

|  |  |   |  |  |           |
|--|--|---|--|--|-----------|
| 1. Report No.<br>FHWA/TX-02/2127-2   |  | 2. Government Accession No.                         |  | 3. Recipient's Catalog No.   |           |
| 4. Title and Subtitle<br>TXDOT ITS DATA ARCHIVING RESEARCH PROJECT<br>ACTIVITIES   |  |   |  | 5. Report Date<br>September 2001   |           |
|  |  |   |  | 6. Performing Organization Code  |           |
| 7. Author(s)<br>Shawn M. Turner, Scott A. Cooner and Robert J. Benz  |  |   |  | 8. Performing Organization Report No.<br>Report 2127-2                             |           |
| 9. Performing Organization Name and Address<br>Texas Transportation Institute<br>The Texas A&M University System<br>College Station, Texas 77843-3135  |  |   |  | 10. Work Unit No. (TRAIS)  |           |
|  |  |   |  | 11. Contract or Grant No.<br>Project No. 0-2127                                    |           |
| 12. Sponsoring Agency Name and Address<br>Texas Department of Transportation<br>Research and Technology Implementation Office<br>P. O. Box 5080<br>Austin, Texas 78763-5080  |  |   |  | 13. Type of Report and Period Covered<br>Research:<br>September 1999 - August 2001 |           |
|  |  |   |  | 14. Sponsoring Agency Code   |           |
| 15. Supplementary Notes<br>Research performed in cooperation with the Texas Department of Transportation and the U.S. Department of Transportation, Federal Highway Administration.<br>Research Project Title: Developing Guidance for Sharing Archived/Warehoused ITS Data  |  |   |  |  |           |
| 16. Abstract<br>The objective of this project was to assemble guidance for the Texas Department of Transportation (TxDOT) in further developing data archiving systems across the state. This guidance, which is contained in the companion Report 2127-3, includes information on: 1) basic principles of data archiving; 2) determining what data and how much to archive; 3) performing quality control to ensure data meets certain quality requirements; and 4) developing a data archiving system using relevant architecture and standards.<br><br>This report (2127-2) summarizes the research activities conducted during the project. These activities were primarily focused on four basic tasks: 1) stakeholder meetings and interviews to discuss issues and ideas about data archiving; 2) preliminary assessment of archived data needs; 3) analysis of archived data for accuracy and quality; and 4) gathering information on data archiving architecture and standards. The authors recommend that TxDOT begin or continue the dialogue to determine who will be responsible for maintaining ITS data archives at the statewide and district level. The development of one or two modest prototypes of single-source data archives will enable TxDOT to demonstrate proof of concept to system developers as well as potential data users. The authors recommend that TxDOT consider the data archiving guidelines in Report 2127-3, as well as the results from ongoing federal activities, when further developing data archiving systems. |  |   |  |  |           |
| 17. Key Words<br>Intelligent Transportation Systems, Data Archiving, Quality Control, Data Standards   |  |   | 18. Distribution Statement<br>No restrictions. This document is available to the public through NTIS:<br>National Technical Information Service<br>5285 Port Royal Road<br>Springfield, Virginia 22161 |  |           |
| 19. Security Classif.(of this report)<br>Unclassified  |  | 20. Security Classif.(of this page)<br>Unclassified |  | 21. No. of Pages<br>28   | 22. Price |



**TXDOT ITS DATA ARCHIVING  
RESEARCH PROJECT ACTIVITIES**

by

Shawn M. Turner, P.E.  
Assistant Research Engineer  
Texas Transportation Institute

Scott A. Cooner, P.E.  
Assistant Research Engineer  
Texas Transportation Institute

and

Robert J. Benz, P.E.  
Assistant Research Engineer  
Texas Transportation Institute

Report 2127-2  
Project Number 0-2127  
Research Project Title: Developing Guidance for  
Sharing Archived/Warehoused ITS Data

Sponsored by the  
Texas Department of Transportation  
In Cooperation with the  
U.S. Department of Transportation  
Federal Highway Administration

September 2001

TEXAS TRANSPORTATION INSTITUTE  
The Texas A&M University System  
College Station, Texas 77843-3135



## **DISCLAIMER**

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 Texas Department of Transportation (TxDOT) or the Federal Highway Administration (FHWA). This report does not constitute a standard, specification, or regulation. The engineer in charge of the project was Shawn Turner, P.E. #82781.

The United States Government and the state of Texas do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of this report.

## ACKNOWLEDGMENTS

The author wishes to acknowledge the support and guidance of Ms. Regina Flores, current project director from TxDOT's Fort Worth District, and Mr. Abed Abukar, former project director currently with Dallas Area Rapid Transit. Additionally, the author acknowledges the comments and input from other members of the project review panel:

- Ms. Natalie Bettger, North Central Texas Council of Governments
- Ms. Sholeh Karimi, City of Arlington
- Mr. Dan Rocha, North Central Texas Council of Governments

The authors also acknowledge the following persons from TTI, who assisted in the research project at various times and with various tasks:

- Mr. Robert Benz
- Mr. Scott Cooner
- Mr. Jason Crawford
- Mr. William Frawley
- Mr. Daryl Puckett

# TABLE OF CONTENTS

|   | <b>Page</b> |
|---|-------------|
| CHAPTER 1. INTRODUCTION .....   | 1           |
| Introduction .....  | 1           |
| Research Project Objectives .....   | 1           |
| Overview of Report .....  | 2           |
| CHAPTER 2. SUMMARY OF RESEARCH ACTIVITIES AND<br>PRELIMINARY FINDINGS ..... | 3           |
| Stakeholder Meetings and Interviews .....                                   | 3           |
| Research Activities .....   | 3           |
| Preliminary Findings .....  | 4           |
| Preliminary Assessment of Data Needs .....                                  | 6           |
| Research Activities .....   | 6           |
| Preliminary Findings .....  | 9           |
| Analysis of Archived Data Accuracy .....                                    | 10          |
| Research Activities .....   | 10          |
| Preliminary Findings .....  | 10          |
| Information on Data Archiving Architecture and Standards .....              | 16          |
| CHAPTER 3. CONCLUSIONS AND RECOMMENDATIONS .....                            | 17          |
| REFERENCES .....  | 19          |





# CHAPTER 1. INTRODUCTION

## INTRODUCTION

As intelligent transportation system (ITS) applications continue to be deployed in Texas, large quantities of data are becoming available from ITS sensor and detector systems. Traffic management centers collect the data, which typically include traffic conditions, incident/accident data, ITS control responses, and other roadway or environmental characteristics, for use in real-time by TxDOT and partner agencies to manage the transportation system. When these data are archived, they can be used for numerous transportation analyses, such as estimation of ITS benefits, computer model calibration, congestion monitoring, transportation planning, or even pavement design. However, little guidance currently exists for data archiving.

## RESEARCH PROJECT OBJECTIVES

The objective of this research project was to assemble guidance that could help TxDOT in further developing data archiving systems across the state. This guidance includes information on:

- basic principles of data archiving;
- determining what data and how much to archive;
- performing quality control to ensure data meets certain quality requirements; and
- developing a data archiving system using the National ITS Architecture and relevant standards.

The research team conducted several activities to develop the project guidance and recommendations:

- conducted a literature and information search and review;
- met with numerous stakeholders in Texas to discuss issues and ideas about data archiving;
- conducted follow-up meetings and a preliminary data needs survey in the Dallas-Fort Worth area to gather more specific information about what data is being collected and what data is most important to archived data users;
- talked with transportation agencies in other states about data archiving in an attempt to determine what works best;
- participated in a national committee that is currently developing data archiving standards; and
- analyzed archived data to ascertain quality and completeness.

This report (2127-2) summarizes these research activities. More importantly, TxDOT Report 2127-3 documents the results and findings from these activities in “field guide” format that is oriented toward providing guidance to TxDOT and other agencies in further developing data archiving systems. Finally, TxDOT Report 2127-S is a project summary report that provides concise findings and recommendations from this project.

## **OVERVIEW OF REPORT**

This report contains three basic chapters that summarize guidance on data archiving systems. The three chapters are as follows:

- **Chapter 1. Introduction** - provides an introduction to the research project objectives and activities;
- **Chapter 2. Summary of Research Activities** - provides a summary of the research activities that were used in developing guidance on data archiving issues; and
- **Chapter 3. Conclusions and Recommendations** - provides conclusions and recommendations based upon the research project findings.

## **CHAPTER 2. SUMMARY OF RESEARCH ACTIVITIES AND PRELIMINARY FINDINGS**

This chapter provides a summary of the research activities that were conducted for TxDOT Project 0-2127. The activities were primarily focused on four basic tasks, which are discussed in this chapter:

- hold stakeholder meetings and interviews to discuss issues and ideas about data archiving;
- conduct preliminary assessment of archived data needs;
- analyze archived data for accuracy and quality; and
- gather information on data archiving architectures and standards.

### **STAKEHOLDER MEETINGS AND INTERVIEWS**

#### **Research Activities**

The research team held numerous stakeholder meetings and interviews throughout Texas to discuss issues and gather feedback about data archiving and the uses of archived data. Table 1 summarizes these meetings and discussions. We also reviewed the data archiving programs in other states and discussed data archiving issues with these other locations. The findings are discussed on the following pages.

**Table 1. Summary of Texas Stakeholder Meetings.**

| Area or Region    | Agencies Represented   | Date(s)                            |
|-------------------|--|------------------------------------|
| Statewide         | TxDOT Transportation Planning and Programming Division<br>TxDOT Traffic Management Division  | November 1999<br>February 7, 2000  |
| Houston           | TxDOT, Houston District, Operations<br>TxDOT, Houston District, Planning<br>Houston-Galveston Area Council   | January 28 and 31, 2000            |
| Austin            | TxDOT, Austin District, Operations<br>TxDOT, Austin District, Planning   | February 23, 2000<br>July 24, 2001 |
| Dallas-Fort Worth | City of Fort Worth<br>Dallas Area Rapid Transit<br>Burlington Northern Santa Fe Railroad<br>North Central Texas Council of Governments<br>TxDOT, Dallas District, Operations<br>TxDOT, Dallas District, Planning<br>TxDOT, Fort Worth District, Operations<br>TxDOT, Fort Worth District, Planning<br>TxDOT Regional Planning Office | June 14 and 20, 2000               |
| San Antonio       | TxDOT, San Antonio District, Planning<br>VIA Metropolitan Transit<br>San Antonio/Bexar County Metropolitan Planning Organization   | March 19, 2001                     |

**Preliminary Findings**

Several areas in Texas (i.e., Austin, Houston, and San Antonio) currently archive some of the data that are collected by ITS. Fort Worth has plans to archive some of its data in the near future, and Houston and Austin would like to further develop and extend their archiving system capabilities. For these three areas that do archive data, most archive freeway detector data to large or numerous text files in the original field-collected format (i.e., lane-by-lane, 20-second or 1-minute time interval). Other groups (such as planners, designers, and air quality analysts) would like to be able to use these archived data, but are unable to because of size and format. To date, most users have been researchers or other “power users” who have specialized database or programming skills.

In stakeholder and other meetings, we heard concerns about the following issues:

- Who should take responsibility for owning and managing the data archives? Won't data archives be expensive because they require large amounts of computer storage space?
- What ITS data are most important to archive, and at what level? What are the archived data users' requirements?

- How “good” or accurate are the ITS data, and how much quality control is performed?
- Isn't some data collection being duplicated between the planning and operating divisions?

In surveying and talking with people outside of Texas, we found several areas that have developed more sophisticated data archiving systems that enable ordinary computer users to access large databases of archived ITS data. The majority of areas, however, were at a similar early stage of data archiving as TxDOT. That is, many other areas simply log original detector data to a text file and are just beginning a dialogue with the many potential users of the data. We conducted in-depth studies of those areas that have already developed effective data archiving systems, which have these characteristics in common:

- A workgroup or agency has taken or been assigned the responsibility of operating and maintaining the archive.
- Most systems have started as modest prototypes focused on a single source of data, with the most widely archived data being traffic data (i.e., vehicle volume, occupancy, and speed) from detector systems.
- The data archiving system had been developed in a way that permitted ordinary users with typical desktop computers to access and analyze the data.
- The most effective method of data access and distribution was either through the Internet or CD-ROM.
- Original data as collected from the field were saved permanently in off-line storage, but data summaries were made available for most users.
- Quality control methods (most relatively simple, but some more complex) were used to flag or remove suspect or erroneous data from the data archive.
- Adequate documentation on the data archive and the corresponding data collection system was provided.

## **PRELIMINARY ASSESSMENT OF DATA NEEDS**

### **Research Activities**

The research team conducted a preliminary assessment of user requirements in the Dallas-Fort Worth (D-FW) area as an extension of the stakeholder outreach meetings. One of the initial steps in the implementation of an archived data management system is identifying the stakeholders. The research team identified stakeholders using the North Texas Regional ITS Program steering committee roster (<http://www.nortex-its.org/>) and from guidance documents from the National ITS Architecture listing typical stakeholders for the archived data user service (ADUS).

The next step after identifying the stakeholders was performing a data inventory and needs assessment. The ADUS contains 29 data sources and 105 data elements. Researchers identified two objectives to guide the inventory and needs assessment:

1. Inventory ADUS data elements collected via ITS technologies:
  - Now - currently gather with existing field devices and systems.
  - Future - have plans in place to gather in the near term.
2. Determine uses and needs for archived ITS data:
  - Would you use it?
  - What is the priority of need?

The first method of gathering information from stakeholders was via focus groups conducted at the TTI Arlington office. Two different focus groups were conducted but only five agencies (Burlington Northern Sante Fe Railroad, City of Fort Worth, Dallas Area Rapid Transit, North Central Texas Council of Governments, and the Texas Department of Transportation Fort Worth District) were represented. Because of the relatively low turnout and difficulty of scheduling meetings, it was decided that another method of gathering feedback was necessary.

Researchers developed a survey instrument based on the ADUS data sources and data elements. Participants were asked to answer four questions about each of the 105 data elements:

1. Does your workgroup or agency currently collect these data? yes or no
2. Will your workgroup or agency collect these data in the future? yes or no
3. Would your workgroup or agency use this archived data element if it were available? yes, no, or maybe
4. Is this archived data element a priority need for your workgroup or agency? none, low, medium, or high

The survey also allowed participants to add any other data sources or elements generated by their ITS devices that weren't already covered in the ADUS.

The survey instrument was distributed via fax to 25 stakeholder agencies. Sixteen agencies returned the survey for a response rate of 64 percent. The survey participants were as follows:

|                                       |  |
|---------------------------------------|--|
| Burlington Northern Sante Fe Railroad | North Texas Tollway Authority              |
| City of Fort Worth                    | North Central Texas Council of Governments |
| City of Farmers Branch                | TxDOT Regional Planning Office             |
| City of Dallas                        | TxDOT Dallas Advanced Planning             |
| City of Arlington                     | TxDOT Fort Worth Traffic Safety            |
| City of Richardson                    | TxDOT Transportation Planning Division     |
| City of Grand Prairie                 | TxDOT Daltans                              |
| D-FW International Airport            | TxDOT TransVISION                          |

Researchers tabulated the 16 completed surveys. A methodology for assessing and comparing the overall priority level for the 105 data elements was developed. The equation below shows calculation procedures for the overall priority level:

$$\text{Priority Level} = (H \times 3) + (M \times 2) + (L \times 1)$$

Where:

H = number of high priority ratings for the data element

M = number of medium priority ratings for the data element

L = number of low priority ratings for the data element

Example:

Data Element = Freeway Vehicle Volume

$$\text{Priority Level} = (9 \times 3) + (6 \times 2) + (1 \times 1) = 40$$

Table 2 provides the top ten data elements ranked in order of overall priority level. The table also lists the frequency of the agencies surveyed that are collecting these elements now or plan to in the future.

**Table 2. Top 10 Data Elements in the Dallas-Fort Worth Data Archiving Survey.**

| Data Element                                    | Number of Agencies Currently Collecting | Number of Agencies Planning to Collect | Priority Level for Archiving |
|---|---|--|------------------------------|
| <b>1. freeway vehicle volume</b>                | 8                                       | 2                                      | 40                           |
| <b>2. freeway vehicle speed</b>                 | 6                                       | 4                                      | 37                           |
| 3(tie). arterial vehicle volume                 | 7                                       | 1                                      | 33                           |
| <b>3(tie). incident location</b>                | 7                                       | 3                                      | 33                           |
| 5. map display of congestion                    | 1                                       | 7                                      | 31                           |
| 6(tie). stops/delay estimates                   | 0                                       | 7                                      | 29                           |
| 6(tie). arterial level of service               | 2                                       | 4                                      | 29                           |
| <b>6(tie). incident time sequence of events</b> | 4                                       | 5                                      | 29                           |
| <b>9. incident cause</b>                        | 5                                       | 3                                      | 28                           |
| 10(tie). freeway lane occupancy                 | 7                                       | 1                                      | 26                           |
| 10(tie). arterial vehicle speed                 | 4                                       | 5                                      | 26                           |
| <b>10(tie). incident type/severity</b>          | 3                                       | 4                                      | 26                           |
| <b>10(tie). police accident report</b>          | 6                                       | 2                                      | 26                           |
| <b>10(tie). work zone location</b>              | 7                                       | 2                                      | 26                           |

Items in bold were rated as “priority items” for the Houston TranStar data warehouse (1).

The analysis of the survey results shown in Table 2 provides several key observations:

- Freeway vehicle volume and speeds emerged as the two top-ranked data elements by survey participants.
- Map display of congestion, stops/delay estimates, and arterial level of service were all rated highly; however, not many of the surveyed agencies collect these data with their existing ITS technologies.
- Transit, weather, commercial vehicle, and emergency response-related data elements were not rated highly by the survey respondents. This outcome could possibly be explained by the lack of participation of stakeholders in these categories.
- The data elements in boldface type correspond to first-tier priority items identified during the Houston user requirements study performed by the Southwest Research Institute (1).



Overall, the survey results are useful as an inventory of existing and planned ITS data available for archiving in the D-FW region and for identifying an initial priority level of data elements for potential inclusion in a regional data archive.

### **Preliminary Findings**

The potential uses and applications of archived ITS data are as diverse as the data user groups that wish to obtain the data. The data needs for some of these applications are currently fulfilled through the manual collection of traffic data, which often suffers from inadequate breadth or depth. Other data needs are met through estimation and computer simulation techniques, while some data needs simply continue to go unmet.

Given the wide variety of potential archived data users, it is likely that all data elements generated by ITS sources could be useful to other archived data users at some time. In developing data archiving and warehousing systems, though, it may be necessary to prioritize certain data elements by their inherent value to archived data stakeholders. If prioritization of data to be archived is necessary, it should be done by archived data managers in consultation with the archived data users. Past experience has indicated that several types of archived data are valuable to more than one data stakeholder group:

- **traffic condition data:** traffic volumes, vehicle speeds and travel times, vehicle classification, and closed-circuit television (CCTV) images;
- **construction and work zone data:** location, time, date, and extent of blockage/closure;
- **traffic incident logs:** time sequence of events (detection, notification, arrival, and clearance), location, extent/severity, and cause; and
- **traffic control responses:** dynamic message sign (DMS) messages, ramp meter timing, etc.

Because of the detailed nature of ITS traffic condition data (typically collected every 20 to 30 seconds by detectors), data aggregation is often a consideration when archiving ITS data. Aggregation refers to the time interval at which data are summarized. For example, several data archiving systems aggregate 20-second speeds and volumes to 5-minute average speeds and volume subtotals. Aggregation primarily saves computer storage space and reduces data processing time when analyzing or further summarizing archived data. Additionally, aggregation is mostly considered only for traffic condition data (i.e., speed, travel time, volume, and occupancy) from detectors or sensors and not for event-based data such as incident response information.

Aggregation levels used in ITS data archiving systems around the country vary considerably, ranging from saving raw data (20 to 30 seconds) to summarizing data (15 minutes). Transportation planners typically require only 15-minute summaries at most, whereas researchers may require the most detailed data possible for sophisticated analyses. Because of the wide-ranging nature of uses for archived data, in some cases the recommendation for aggregation level

is “save as much as you can afford.” The particular solution for an area will depend upon local capabilities, needs, and resources. To date, the focus in many traffic management centers has been mostly on simple solutions, such as selecting 5- or 15-minute aggregation levels based on existing data needs. Of the traffic management centers that do archive data, several save the original field-collected data in an off-line format, such as compressed files on magnetic tape cartridges or CD-ROMs.

## **ANALYSIS OF ARCHIVED DATA ACCURACY**

### **Research Activities**

In several stakeholder meetings, accuracy was mentioned as a concern when using archived data. In this sense, accuracy refers to the sensor’s ability to truly reflect actual traffic conditions (e.g., reported vehicle counts closely approximate actual number of vehicles).

A ground truth method has been used in several accuracy assessments of archived operations data (2,3,4). These accuracy assessments have shown varying results. In San Antonio, one detector location had vehicle counts within  $\pm 3$  percent of ground truth, whereas another detector location had vehicle counts that range from +20 percent to -38 percent of ground truth. Similar findings have been made in Atlanta, Orlando, and New York.

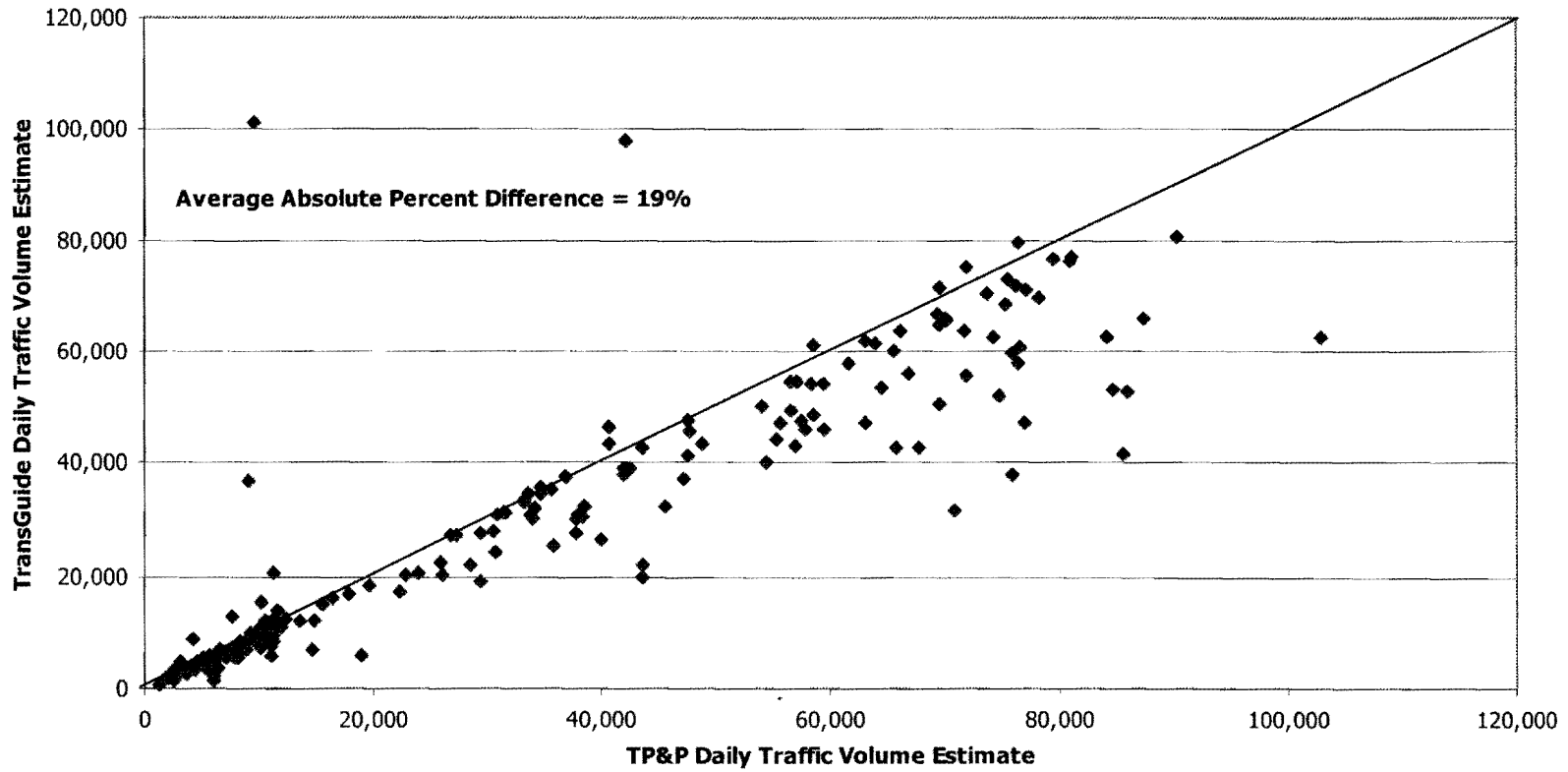
The research team made several similar comparisons between traffic volume data collected by traffic management centers and planning groups or other data sources. In San Antonio, traffic volumes from the TransGuide® traffic management center were compared to volumes collected by TxDOT’s Transportation Planning and Programming (TP&P) Division. In this instance, ground truth for the traffic volumes was not known and researchers directly compared the traffic volume counts from the two sources. In Houston, traffic volumes from the TranStar traffic management center were compared to ground truth counts obtained by manually reviewing videotape.

### **Preliminary Findings**

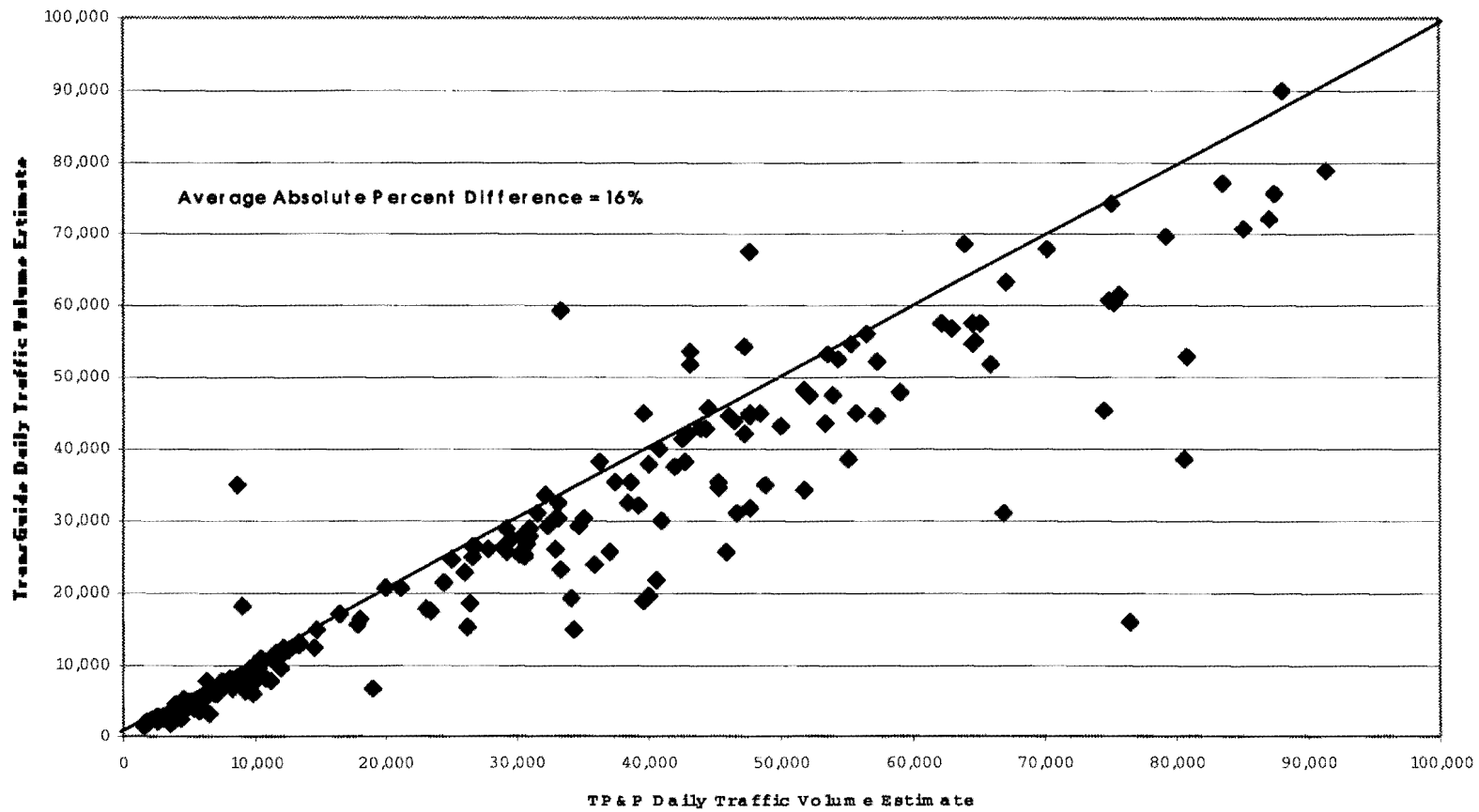
Figures 1 and 2 show the results of traffic volume comparisons in San Antonio for 1998 and 1999. For example, Figure 1 shows a single daily traffic volume count collected by TP&P in 1998 as compared to the equivalent daily traffic volume count from TransGuide data. The TP&P traffic counts, which are collected using pneumatic tubes, are used as the basis for average daily traffic volume estimates. The comparisons for both 1998 and 1999 show consistent results, i.e., the TP&P traffic volumes were on average about 16 percent less than traffic volumes collected by TransGuide.

Figures 3 and 4 show the results of traffic volume comparisons in Houston. In this comparison, the TranStar traffic counts from several different sets of loop detectors are compared to multiple manual counts of videotapes. In Figures 3 and 4, each data point represents a 15-minute traffic

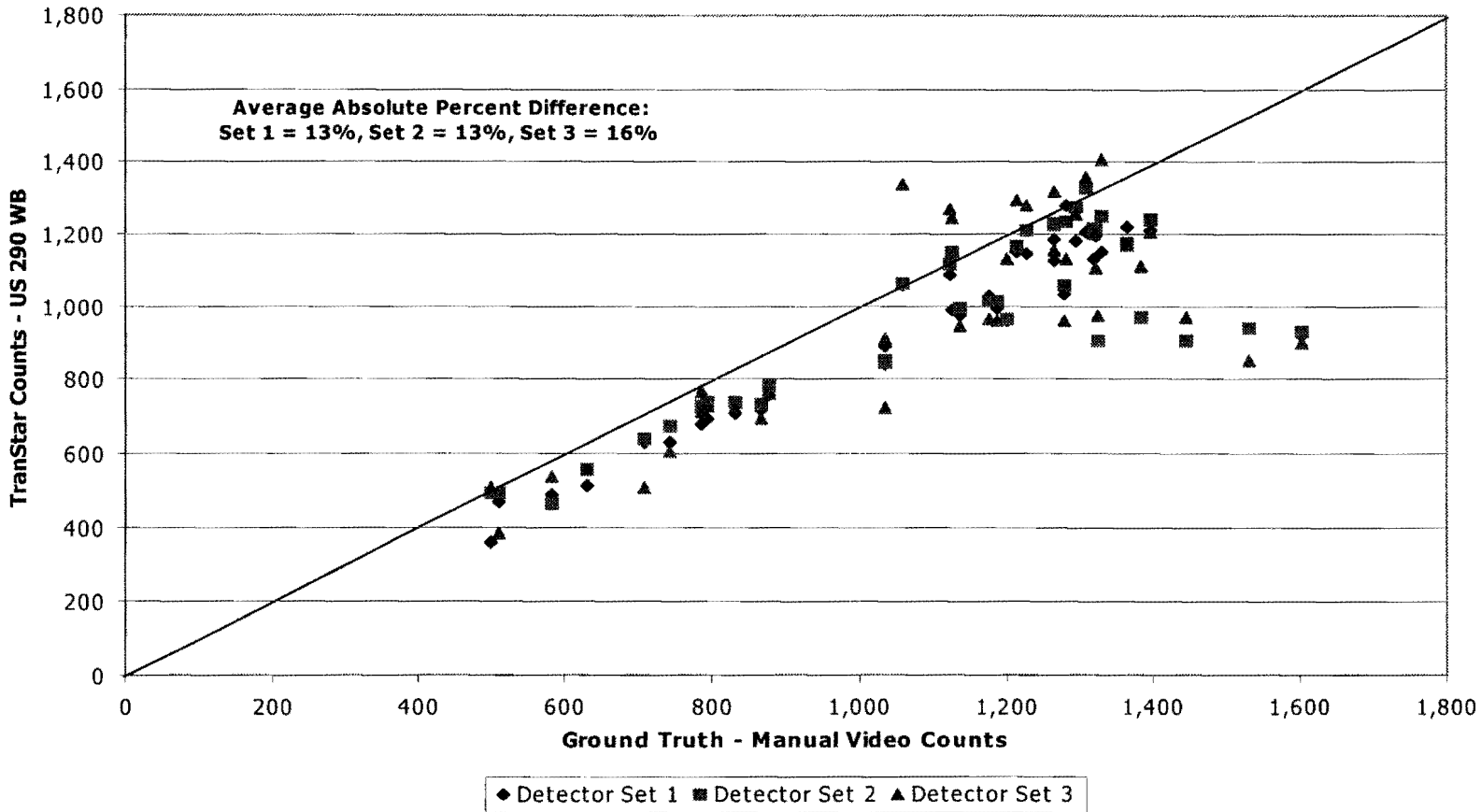
volume count across all directional freeway mainlanes. These comparisons show reasonably consistent results. For westbound detectors (Figure 3), the TranStar traffic volume counts are, on average, within about 16 percent of ground truth. For the eastbound detectors (Figure 4), the TranStar traffic volume counts are within 8 percent of ground truth traffic volumes.



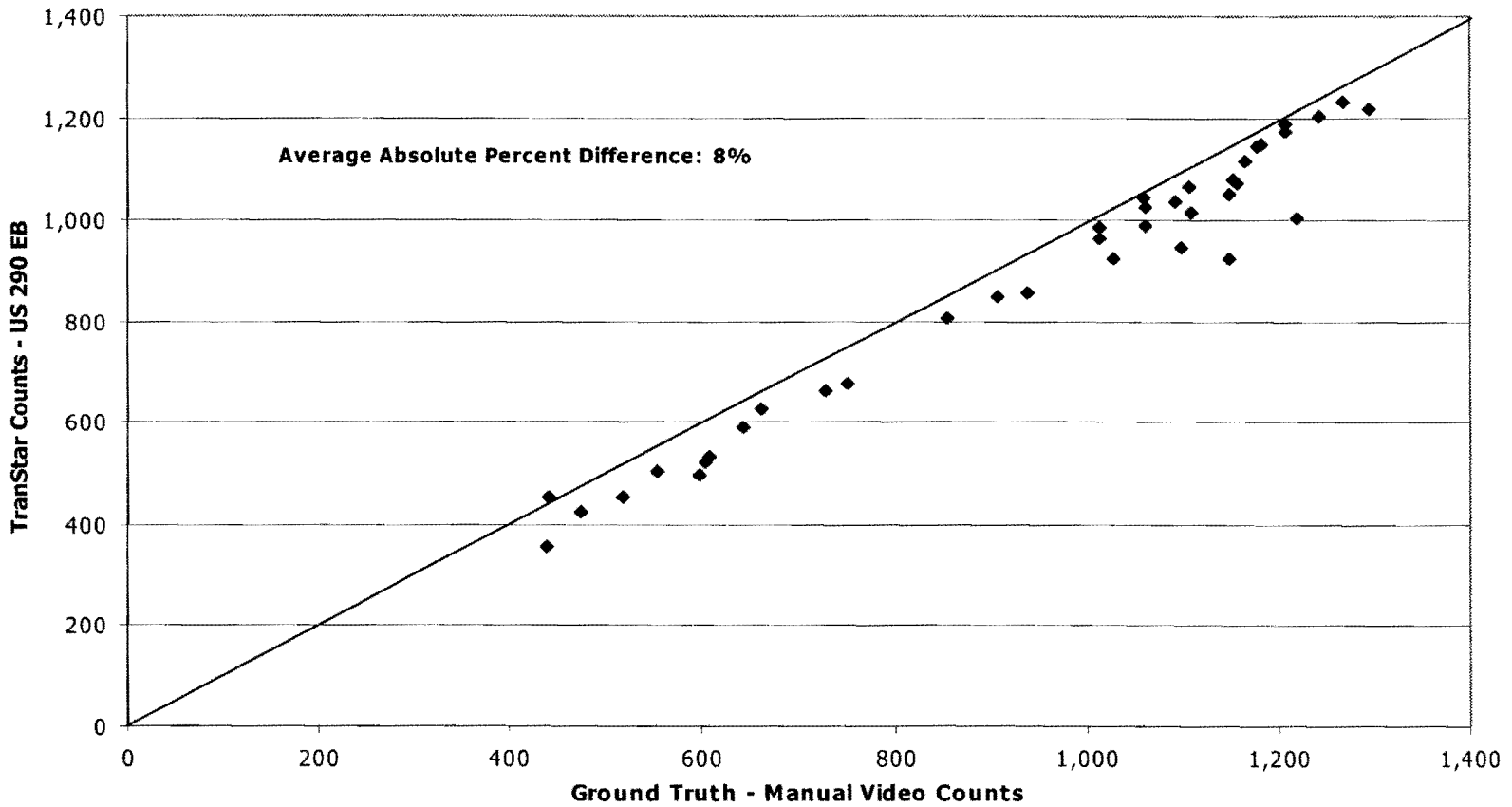
**Figure 1. Comparison of TransGuide and TP&P Daily Traffic Volume Estimates on San Antonio Freeways and Ramps, 1998.**



**Figure 2. Comparison of TransGuide and TP&P Daily Traffic Volume Estimates on San Antonio Freeways and Ramps, 1999.**



**Figure 3. Comparison of TranStar 15-Minute Traffic Volume Counts to Ground Truth, US 290 Westbound near Pinemont in Houston.**



**Figure 4. Comparison of TranStar 15-Minute Traffic Volume Counts to Ground Truth, US 290 Eastbound near Pinemont in Houston.**

## INFORMATION ON DATA ARCHIVING ARCHITECTURE AND STANDARDS

Several key references regarding data archiving architecture and standards are summarized in Report 2127-3. These references are provided here for readers wanting additional information:

- *ITS as a Data Resource: Preliminary Requirements for a User Service*, Report No. FHWA-PL-98-031, April 1998 (<http://www.fhwa.dot.gov/ohim/its/itspage.htm>).
- “Archived Data User Service (ADUS): An Addendum to the ITS Program Plan,” Version 3, September 1998 ([http://www.itsdocs.fhwa.dot.gov/jpodocs/repts\\_pr/41401!.htm](http://www.itsdocs.fhwa.dot.gov/jpodocs/repts_pr/41401!.htm)).
- ITS America’s Archived Data User Service Resource Page, <http://www.itsa.org>, search for “Archived Data User Service Resource Page.”
- Archived Data User Service in the National ITS Architecture, <http://www.iteris.com/itsarch>.
- *ITS Data Archiving Resources: Resources for Implementing ADUS*, CD-ROM developed for Federal Highway Administration by Texas Transportation Institute, 2000.

Report 2127-3 contains information on the Archived Data User Service in the National ITS Architecture, as well as the relevant archived data standards that are being developed.



### CHAPTER 3. CONCLUSIONS AND RECOMMENDATIONS

This chapter provides a summary of conclusions and recommendations based upon this research project. Report 2127-3 contains the guidance information that has been prepared based upon the research activities.

Based upon the findings of the research team, we offer the following conclusions and recommendations:

- **TxDOT should begin (or continue in some regions) the dialogue to determine who will be responsible for maintaining ITS data archives at the regional and statewide level.** The responsibility for data archives should be considered part of a regional traffic management center's "concept of operations." At the statewide level, responsibility should be assumed for archived ITS data of statewide interest, such as traffic volumes and speeds on TxDOT-maintained roadways. In some districts, such as Houston or Fort Worth, planning has already started for district-level or regional data archives. In other regions and at the statewide level, no one has assumed this responsibility for further development of data archives.

The dialogue about data archiving at the regional level should include the TxDOT district, the metropolitan planning organization, city and county transportation agencies, transit agencies, and other interested stakeholders as identified in Report 2127-3. Similarly at the statewide level, TxDOT's TP&P Division, the Texas Natural Resource Conservation Commission (TNRCC), and other interested statewide groups or agencies should be engaged in conversations about maintaining archived ITS data in statewide information systems.

- **TxDOT should develop one or two modest prototypes of single-source data archives (freeway detector data being the most desirable) to demonstrate proof of concept to system developers as well as potential data users.** These prototypes will make the existing detector data files (which are large, multi-million record text files) easily accessible to typical computer users, thus enabling most users to better understand the type and quantity of data being collected in Texas' traffic management centers. The prototypes will also help system developers and integrators to better understand the issues of larger data archiving efforts. This "start small but think big" approach comes from other industries, where large, complex data warehousing efforts have failed or struggled for years trying to "be all things to all people."

We recommend that data archiving efforts in Texas focus on permanently storing freeway detector data (vehicle volume, occupancy, and speed) at a common time interval, such as 5 minutes by lane, then making various data summaries and reports available through an easy-to-use interface (spreadsheet or web browser). Once experience has been gained,

enhancements can include integration of other databases (such as roadway incidents and work zones) as well as more sophisticated analysis and reporting features.

- **TxDOT should consider the data archiving guidelines in Report 2127-3, as well as the results from ongoing federal activities, when further developing data archiving systems.** The guidelines (Report 2127-3) address several relevant issues in the development of data archiving systems, such as a) basic principles of data archiving; b) determining what data and how much to archive; c) performing quality control to ensure data meets certain quality requirements; and d) developing a data archiving system using the National ITS Architecture and relevant standards. Beyond the guidelines that were developed in this TxDOT project, however, the Federal Highway Administration has a data archiving program that is producing useful information for implementing ADUS as documented in the National ITS Architecture.

## REFERENCES

1. *TranStar Data Warehouse User Data Needs*. TDW-Version 1.0, Prepared by Southwest Research Institute for Texas Department of Transportation, March 13, 2000.
2. Turner, S.M., W.L. Eisele, B.J. Gajewski, L.P. Albert, and R.J. Benz. *ITS Data Archiving: Case Study Analyses of San Antonio TransGuide® Data*. Report No. FHWA-PL-99-024. Federal Highway Administration, Texas Transportation Institute, College Station, Texas, August 1999.
3. Schmoyer, R., P. Hu and R. Goeltz. *Statistical Data Filtering and Aggregation to Hour Totals of ITS Thirty-Second and Five-Minute Vehicle Counts*. TRB Preprint 01-0873. Paper presented at the 80<sup>th</sup> Annual Meeting of the Transportation Research Board, Washington, DC, January 2001.
4. Grant, C., B. Gillis and R. Guensler. Collection of Vehicle Activity Data by Video Detection for Use in Transportation Planning. In *ITS Journal*, Volume 5, Number 4, 1999.