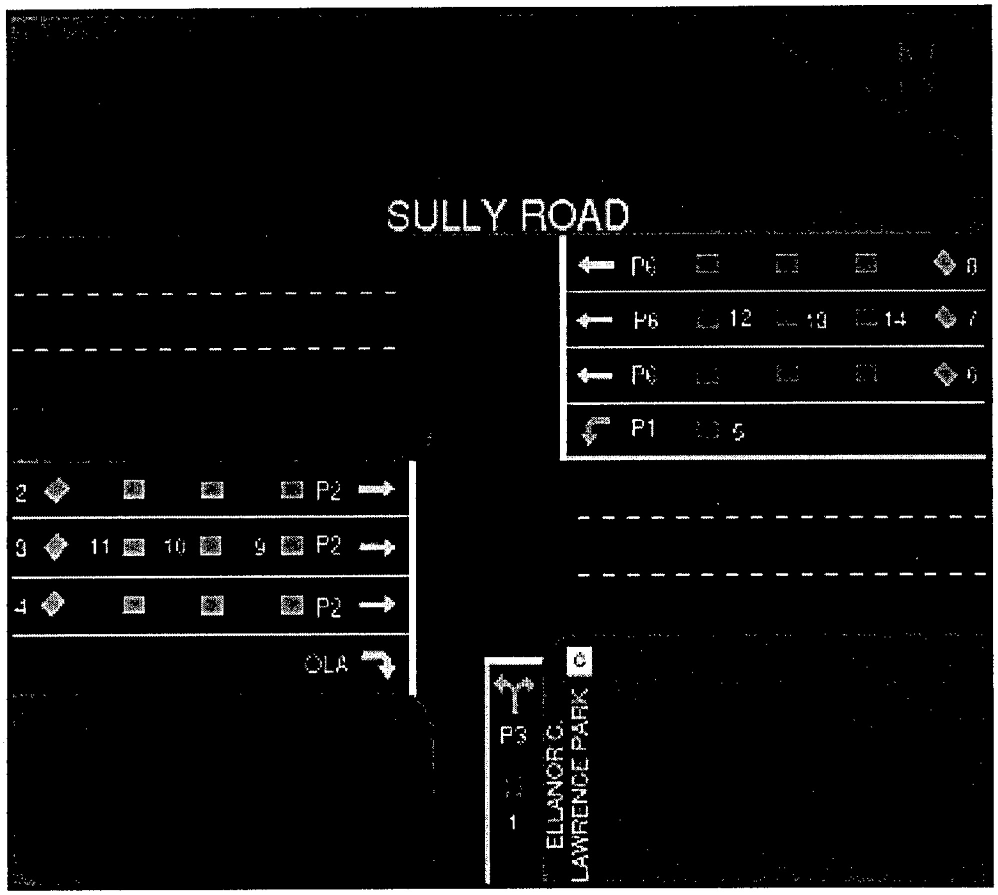

Figure 5: Intersection diagram for Sully Rd & Eleanor Lawrence Park.

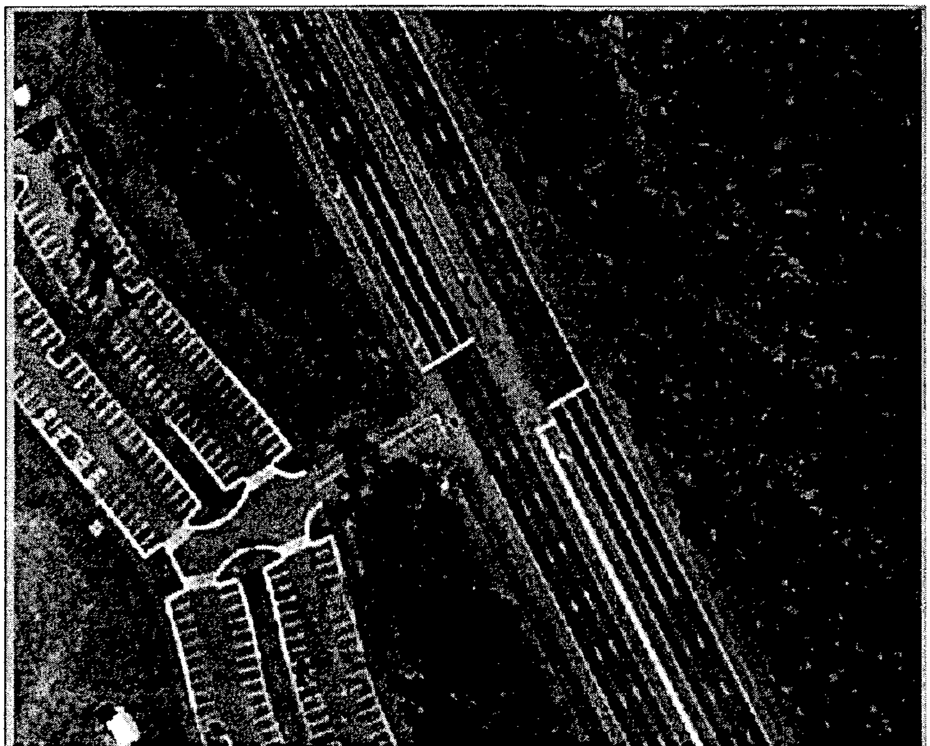

Figure 6: Aerial photograph of Sully Rd and Eleanor Lawrence Park entrance.

Site 2 - Route 7100: Fairfax County Parkway & John Kingman Road.

At this intersection, inspection of the geometry shown in the Figures 7 and 8 reveal that traffic on a total of five major approach lanes, including one double left-turn lane, will not be counted by the system detectors. There are no entrances between this intersection and the corresponding count station, and the distance is the least of all the locations sampled (0.18 mi). It is therefore likely that the difference in volumes at this location is due to the intersection geometry and the arrangement of system detectors. Thus errors are expected to be positive, meaning an undercount due to vehicles entering the roadway segment from lanes without system detectors.

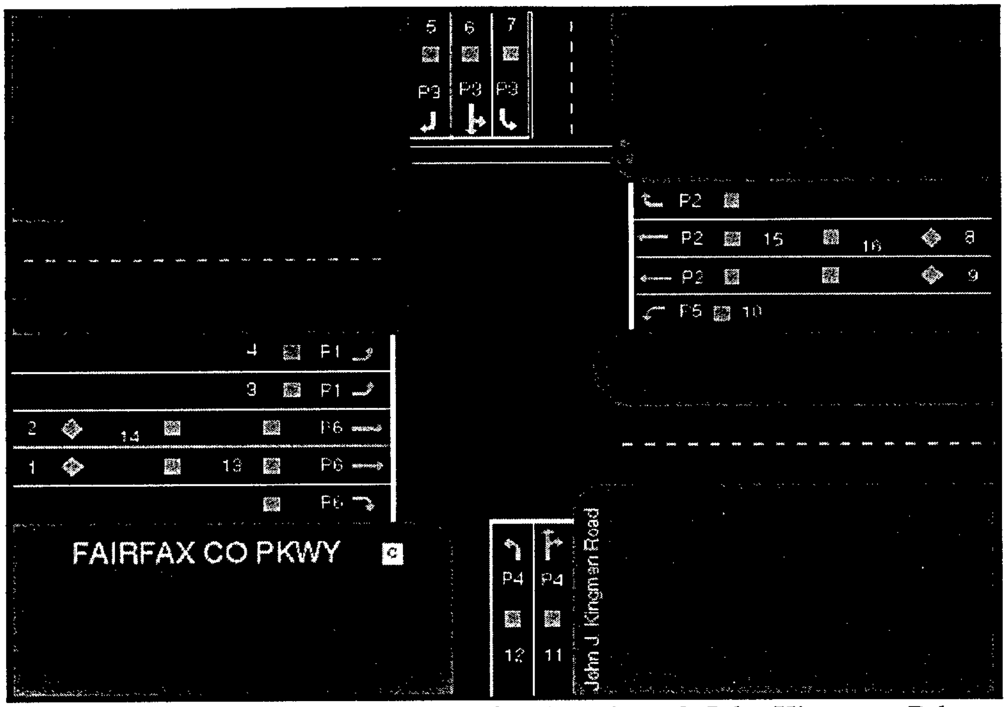

Figure 7: Intersection diagram of Fairfax Co. Pkwy & John Kingman Rd.

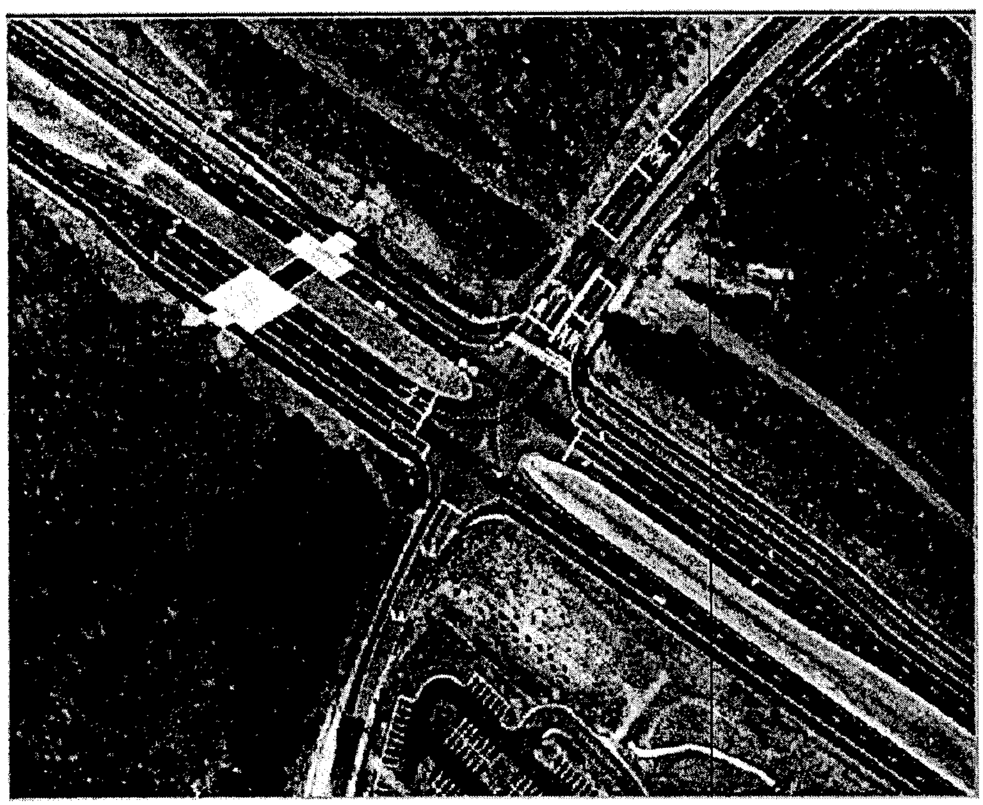
Figure 8: Aerial photographyof Fairfax Co. Pkwy & John Kingman Rd.

Site 3 - Route 1: Richmond Highway & Woodlawn Road.

This intersection, like the intersection of Sully Road and Eleanor Lawrence Park has only one minor approach, which leads to a cultural recreation area, in this case Woodlawn Plantation. It is therefore unlikely that intersection geometry, shown in Figures 9 and 10, will contribute significantly to the error, and errors recorded will more likely be the result of actual differences in traffic volume between the two locations.

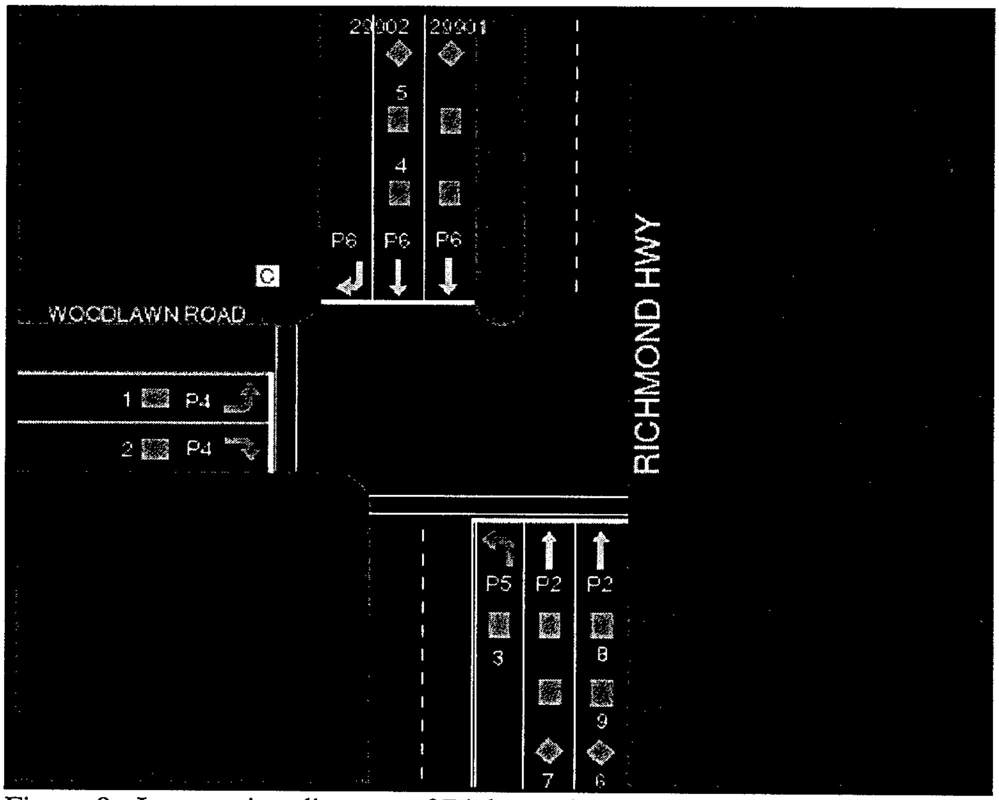

Figure 9: Intersection diagram of Richmond Hwy & Woodlawn Blvd.

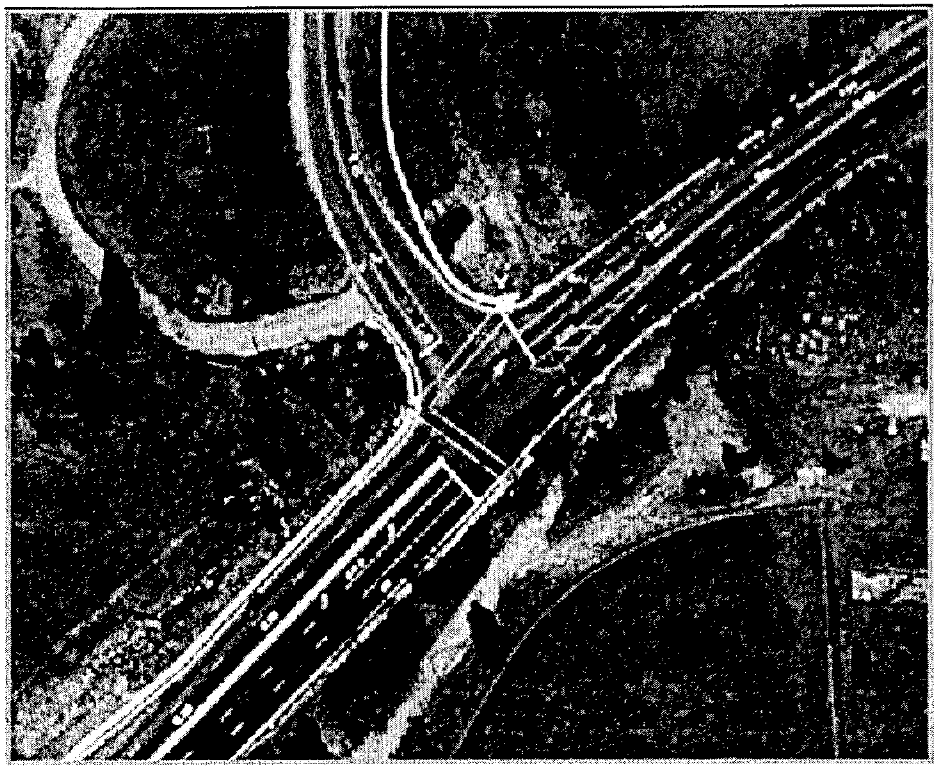
Figure 10: Aerial photograph of Richmond Hwy & Woodlawn Blvd.

Site 4 - Route 7: Leesburg Pike and Tyson's Corner Entrance

The aerial photograph, Figure 12, clearly shows that this intersection is at a major commercial entrance (Tyson's Corner Shopping Center), and the intersection diagram, Figure 11, shows five approach lanes whose traffic would be missed by the system detectors. Errors at this site could be due to the interchange and the 0.65 mile between the intersection and the permanent count station and to the system detector configuration. Errors at this study location are expected to be higher than at the other locations studied in this report.

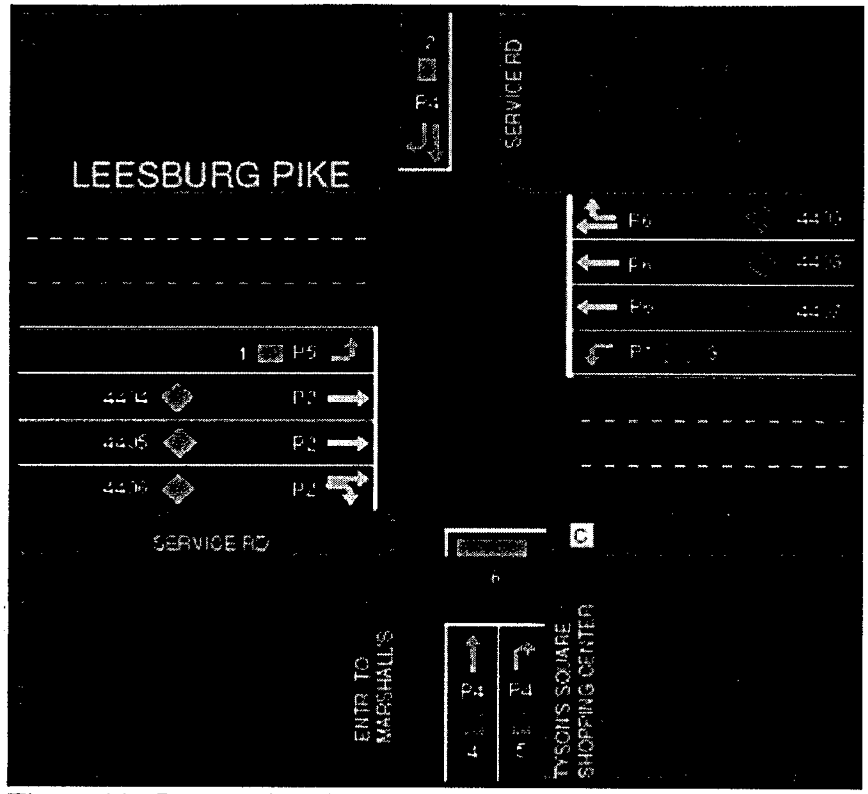

Figure 11: Intersection diagram for Leesburg Pike & a commercial entrance.

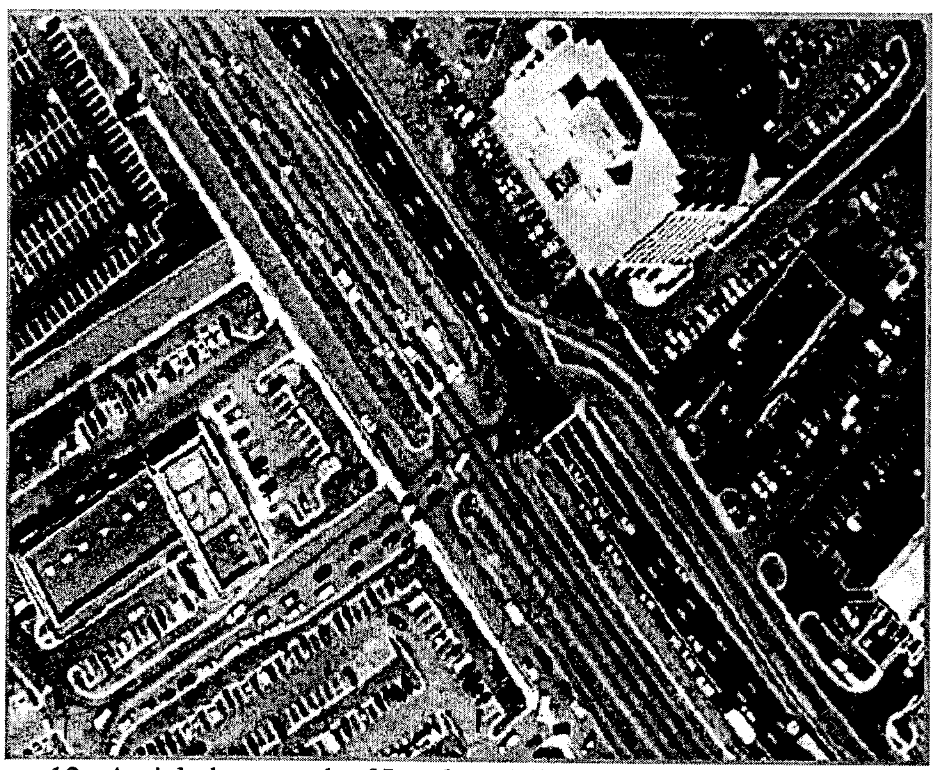

Figure 12: Aerial photograph of Leesburg Pike & a commercial entrance.

Site 5 - Route 7100: Fairfax County. Parkway and Lee Jackson Highway

Inspection of the aerial photograph, Figure 14, shows that this intersection is part of a grade-separated interchange. The intersection diagram, Figure 13, shows that the system detectors count none of the traffic on the ramps. These ramps lead to Lee Jackson Highway, a major arterial. It is therefore likely that the error is due to the system detector configuration and not to the distance or number of entrances between the intersection and the corresponding permanent count station.

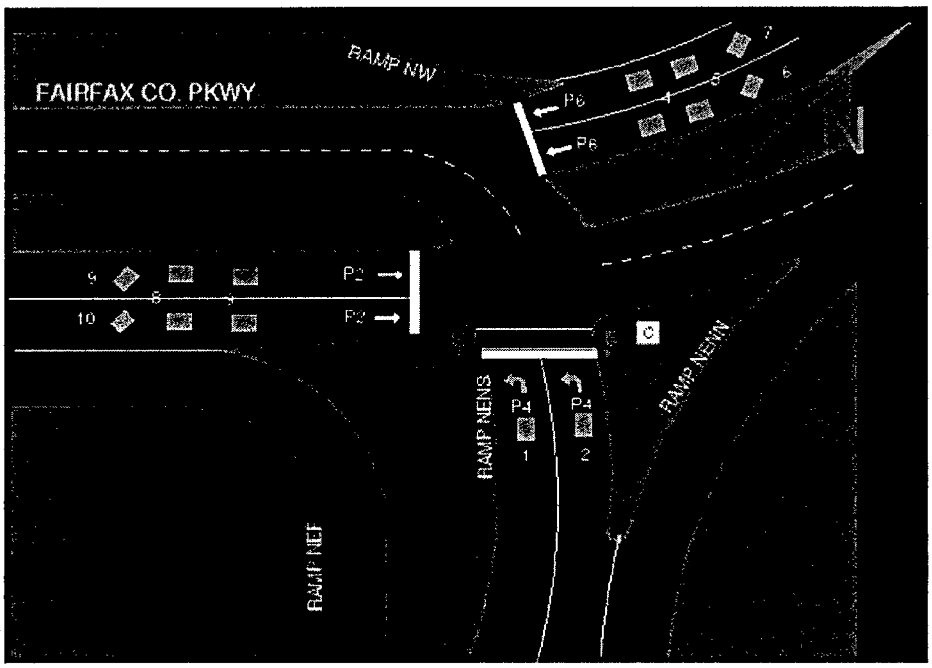

Figure 13: Intersection diagram for Fairfax Co Pkwy & Lee-Jackson Hwy.

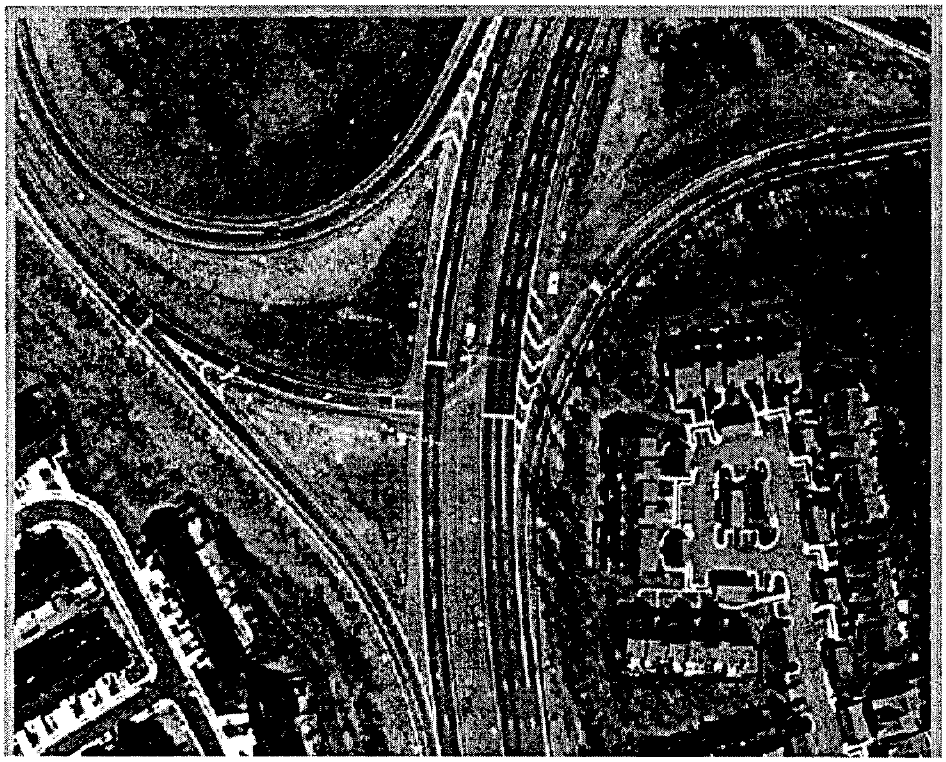

Figure 14: Aerial photograph of Fairfax Co Pkwy & Lee-Jackson Hwy.

Daily Volume Comparisons

Table 2 presents the average daily volumes collected at each of the five sites from the NVSTSS and TMS, the mean percent error, the standard deviation of the percent error, and the 99% confidence interval for the mean errors. Errors that are negative indicate that the NVSTSS volume counts are higher than the TMS volume counts, while positive errors indicate that the NVSTSS count is lower than the adjacent TMS volume count. There is a 99% probability that the "true" mean error will lie between the lower and upper limits of the confidence interval presented in the table. The "# Days" column indicates the number of daily volume counts analyzed.

Table 2: Summary of Results

	Average Volumes		Percent Error			99% Confid. Interval	
Site	TMS	NVSTSS	Mean	Std Dev	# Days	Lower	Upper
1	51675	54124	-4.71	0.93	8	-5.85	-3.56
2	13351	12516	5.84	2.51	93	5.16	6.53
3	31606	34457	-9.07	2.07	106	-9.59	-8.54
4	69564	79311	-13.92	1.63	33	-14.69	-13.14
5	46450	40215	13.98	3.90	35	12.18	15.77

In general, the results from this case study are encouraging. Using NVSTSS detectors, which were designed only to support real-time signal operations, it was possible to generate daily volume counts that were reasonably close to "official" daily volume estimates collected by nearby permanent count stations. Mean absolute errors ranging from 4.71 to 13.98 percent would be acceptable in many traditional transportation applications, such as transportation planning. Furthermore, the error is likely much greater than this at sites that are "monitored" using 48-hour tube counts every 2 to 3 years. Thus, one can conclude from this case study that it is reasonable for transportation agencies to use data from ASCSs to supplement permanent count stations.

Beyond the aggregate "error" measures, it is important to understand what factors result in the differences between the daily volume counts of the permanent count stations and NVSTSS system detectors. Aside from equipment errors, communications problems, and situations that result in inaccurate field measurements recorded in the databases, there are two main variables that are expected to result in the differences. One cause of error is the set of traffic sources (locations that "generate" traffic – such as residential areas) and sinks (locations that "attract" traffic such as shopping centers) that may exist between the permanent count station and the NVSTSS system detectors. It is reasonable to assume that the distance between the NVSTSS system detectors and count station may serve as an effective proxy measure for traffic sources and sinks. It is expected that the magnitude of errors due to traffic sources and sinks between the intersection and the corresponding county station will increase with distance. The advantage of using distance, rather than counting entrances and intersections on a map, lies in simplicity of measurement and the fact that the map may not represent all access points such as driveways and commercial entrances.

The second contributing factor to the differences in traffic volumes is the layout of the NVSTSS system detectors and the intersection geometry. As discussed in the previous section, the placement of signal system detectors resulting from the original system design (which was intended only to support signal operations) may preclude the measurement of certain traffic movements. Because only major approach system detector data were available, the intersection itself represents a potential traffic source or sink because the systems detectors are installed upstream of the intersections. Vehicles turning onto the major street from the minor streets enter the segment downstream of the detectors; thus they are not counted. Since system detectors are located downstream of the taper point of turning lanes, vehicles that are turning off the major approach are also not counted. Thus, errors introduced by system detector configuration will generally be positive, signifying an undercount.

Of the five locations analyzed, the volume counts from two sets of NVSTSS system detectors were lower than the volume counts recorded at the nearby permanent count stations. As explained above, this was expected based on the intersection geometry and system detector configuration. An analysis of the arrangement of the system detectors and the intersection geometry at the intersection of Fairfax County Parkway and John Kingman Road (Site 2) and the intersection of Fairfax County Parkway and Lee-Jackson Highway (Site 5) shows that at these sites the difference in volume counts is likely caused by vehicles entering and exiting the roadway segment from lanes that are not equipped with a system detector.

At the other three locations, Sully Road and Eleanor Lawrence Park (Site 1), Richmond Highway and Woodlawn Boulevard (Site 3), and Leesburg Pike and the entrance to Tyson's Corner Shopping Center (Site 4), the volumes reported by the NVSTSS were higher than the volumes reported by the permanent count station. The resulting negative errors are most likely to be caused by vehicles leaving the traffic roadway segment at points located between the intersection and the permanent count stations. The relationship between the percent error and distance for these three locations is shown in Figure 15. Based on this figure, one can conclude that distance is a reasonable indicator of the ability of a set of signal system detectors to collect traffic monitoring data if the detectors are configured properly.

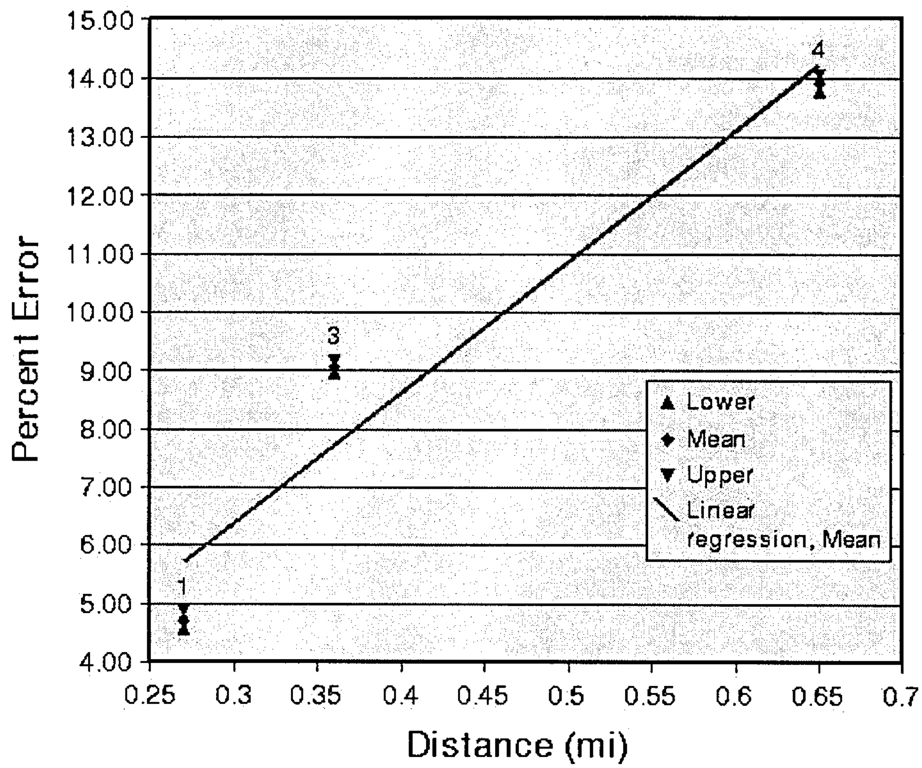

Figure 15: Impact of distance on error.

CONCLUSIONS

1. Upstream signal system detectors can be used to generate accurate daily volume counts. However, as in the case of VDOT's NVSTSS, reliance on upstream system detectors may severely limit the number of intersections that can be used to generate daily volume counts. This is because not all intersections have upstream detectors on every approach through lane. Furthermore, the accuracy of the volumes calculated from upstream system detectors depends upon the configuration of the system detectors with respect to the overall intersection geometry. In cases where the traffic entering or exiting the roadway link from the approaches without system detectors represents a significant portion of the total volume, the volume calculated from the upstream system detectors will be erroneously low. This can

be particularly problematic when the signal is part of an interchange, because the ramps typically do not have system detectors installed on them.

2. If the NVSTSS intersections controlled by the MIST system can be successfully integrated into an ADMS, a substantial savings in data collection costs may be realized by VDOT. Currently there are only 15 arterial permanent count stations in the Northern Virginia District; thus most of the traffic volumes published by VDOT are based on 48-hour traffic volume counts. These counts entail the installation and removal of temporary equipment, which is hazardous and costly. By using the data from the NVSTSS intersections, VDOT may be able to reduce the number of 48-hour counts needed, thus reducing the costs of the traffic monitoring program.

3. Daily volume counts calculated from signal system data collection equipment may be difficult to validate without actually conducting a traffic count immediately adjacent to the intersection under study. This research demonstrates that validating intersection daily volume counts by comparing them to daily volume counts from nearby permanent count stations can be complicated by the fact that the actual traffic volumes at the two locations could be different, depending on how many vehicles enter or exit the roadway between the intersection and the permanent count station. The distance, and the number of access points identified on maps between the intersection and permanent count station, may be used as predictors of the magnitude of difference in traffic volume.

4. This research demonstrates that ADUS functionality does not inherently result from successful advanced signal system deployment. Although advanced signal systems frequently have data collection capabilities, the usefulness of the data depends heavily upon the configuration and deployment of the data collection equipment. Furthermore, this research shows that ADUS functionality requires more extensive deployment of surveillance equipment than is currently the practice in advanced traffic control systems.

5. In addition to the equipment considerations, the data as collected by the signal control system need a great deal of processing to translate it into information such as transportation system performance measures. This processing entails screening the data for validity and aggregating the data spatially and or temporally, as specified by the data product needs of the end users.

6. Transactional databases, while necessary for efficient capture and initial storage of detector data, cannot adequately provide the functionality required by the ADUS. As mentioned above, a significant amount of data processing is required to generate the archived data products such as daily traffic counts, VMT, and other performance measures. To fully achieve the ADUS functionality specified in the National ITS Architecture, a complete archived data management system would have to be deployed. Such a system should have a transactional database to acquire the data from the field equipment, a "staging area" database where the raw data are screened and prepared for inclusion into the final component, the data warehouse. The data warehouse component provides both data analysis functions and serves as the foundation for the user interface.

RECOMMENDATIONS

1. VDOT should actively seek to refine the procedures presented in this report to support traffic counting needs at signal system locations with proper equipment configuration.

2. The needs of ADUS stakeholders should be considered when designing the data collection components of an advanced signal control system. Specifically, the need to deploy additional detectors, beyond what is required for signal control and optimization, should be carefully investigated, and the location of detectors with respect to intersection geometry should also be examined. Ideally, when a new system is being deployed, the design for the detector installation should be such that all traffic volumes on the roadway links can be accurately counted and used in an archived data management system. This may entail the addition of comparatively few detectors, or merely an adjustment to their location with respect to intersection geometry. For existing systems, the feasibility of using additional existing detectors, such as local detectors, should be compared with the cost and utility of installing additional system detectors.

3. A complete data management system, with subsystems for initial, data collection and storage, data screening and validation, and data aggregation and analysis should be developed for each major VDOT traffic control system in order to "maximize the integration of information with other data sources and systems" (USDOT, 2001). For example, an ADMS should have a database for initial data capture and storage, a separate "staging area" database for screening and validating the data, and a data warehouse that stores the validated data. The OLAP and graphical user interface tools of the data warehouse should ensure that the system meets the principles of accessibility and pertinence by providing end users with a graphical tool to analyze the data according to their specific needs.

4. Additional data should be archived from the NVSTSS local detectors. The objective of this additional data storage would be to produce a better volume count by capturing vehicles that are currently not being counted. If local detectors are determined to be sufficiently reliable and feasible to acquire, their use could dramatically expand the number of intersections where daily volume counts can be calculated. In addition, local detector data would enable turning movement counts and volume counts for minor approaches to be calculated.

REFERENCES

American Association of State Highway and Transportation Officials. *AASHTO Guidelines for Traffic Data Programs*. American Association of State Highway and Transportation Officials, Washington, DC, 1992.

Chapparal Corporation. Traffic Statistics 101. www.ntl.bts.gov/tris October 29, 2001.

Corey, M. J., M. Abbey, I. Abramson, and B. Taub. *Oracle8 Data Warehousing*. Osbourne/McGraw-Hill, Berkeley, CA, 1998.

Hauser, T. Ann. *Data Mining Tools for the Support of Traffic Signal Timing Plan Development in Arterial Networks.* Graduate Master's of Science Thesis Presented to the faculty of the School of Engineering and Applied Science of the University of Virginia, May 2001.

ITS America. ITS offers a new approach. www.ista.org. October 29, 2001.

Mitretek Systems Inc. *ITS Benefits: Data Needs Update.* U. S. Department of Transportation Joint Program Office for Intelligent Transportation Systems, Washington, D.C., August 2000.

Turner, S.M, , R. E. Brydia, and J. C. Liu. *ITS Data Management System: Year One Activities.* Texas Transportation Institute, College Station, August 1997.

U.S. Department of Transportation Joint Program Office for Intelligent Transportation Systems. ITS Data Archiving : Five-Year Program Description. www.its.dot.gov October 29, 2001.

U. S. Department of Transportation Joint Program Office for Intelligent Transportation Systems. Metropolitan ITS Deployment Tracking : Extract of Data on Traffic Signals www.its.dot.gov October 29, 2001.

U. S. Department of Transportation Joint Program Office for Intelligent Transportation Systems. *National ITS Architecture ver 3.0.* www.odetics.com/itsarch October 29, 2001.

Winick, R. M. Ph.. Using Operational Data to Improve Information for Planners and Decision Makers. Institute of Transportation Engineers. Meeting and exhibit (2000 : Nashville, Tenn.). Compendium of papers. Institute of Transportation Engineers, Washington, D.C. 2000.