TECHNICAL REPORT STANDARD PAGE

1. Report No. FHWA/LA- 314	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Development of CMS Monitoring Procedures	5. Report Date April 1998	
	6. Performing Organization Code	
7. Author(s) Cesar A. Quiroga and Darcy Bullock	8. Performing Organization Report No.	
9. Performing Organization Name and Address Remote Sensing and Image Processing Laboratory Louisiana State University Baton Rouge, LA 70803	10. Work Unit No.	
	11. Contract or Grant No.	
12. Sponsoring Agency Name and Address Louisiana Transportation Research Center 4101 Gourrier Avenue Baton Rouge, LA 70808	13. Type of Report and Period Covered Final Report February 1997 - April 1998	
	14. Sponsoring Agency Code	

15. Supplementary Notes

Conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.

16. Abstract

This report describes a set of CMS monitoring procedures, collectively called Travel Time with GPS (TTG), which can be implemented on low cost, commonly available PCs. The main contribution is a GPS-GIS methodology which can be used by district engineers, city traffic officials, and MPO planners to conduct travel time studies and monitor congestion at various levels of resolution (from system to local). TTG allows users 1) to process GPS travel time data in a coherent, controlled way; 2) to store the data in a relational database configuration, which is critical for data archival purposes and the development of client-server applications; and 3) to efficiently retrieve and use the data for analysis and reports. TTG is built around a general data model that includes a spatial model, a geographic relational database, and a procedure for linearly referencing GPS data using GIS dynamic segmentation tools.

The report includes an updated methodology for estimating required sample sizes. This methodology uses t-distribution parameters and a more reliable procedure to estimate average sample ranges than that suggested in current ITE guidelines. The report also includes a new methodology to measure delay at signalized intersections based on linearly referenced GPS data. This methodology is compatible with new Highway Capacity Manual procedures that focus on the measurement of control delay (which includes deceleration delay, stopped delay, and acceleration delay) as opposed to just stopped delay. Because it uses automated position and speed data collection devices (GPS receivers), the methodology can be used in situations where automatic and/or real-time delay measurements are needed.

The implementation of TTG in this project uses a TransCAD - Access environment. This environment allows users to process vast amounts of GPS data and produce speed and travel time reports quickly and cost-effectively. The TransCAD - Access environment works well in most cases. However, there are problems, particularly with the current version of TransCAD (3.1), that may have an impact on results and/or performance. These problems are particularly significant when TransCAD-generated data must be exchanged with other applications.

17. Key Words Congestion Management, GPS, GIS, travel time studies, data collection, data reduction, data reporting, TTG	18. Distribution Statement Unrestricted.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	87	

Form DOT F1700.7 (1-92)

DEVELOPMENT OF CMS MONITORING PROCEDURES

by

Cesar A. Quiroga
Research Associate
Remote Sensing and Image Processing Laboratory
Louisiana State University
Baton Rouge, LA 70803

Darcy Bullock
Associate Professor
School of Civil Engineering
Purdue University
West Lafayette, IN 47907

LTRC RESEARCH PROJECT 97-7SS
STATE PROJECT 736-99-0450

conducted for

LOUISIANA DEPARTMENT OF TRANSPORTATION AND DEVELOPMENT
LOUISIANA TRANSPORTATION RESEARCH CENTER

April 1998

ABSTRACT

This report describes a set of CMS monitoring procedures, collectively called Travel Time with GPS (TTG), which can be implemented on low cost, commonly available PCs. The main contribution is a GPS-GIS methodology which can be used by district engineers, city traffic officials, and MPO planners to conduct travel time studies and monitor congestion at various levels of resolution (from system to local). TTG allows users 1) to process GPS travel time data in a coherent, controlled way; 2) to store the data in a relational database configuration, which is critical for data archival purposes and the development of client-server applications; and 3) to efficiently retrieve and use the data for analysis and reports. TTG is built around a general data model that includes a spatial model, a geographic relational database, and a procedure for linearly referencing GPS data using GIS dynamic segmentation tools.

The report includes an updated methodology for estimating required sample sizes. This methodology uses t-distribution parameters and a more reliable procedure to estimate average sample ranges than that suggested in current ITE guidelines. The report also includes a new methodology to measure delay at signalized intersections based on linearly referenced GPS data. This methodology is compatible with new Highway Capacity Manual procedures that focus on the measurement of control delay (which includes deceleration delay, stopped delay, and acceleration delay) as opposed to just stopped delay. Because it uses automated position and speed data collection devices (GPS receivers), the methodology can be used in situations where automatic and/or real-time delay measurements are needed.

The implementation of TTG in this project uses a TransCAD - Access environment. This environment allows users to process vast amounts of GPS data and produce speed and travel time reports quickly and cost-effectively. The TransCAD - Access environment works well in most cases. However, there are problems, particularly with the current version of TransCAD (3.1), that may have an impact on results and/or performance. These problems are particularly significant when TransCAD-generated data must be exchanged with other applications.

ACKNOWLEDGMENTS

This project was supported by the Louisiana Department of Transportation and Development (DOTD), through the Louisiana Transportation Research Center (LTRC Research Project 97-7SS; State Project 736-99-0450). The help of Jim Joffrion, Art Rogers, Ramya Sarma, and Caroline Heifner is gratefully acknowledged.

IMPLEMENTATION STATEMENT

The deliverable of this research project is a set of procedures for monitoring traffic congestion in urban areas using global positioning system (GPS) and geographic information system (GIS) technologies. The set of procedures is collectively called Travel Time with GPS (TTG) and is implemented in the form of computer programs and spreadsheet macros which can be run on ordinary personal computers.

For running the programs and macros we developed, the following hardware and software are required:

- IBM Compatible PC-486 or higher, at least 16 Mbyte RAM running Windows 3.1 or higher, Windows 95, or Windows NT 3.51 or higher
- Caliper TransCAD version 3.1 (or TransCAD for short)
- Microsoft Access version 2.0 (or Access)
- Microsoft 16-bit Open Database Connectivity (ODBC) Driver Pack version 2.0
- Microsoft Excel version 7.0 (or Excel)
- TTG CD. This CD contains the TTG add-in, Access database file and other utilities, as well as a sample application. More specifically, the CD contains the following directories and files:
  - ***readme.txt*** This file is a text file describing the contents of the CD.
  - ***INSTALL*** This directory contains all subdirectories and files needed to setup a new TTG project.
  - ***SAMPLE directory*** This directory contains a sample TTG project based on two GPS data files in Baton Rouge.
  - ***TTG directory***. This directory contains database tables, maps, and other files generated for the entire 428 data file set in Baton Rouge.

Note: The TTG CD and accompanying user manual are available at the Louisiana Transportation Research Center (LTRC) upon request.

TTG provides users with a PC-based, relatively inexpensive environment to process vast amounts of GPS data and produce speed and travel time reports quickly and cost-effectively. The estimated cost of the required commercial software is about $3,500.

We found Access to be adequate for handling linearly referenced GPS data. Appending data and building queries becomes slower as the number of records in the database increases. However, we believe this should not be a major problem given the kind of computer technology that is available today. As with all engineering software, we found TransCAD to have some deficiencies that could produce erroneous results if care is not taken in processing the data. These deficiencies become particularly significant when TransCAD-generated data must be exchanged with other applications. Some other deficiencies may affect performance but not necessarily numerical results. A complete description of the problems we found is included in the Evaluation of the TransCAD - Access Environment section.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

One of the mandates of the Intermodal Surface Transportation Efficiency Act (ISTEA) of 1991 was the development of Congestion Management Systems (CMSs) *[1]*. Also included in the mandate was the development of statewide highway traffic monitoring systems. An important component in the development of these systems is the capability to measure travel time (and by extension delay and speed) accurately and reliably. This is particularly critical at a time when the construction of new facilities nationwide has virtually stopped and, as a result, most of the interest is now in the implementation of more localized solutions to mitigate congestion.

In the CMS project conducted for Baton Rouge, Shreveport, and New Orleans in 1995 and 1996, we developed a new GPS-GIS methodology for conducting travel time studies *[2]*. Compared to traditional approaches, the new methodology provided consistency, automation, finer levels of resolution, and better accuracy in measuring travel time and speed. These characteristics enabled the detection of traffic flow variability along routes much more efficiently than with traditional techniques. The new methodology automated the data collection and data reduction procedures, and provided improved procedures for documenting and analyzing travel time and speed data. As a result, we were able to collect and process large amounts of reliable travel time and speed data.

The GPS equipment we used for data collection was quite inexpensive (less than $2,000, including a GPS receiver, a differential correction unit, and a laptop for data storage). However, the GIS system that we used for data reduction was very expensive. When the project was first begun in January 1995, we selected an Intergraph GIS platform (approximate cost $15,000) because this was the only way to ensure compatibility with the software used by the Louisiana Department of Transportation and Development (DOTD) cartographic unit in Traffic and Planning. However, in recent years, several inexpensive GIS packages (less than $3,000) which can run on ordinary personal computers (PCs) have emerged. Therefore, it would be of interest to adapt the current GPS-GIS based methodology to one of these cheaper packages so that a much wider audience including district engineers, city traffic officials, and metropolitan planning organization (MPO) planners can have the necessary tools for efficiently conducting travel time studies and monitoring congestion. Because these new potential users may have markedly different data needs, the modified GPS-GIS methodology would have to provide users with an expanded capability to monitor traffic conditions at various levels of resolution (from system to local, say at a signalized intersection).

OBJECTIVE

The goal of this research study is to develop a set of CMS monitoring procedures which can be implemented on low cost PCs commonly available. Its main contribution is a GPS-GIS methodology which can be used by district engineers, city traffic officials, and MPO planners for conducting travel time studies and monitoring congestion at various levels of resolution (from system to local). Specific objectives are summarized as follows:

- An efficient PC-based procedure for producing GPS-based highway network maps suitable for travel time studies. This is necessary so that traffic engineers throughout the districts can rapidly use this technology for evaluating and optimizing traffic flow along arterial streets and highway sections.
- An updated procedure for estimating sample size requirements during the data collection process.
- A PC-based procedure and accompanying software based on dynamic segmentation for generating a linear referencing system to all GPS travel time and speed data along arterial streets and highway sections. This is necessary so that traffic flow monitoring at various levels of resolution is possible. The procedure is a generalization of the procedures developed during the first phase of CMS project because of the capability to define any segmentation scheme (current segments are nominally 0.2 miles long).
- A procedure for reporting segment travel time and speed data both in graphical and tabular formats. Again, this is a generalization of the procedures developed during the first phase of the CMS project.

SCOPE

This research study is concerned with the development of a set of procedures for monitoring congestion using GPS and GIS. These procedures are meant to be used more as a planning tool than for everyday traffic monitoring. Under this assumption, a series of travel time runs are conducted along specified routes at specific time intervals, for example, every two years. The procedures can also be used for before-and-after studies where the need for specific travel time runs is clearly defined. In either case, GPS receivers are used to collect data at short time intervals, say every one second. This GPS data is kept, processed, and stored in a relational database for analysis purposes. For everyday traffic monitoring, collecting GPS data every one second is still feasible, but keeping and processing all this data is probably inefficient and costly. In this case, the data reduction procedure may be simplified so that only link or section aggregated data is kept in the database.

This research study focuses on travel time and speed only and not on other variables like flow rates, traffic composition, availability of alternative modes of transportation, and characteristics of the highway network, all of which play an important role in the process of assessing congestion in urban areas. Likewise, this research project does not address specific implementation issues like ranking of congested sections or coordination with other transportation management systems. DOTD is addressing these issues through a separate contract with a consultant (Parsons Brinckerhoff Quade and Douglas).

METHODOLOGY

Evaluation of GIS Packages

We made a comparison among existing GIS packages to select the most suitable software configuration for the project. More specifically, we evaluated the following GIS packages: ArcView version 3, MapInfo version 4.1, Maptitude version 3, and TransCAD version 3.1. We considered the following issues: general capabilities, linear referencing capabilities, compatibility with Intergraph's Modular GIS Environmet (MGE) platform, capability to generate line diagrams (diagrams in which one axis is distance and the other axis is an attribute such as speed or travel time), export/import capabilities, and cost.

Table 1 compares the capabilities of the four GIS packages considered. ArcView and TransCAD appeared to be the most suitable packages for our application. ArcView had the advantage of being less expensive than TransCAD. However, TransCAD appeared to provide more powerful network and routing capabilities than ArcView. In addition, TransCAD supported the OLE specification (which ArcView did not) and, consequently, it showed potential for the production of tabular and graphical reports involving other Windows applications. Furthermore, TransCAD had travel demand analysis capabilities which are of interest to DOTD and several MPOs in Louisiana (even though they were not a requirement in our project). For all these reasons, we chose TransCAD as our GIS platform.

Spatial Model

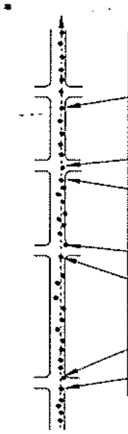
GPS receivers record location in latitude-longitude pairs. Furthermore, GPS data files tend to have huge numbers of records, particularly if data is collected at short time intervals, for example every one second. As a result, formal procedures for linearly referencing, storing, and retrieving the GPS travel time and speed data efficiently become essential.

One way to circumvent the GPS data storage problem is by aggregating the GPS data into highway segments or links so that only segment (or link) travel time and speed data are stored in the database. One of the drawbacks of this approach is that the rich detail of the original data is lost because only segment data are stored in the database (the original GPS files may still exist, but they are not really usable unless they are retrieved and processed from scratch). Some of the information contained in the original GPS data include acceleration and deceleration patterns, control delay, and stopped delay, all of which occur regardless of any highway segmentation scheme considered. In order to access this information it is necessary to store all GPS point data in the database and provide a linear reference to each GPS point before attempting any GPS data aggregation. This referencing can be performed with the help of GIS dynamic segmentation tools. Unfortunately, using this capability has been, until recently, out of the reach for most agencies because of high data storage and processing demands. These limitations are quickly disappearing, though, as more affordable computers with much larger data storage capabilities and faster processors enter the market.

Table 1
Comparison between GIS packages

Description	ArcView version 3	MapInfo version 4.1	Maptitude version 3	TransCAD version 3.1
General capabilities	Generic desktop mapping and GIS tool. Provides mapping functionality, tabular data management, and the possibility of creating customized applications. It is a "front-office" application in the Arc/Info group of products.	Business-oriented desktop mapping tool. It is mainly used for area wide geographic analyses such as buffering, aggregating and disaggregating areas, and redistricting.	Business-oriented desktop mapping tool, and is mainly used for area wide geographic analyses.	Transportation-specific GIS package. Combines GIS and transportation planning functions. A limited version called Base TransCAD handles networks, route systems, and mileposts, but not travel demand modeling procedures.
Linear referencing capabilities	Geocodes graphical elements by street address, X-Y point locations (which is suitable for handling GPS data), and route-mapped point and linear events. ArcView Network Analyst is an extension that provides routing capabilities to ArcView including computation of shortest and optimum routes and definition of accessibility areas by drive time. It also provides customizable directions including reference landmarks.	Geocodes graphical elements by street address, ZIP codes, census tract, or user-defined match options. Software does not have dynamic segmentation capabilities. Blue Marble GPS software included in the tracks GPS data.	Geocodes graphical elements by street address, ZIP codes, census tract, or user-defined match options. It also allows users to determine shortest and fastest routes. Maptitude GPS is an add-in that performs real-time location tracking and data acquisition using laptop-based GPS receivers.	Linearly references attribute data, allows directions of flow at individual links, delays and restrictions for turn movements at intersections, routes as collections of geographic features, and builds spatial queries based on linearly-referenced attribute data. The software also handles address matching tools and geocoding tools, and provides interfaces to data collection tools including GPS.
Compatibility with MGE platform	Reads map data directly from Arc/Info coverages, AutoCAD (.dxf and .dwg files), and MGE *.dgn* files in their native format. The software creates virtual views of original files.	Imports MGE *.dgn* files into MapInfo format.	Reads MGE *.dgn* files with an add-in product. The add-in imports *.dgn* files and creates geographic files in Maptitude format.	Reads MGE *.dgn* files with an add-in product. The add-in imports the *.dgn* file and creates a geographic file in TransCAD format.
Capability to generate line diagrams	Does not generate line diagrams. However, it does generate line charts which represent trends over time.	Does not generate line diagrams.	Does not generate line diagrams.	Generates line diagrams (called strip charts). It is possible to customize charts using the GIS Developing Kit (GISDK), which is the developing language included in TransCAD.
Export/ import capabilities	Reads data in Oracle, Access, dBASE, and text file formats. It also reads from and writes to Open Database Connectivity (ODBC)-compliant databases. It does not support Object Linking and Embedding (OLE) and, therefore, only attribute data (i.e. no maps) can be passed to client applications.	Reads data directly from Microsoft Excel, Lotus 1-2-3, dBASE, Access, and text files. It also reads from and writes to ODBC-compliant databases.	N/A	Reads data directly from dBase and text files. It also reads from and writes to ODBC-compliant databases. It supports OLE allowing attribute data and maps to be passed to client applications.
Cost	$2,690. Includes ArcView version 3 and Network Analysis.	$1,295. Includes MapInfo Professional version 4.1 and Blue Marble GPS software.	$1,385. Includes Maptitude version 3, GPS add-in, and DGN reader add-ins.	$3,490. Includes Base TransCAD version 3.1 and DGN reader add-in.

Linearly referencing GPS data involves computing cumulative distance or milepost values for all GPS points that are located along a specified route. As an illustration, figure 1 shows an example test vehicle trajectory on a highway route, with GPS data (location and speed) recorded every one second. Location is recorded in latitude-longitude format. Using GIS dynamic segmentation tools, we can map GPS points to the highway network and calculate milepost data (figure 1, Milepost column). With this information, we can calculate distance traveled and travel time. For example, based upon the milepost attributes recorded at 8:30:26 and 8:30:53, we see the vehicle traveled 0.142 miles in 27 seconds.



GPS point data

Time	Latitude	Longitude	Speed (mph)	Milepost
8:30:53	30.4234600	-91.1555914	24.2	0.142
:	:	:	:	:
8:30:48	30.4229991	-91.1555900	18.5	0.110
8:30:47	30.4229523	-91.1555915	15.5	0.107
:	:	:	:	:
8:30:39	30.4224010	-91.1555890	19.2	0.069
8:30:38	30.4223652	-91.1555920	21.2	0.067
:	:	:	:	:
8:30:27	30.4213956	-91.1555895	20.5	0.002
8:30:26	30.4213598	-91.1555915	20.2	0.000

Figure 1
GPS data recording and linear referencing on a highway route

In order to linearly reference travel time GPS data efficiently, it is essential to use a good base vector map to link the GPS records to the database. One approach is to use digitized quad maps or topologically integrating geographic encoding and referencing (TIGER) files (figure 2a), but such maps only provide a crude representation of the network. As a result, problems like misplaced or incorrectly modeled geographic features are common. For example, misplaced signalized intersecting streets have the effect of distorting the relative position of the GPS data with respect to the signalized intersection. A misplacement of only 0.03 mi (50 m) (compare the relative position of the Bon Marche Mall intersection with respect to the Lobdell intersection in figure 2a and figure 2b) could result in GPS data being mapped to the wrong side of the signalized intersection. Incorrectly modeled features have the effect of masking or distorting actual traffic conditions. For example, figure 2a suggests that Wooddale Boulevard ends at Florida Boulevard. However, Wooddale Boulevard actually crosses Florida Boulevard (figure 2b).

Problems with incorrectly modeled features are particularly evident in the case of controlled access facilities. As shown in figure 3a, highway interchanges are usually represented by single straight lines or by lines connecting bi-directional vector elements. As a result, when GPS points are displayed on GIS maps containing this kind of features, the offset between GPS points and the physical discontinuity features is usually quite significant. In contrast, figure 3b shows a digital representation of the same area based on GPS data. Notice that the offset of some of the physical discontinuities in figure 3a with respect to those of figure 3b is quite large: between 0.1 and 0.2 miles. This would also be the offset associated with GPS travel time data when trying to map this data to physical discontinuities such as those shown in figure 3a.

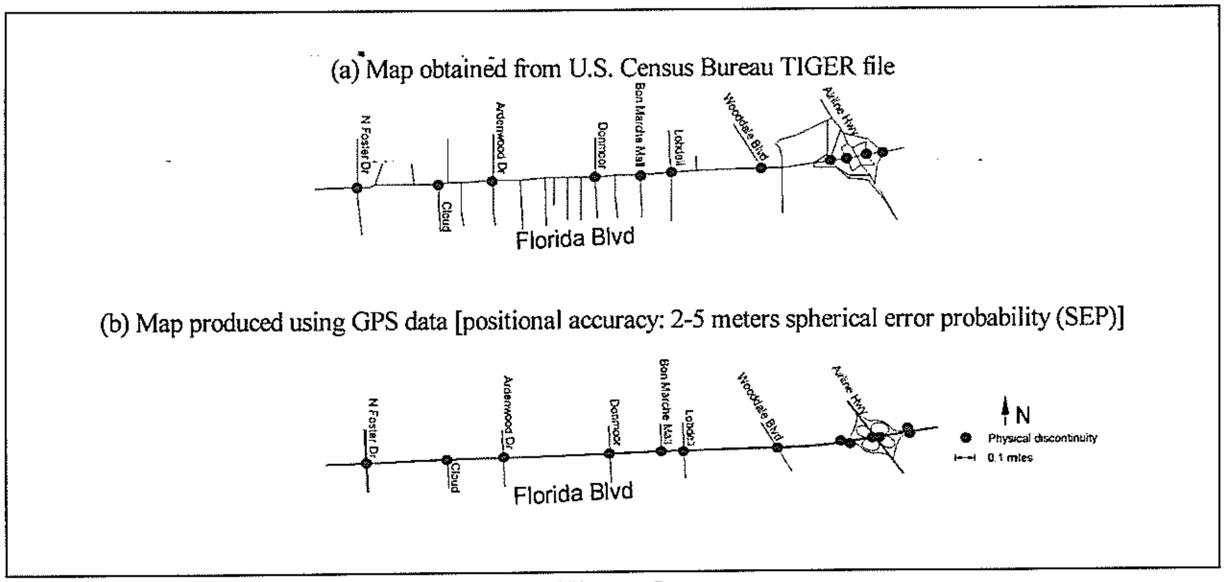


Figure 2
Digital representation of Florida Boulevard between North Foster Drive and Airline Highway in Baton Rouge, Louisiana

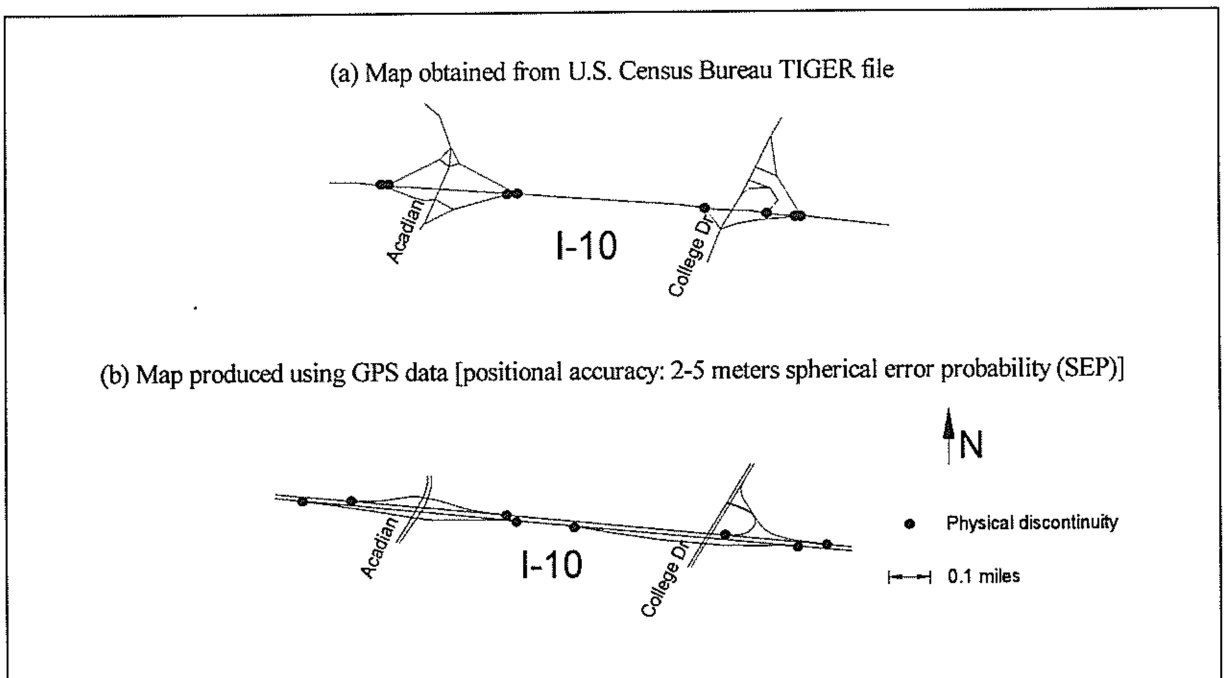


Figure 3
Digital representation of I-10 between Acadian Thruway and College Drive in Baton Rouge, Louisiana

To address the problem of poor existing GIS maps, we developed a procedure to construct GPS-derived network maps for processing GPS travel time data. This procedure involves the following general steps:

1. Drive a GPS-equipped vehicle to survey both directions of travel, as well as the cross streets at signalized intersections and other physical discontinuities (figure 4a). Normally, it is sufficient to provide differentiation by direction of travel only (unless lane differentiation is required). This means that, in most cases, it is sufficient to do the survey by driving the probe vehicle on the middle lane of each direction of travel using GPS equipment capable of providing a positional accuracy of around 2-3 m. This GPS equipment would be the same or have the same characteristics as that used for collecting actual travel time data, therefore ensuring positional accuracy compatibility between the GPS-based travel time data and the underlying GIS map. In general, all GPS data must be differentially corrected to correct for errors such as selective availability (SA) errors, satellite errors (atomic clock errors, frequency offsets, and hardware delay), and atmospheric errors (mainly ionospheric errors) *[3] [4]*.
2. Import the differentially corrected GPS (DGPS) data resulting from the initial survey into a GIS and generate a directional centerline highway network map (figure 4b). The network map generation process involves using the lowest possible number of outline points without jeopardizing positional accuracy. A good rule of thumb is to draw all graphical features with the lowest number of outline points such that the offset between the line work and actual GPS travel time data is less than the positional accuracy associated with the GPS data. In the case of relatively straight highway segments, only a few points are needed to represent the directional centerline alignment properly. In the case of curved segments and physical discontinuities like ramps and interchanges, a larger number of outline points are needed to represent the feature properly. In general, though, it is almost always possible to generate linear features with a much lower number of outline points than the number of GPS points collected during the survey. The end result should be a workable, low-burden base map with sufficient positional accuracy for processing GPS travel time data.
3. Formalize the location of physical discontinuities by locating checkpoints at all major physical discontinuities so that all roads, ramps, and service roads can be treated as identifiable entities (figure 4c). The resulting linear features are called links.
4. Assign unique identification numbers to each link (figure 4d). The resulting map is then ready for mapping GPS travel time data to the network (figure 4e). In the case of TransCAD, it is advisable to use a different identification field than that used by TransCAD to index graphical features automatically as they are generated. The TransCAD automatic indexing field is called ID. If for any reason a user needs to compact a geographic file in TransCAD (i.e. converting a standard geographic file into a compact file and then back into a standard geographic file), TransCAD redefines new ID values for all graphical features (the old ID values are stored in DATAx columns), therefore causing additional database complexities.

As we will discuss with greater detail later, mapping GPS data to the network allows us to compute a milepost value to each GPS point which, in turn, allows us to compute speeds and travel times along routes using any linear aggregation level we specify. For this purpose, it is convenient to define a consistent naming convention so that we know exactly what each aggregation level represents. In Baton Rouge, we have proposed to use the following four levels of aggregation, as shown in figure 5: segments, links, sections, and routes.

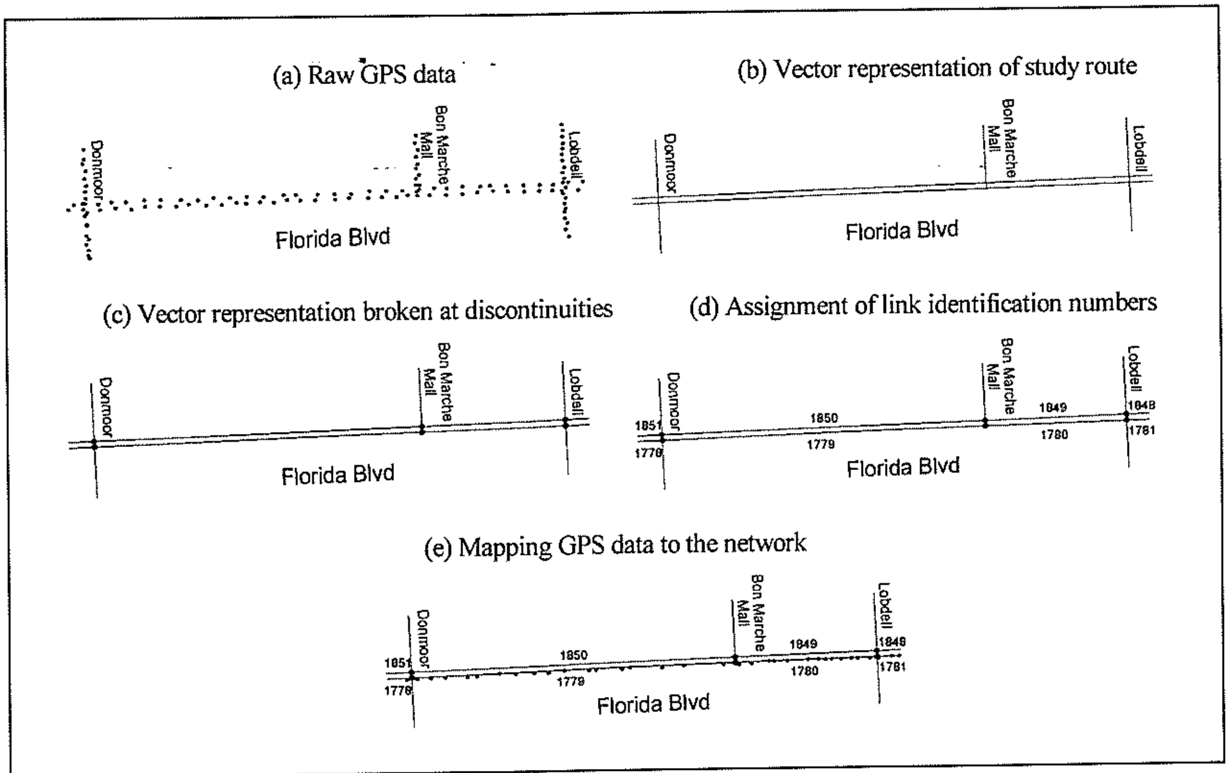


Figure 4
Sample network map geocoding and segmentation for a short section of Florida Boulevard in Baton Rouge, Louisiana

A segment is a short portion of highway (nominally 0.2 mi in length) that is associated with the segmentation scheme we followed during the first phase of the CMS project. Segments are directional. For example, as shown in figure 5a, segment 12444 always represents the 0.2 mi portion of I-10 eastbound (EB) right before the I-10 and I-12 split in Baton Rouge. Similarly, segment 12448 always represents the portion of I-12 westbound (WB) right before I-12 WB merges with I-10 WB in Baton Rouge. Notice that other segmentation schemes are also possible. In these cases, the resulting elements would also be called segments but, for consistency, some qualifier should be required to avoid any confusion. Examples of alternative segmentation schemes would include the use of 0.05-mi segments and 0.10-mi segments. Notice also that segments are designed so that they do not cross physical discontinuity points such as on-ramps, off-ramps, lane drops, and major signalized intersections. This simplifies the process of data aggregation into longer elements such as links, sections, or corridors.

A link is a portion of highway located between two contiguous physical discontinuity points along a route. A link is a special type of segment that is defined by the distance between contiguous physical discontinuity points. Since these distances are rarely, if ever, uniform, it would be impractical to use the x-mi segment naming convention. Instead, we use a simple name: link. Like segments, links are directional. For example, as shown in figure 5b, link 1240 always represents the portion of I-12 EB between the I-10 and I-12 split and the I-10 WB on-ramp. Links are the basic features we are using in this project for linear referencing GPS data along routes.

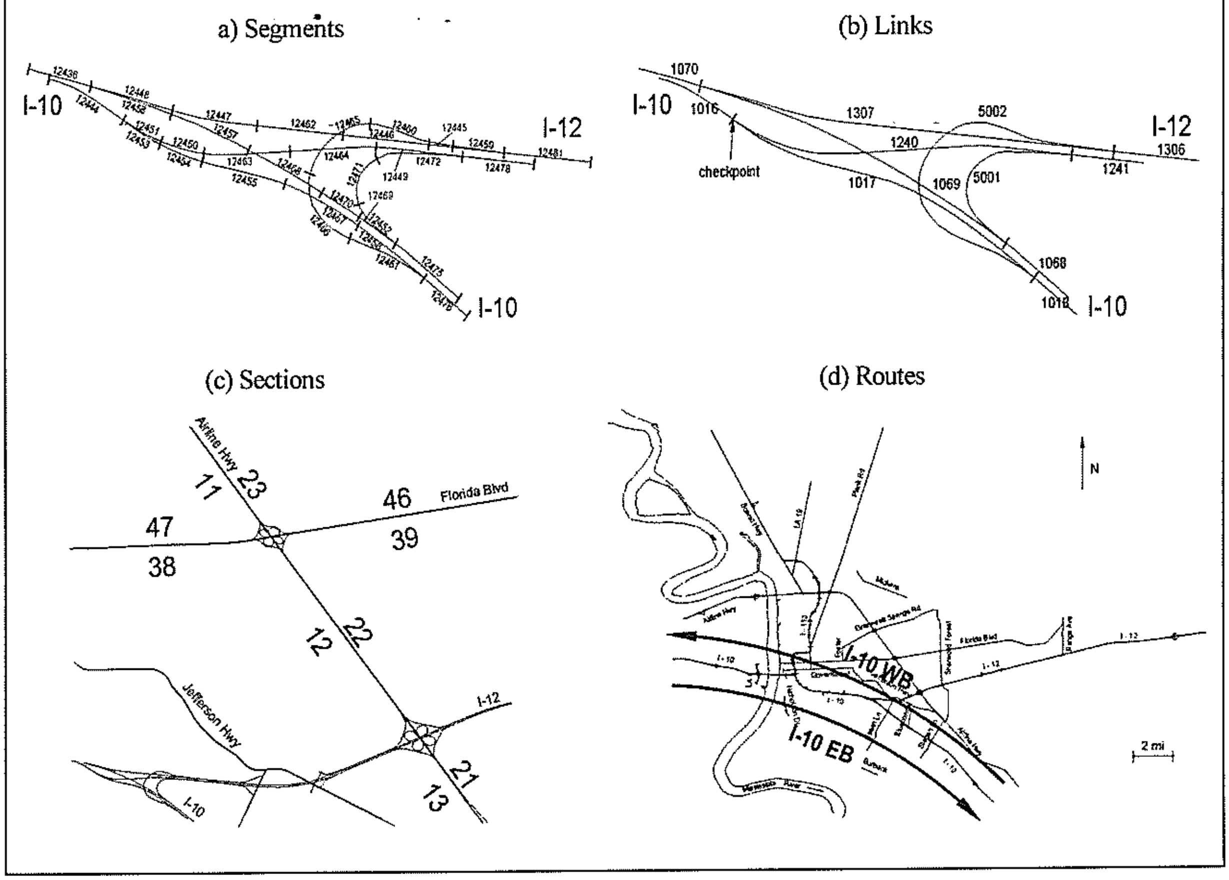


Figure 5
Aggregation levels in Baton Rouge, Louisiana

A section is a portion of highway located between two contiguous CMS congestion corridors. This scale is currently being used in Baton Rouge to provide a preliminary assessment of critical congestion areas. In general, a section is composed of several links. Like segments and links, sections are directional. For example, as shown in figure 5c, section 12 would always represent the portion of Airline Highway southbound (SB) between the Florida Boulevard WB off-ramp and I-12 WB off-ramp.

A route is a portion of highway that denotes the largest aggregation level for analysis purposes. Examples of routes are I-10 EB and I-10 WB (figure 5d). In general, routes are based on the initial selection of 22 congestion corridors made by the Baton Rouge Capital Region Planning Commission (CRPC).

Data Collection Methodology

GPS equipment

As mentioned above, the GPS equipment must be able to provide a positional accuracy of around 2-3 m. The GPS equipment we used in the first phase of the CMS project had a positional accuracy of 2-5 m spherical error probability (SEP) and a speed accuracy of 0.1 mph (1 sigma) *[2]*. It had the capability to collect and output the following data every one second: location (in latitude/longitude), time (in Universal Coordinated Time - UTC), speed (in mph), and differential correction status.

GPS technology is advancing at a very rapid pace. New GPS equipment configurations will likely be better and more efficient than what we used three years ago. In the short run, GPS equipment configurations will still be composed of GPS receivers and differential correction receivers. However, new GPS receivers have enhanced data storage capabilities, possibly rendering the use of on-board laptop computers unnecessary. Likewise, differential correction broadcasts can now be obtained from sources like the U.S. Coast Guard DGPS network or satellites, and not just from commercial, sometimes unreliable FM subcarriers.

Travel time runs

A set of runs must be scheduled to measure travel time, speed, and delay. In general, runs are made during the a.m. peak, p.m. peak, off-peak, and other specific traffic conditions. Recurrent congestion in a metropolitan area like Baton Rouge tends to be spread over a two to three hour period both in the morning and in the afternoon, although significant variations in traffic conditions do exist within each congestion period. Strictly speaking, runs should be scheduled at regular time intervals within each congestion period, for example every half hour, to determine which time period is the most congested. During the first phase of the CMS project, we detected particularly heavy conditions from 7:00 to 8:00 a.m. and from 4:30 to 5:30 p.m. At the very minimum, therefore, runs for congestion monitoring purposes should be scheduled within these time periods.

For each time period, the number of runs (or sample size n) must comply with acceptable error tolerance specifications. A commonly used formulation for estimating n is included in the *Institute of Transportation Engineers (ITE) Manual of Transportation Engineering Studies [5]*. Unfortunately, this formulation seriously underestimates required sample sizes (or, alternatively, it overestimates confidence levels) *[6]*. In order to provide more reliable estimates, we corrected and extended the ITE formulation. In the updated formulation, the required sample size n is given by

$$n = \left[\frac{t_\alpha \overline{R}}{d\varepsilon}\right]^2 \tag{1}$$

where:

α = significance level
t_α = two-tailed t distribution statistic for a confidence level of 1-α
$\overline{R}$ = sample range based on available data
d = ratio of $\overline{R}$ to σ (standard deviation of the population)

ε = user-selected allowable error in the estimate of the mean speed (or interval half-length).
The sample range $\overline{R}$ is computed as

$$\overline{R} = \max_{i=1}^{m} v_i - \min_{i=1}^{m} v_i \tag{2}$$

where:

m = sample size of available data
v_i = i^{th} segment speed observation of the initial study.

If there is no available data to estimate $\overline{R}$, m = 0. In this case, $\overline{R}$ can be estimated based on a quick assessment of local traffic conditions. The t_α statistic and the ratio d are functions of n, and, as a result, an iterative procedure must be followed to solve for n. As an aid to users, table 2 shows values of n for various combinations of $\overline{R}$ and ε, assuming confidence levels of 99.73 percent, 95 percent, 85 percent, and 75 percent. For example, assume $\overline{R}$ = 9 mph and $\varepsilon = \pm 5$ mph (this value is recommended for planning studies in the ITE guidelines *[5]*). From table 2, n would be 9 (for a 99.73 percent confidence level), 5 (for a 95 percent confidence level), 4 (for an 85 percent confidence level), and 3 (for a 75 percent confidence level).

Figure 6 shows a suggested workflow to be used with the new formulation. The first step [box (1) in figure 6] involves selecting a permitted error ε in the estimate of the mean speed. Following ITE guidelines, the suggested ranges of ε are as follows *[5]*:

- Transportation planning and highway needs studies: ±3 mph to ±5 mph
- Traffic operations, trend analysis, and economic evaluations: ±2 mph to ±4 mph
- Before-and-after studies: ±1 mph to ±3 mph.

The second step (2) involves defining an appropriate confidence level (1-α) based on specific needs, budgetary constraints, and other considerations. For example, if the agency in charge of the travel time study can allocate sufficient funds for a relative large number of runs, it is quite likely that an appropriate confidence level might be around 95 percent or even higher. However, if the agency is facing budget constraints, it might choose to make a relatively low number of runs. In confidence level terms, the agency may be forced to accept a lower confidence level, perhaps as low as 75 percent.

The third step (3) involves estimating an average sample range $\overline{R}$. If no previous travel time data exists for the route under study, i.e. if m = 0 (4), $\overline{R}$ can be estimated based on a quick assessment of local traffic conditions (5). If previous travel time data does exist for the route under study (i.e. m > 0), we can use equation (2) to estimate $\overline{R}$ (6). If more than one set of runs exists for the route, we first calculate $\overline{R}$ for each set of runs (6), and then we average all sample ranges (7).

Table 2
Sample size requirements for travel time studies (based on equation (1))

a) Confidence level (1-α): 99.73%

Average Range $\overline{R}$ (mph)	Specified Permitted Error ε				
	±1 mph	±2 mph	±3 mph	±4 mph	±5 mph
1	6	5	4	4	4
2	9	6	5	5	4
3	13	8	6	5	5
4	17	9	7	6	6
5	21	11	8	7	6
6	26	13	9	8	7
7	32	15	10	8	7
8	37	17	12	9	8
9	44	19	13	10	9
10	50	21	14	11	9
11	57	24	15	12	10
12	65	26	17	13	11
13	73	29	18	14	11
14	81	32	20	15	12
15	89	35	21	16	13
16	98	37	23	17	14
17	>100	41	25	18	14
18	>100	44	26	19	15
19	>100	47	28	20	16
20	>100	50	30	21	17
25	>100	69	40	28	21
30	>100	89	50	35	26

b) Confidence level (1-α): 95%

Average Range $\overline{R}$ (mph)	Specified Permitted Error ε				
	±1 mph	±2 mph	±3 mph	±4 mph	±5 mph
1	4	3	3	3	3
2	6	4	3	3	3
3	8	5	4	4	3
4	10	6	5	4	4
5	12	7	5	4	4
6	15	8	6	5	4
7	18	9	6	5	5
8	21	10	7	6	5
9	24	11	8	6	5
10	27	12	8	7	6
11	31	13	9	7	6
12	34	15	10	8	6
13	38	16	11	8	7
14	43	18	11	9	7
15	47	19	12	9	8
16	51	21	13	10	8
17	56	22	14	10	8
18	61	24	15	11	9
19	66	25	16	12	9
20	71	27	17	12	10
25	99	36	22	15	12
30	>100	47	27	19	15

c) Confidence level (1-α): 85%

Average Range $\overline{R}$ (mph)	Specified Permitted Error ε				
	±1 mph	±2 mph	±3 mph	±4 mph	±5 mph
1	3	3	2	2	2
2	4	3	3	3	3
3	6	4	3	3	3
4	7	4	4	3	3
5	9	5	4	3	3
6	10	6	4	4	3
7	12	6	5	4	4
8	14	7	5	4	4
9	16	8	6	5	4
10	18	9	6	5	4
11	20	9	7	5	5
12	23	10	7	6	5
13	25	11	7	6	5
14	28	12	8	6	5
15	30	13	9	7	6
16	33	14	9	7	6
17	36	15	10	7	6
18	39	16	10	8	6
19	42	17	11	8	7
20	45	18	11	9	7
25	62	24	15	11	9
30	81	30	18	13	10

d) Confidence level (1-α): 75%

Average Range $\overline{R}$ (mph)	Specified Permitted Error ε				
	±1 mph	±2 mph	±3 mph	±4 mph	±5 mph
1	3	2	2	2	2
2	4	3	3	2	2
3	5	3	3	3	2
4	6	4	3	3	3
5	7	4	3	3	3
6	8	5	4	3	3
7	9	5	4	3	3
8	11	6	4	4	3
9	12	6	5	4	3
10	14	7	5	4	4
11	15	7	5	4	4
12	17	8	6	5	4
13	19	9	6	5	4
14	20	9	6	5	4
15	22	10	7	5	5
16	24	11	7	6	5
17	26	11	8	6	5
18	28	12	8	6	5
19	30	13	8	6	5
20	33	14	9	7	6
25	44	18	11	8	7
30	58	22	14	10	8

The fourth step (8) involves estimating the required sample size n using either table 2 or equation (1). If the required sample size is larger than the number of runs already made (9), we schedule new or additional runs (10) under similar traffic and environmental conditions until the required sample size requirement is met (11) (6) (7) (8) (9).

The last step (12) involves computing representative statistical information about the speed data collected in the field. Once the sample has been collected, it makes sense to calculate the corresponding average speed $\bar{v}$, standard deviation s, and significance level α. We can also calculate $\bar{R}$. However, since we are now dealing with a sample already collected, it may be preferable to compute widely accepted, used parameters such as $\bar{v}$ and s. With s, we can provide a final check for α. To do this, we calculate t_α as

$$t_\alpha = \frac{\sqrt{n\varepsilon}}{s} \qquad (3)$$

and look up α in any t distribution table*[7]*. This value of α (or final significance level) has to be less than or equal to one minus the specified confidence level used in the travel time study.

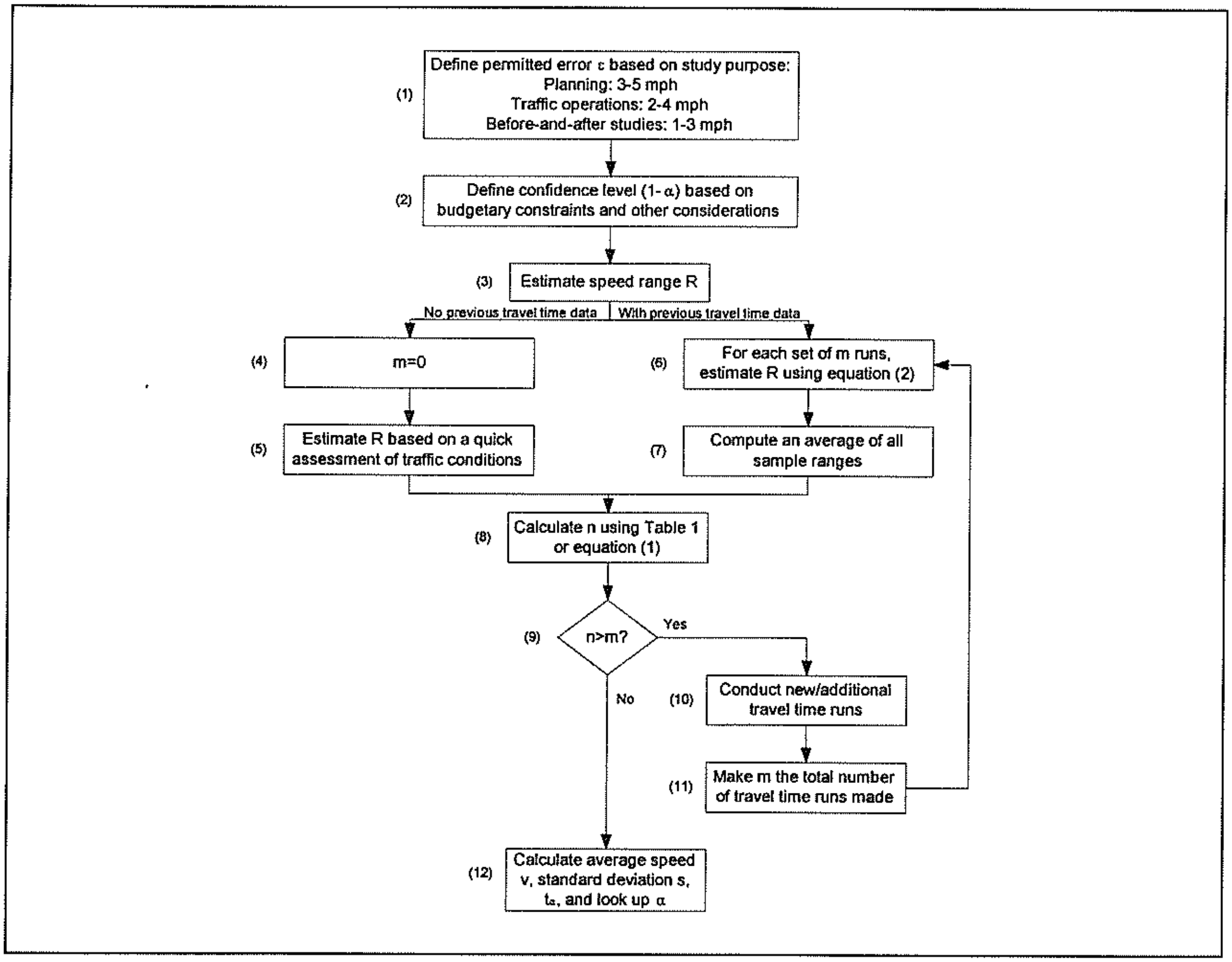


Figure 6

Recommended workflow to use with updated formulation for estimating n

Regardless of required sample size, standard procedures must be followed to ensure that all GPS data is consistent and that GPS files are properly handled and processed. The data reduction methodology discussed below addresses these issues.

When scheduling runs, it is advisable to break the highway network into sections which can be driven in relatively short periods of time. For example, the I-10 corridor in Baton Rouge is about 19 miles long, but driving this distance during congestion periods might easily take one hour or longer (on one direction only). Because traffic conditions can change significantly during one hour, it is best to divide the 19-mile stretch into, say, two or three smaller sections, each of which can be driven in less than half an hour. To ensure complete coverage, the sections must overlap. In the case of the I-10 corridor, then, two possible driving sections might be from LA 415 to College Drive and from Acadian Thruway to LA 42.

Because data collection with GPS is network and vehicle independent, individual GPS data files can contain data for several routes. There is no theoretical limit to the number of points collected and routes driven with a single GPS data file other than physical data storage capacity. In fact, the application we developed in this project allows users to handle multiple routes per GPS data file. However, for efficiency, it is advisable to limit the number of routes included in a single GPS data file to two: one for each direction of travel associated with a corridor. This way, GPS data collectors can drive both directions of travel associated with a specific section of highway using a single GPS data file. If the need arises, the technician can easily drive on the same section of highway several times, i.e. doing several loops, using a single GPS data file.

Database Management System

Each GPS data file contains data such as time stamps, speed, latitude, longitude, and satellite navigational data at regular time intervals, say every one second. This data needs to be linearly referenced so that GPS data can be associated with routes on the highway network. To manage all this information properly, we developed a geographic relational database. As shown in figure 7, the database schema is composed of 13 tables. A summary of the associated attributes is shown in table 3, and a sample of records is shown in figure 8.

The database schema described in figures 7 and 8 and table 3 assumes that a link code is explicitly associated with each GPS point (LinkCode attribute in table GPS_DATA). This link code results from the linear referencing process and is the same as that associated with links in the highway network map in TransCAD (see for example LinkCode = 1779 in figure 4 and table GPS_DATA of figure 8). We included attribute LinkCode in table GPS_DATA to aid users in the process of building queries which assume the same highway segmentation scheme as that used in generating the highway network map.

Strictly speaking, however, all we need from the linear referencing process is a route code and a milepost value for each GPS point (attributes RouteCode and MilePost in table GPS_DATA). With this information and any table containing the beginning and ending mileposts associated with links or segments along routes, we can produce generic GPS data tables for any highway segmentation scheme we want, and then use these tables for generating segment aggregated travel time and speed data (figure 9). This is particularly useful if we need to compute travel time and speed at a finer resolution than the one obtained with the original highway links.

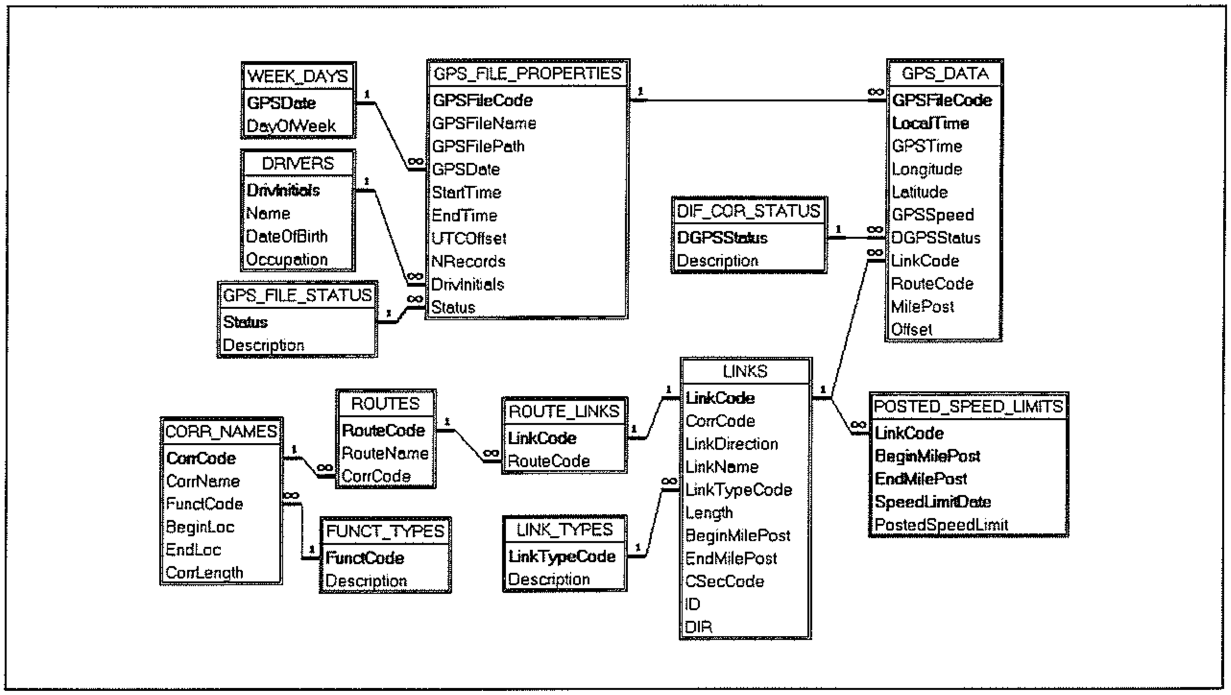


Figure 7
Geographic database schema

We used TransCAD in conjunction with Access for implementing the database schema described previously. Like some other desktop GIS packages, TransCAD uses an open architecture that allows users to open files and access data in a variety of formats. As part of this architecture, geographic files, maps, and tables are built using a fairly large number of associated files. For example, a geographic file may easily involve 10-20 associated files including graphical elements, tables, indexes, and data dictionaries. By design, most database files in TransCAD are in dBase IV format (*.dbf* extension), which means that each database table is stored in a separate dBase file.

While this kind of open architecture is intended to provide flexibility to typical TransCAD users, it can also complicate data management problems in an environment where tens or hundreds of GPS data files are being generated and processed. For example, for data quality control, it is much more efficient to have all linearly GPS data stored in a centralized database environment that allows users to implement data integrity constraints fully, instead of having GPS data scattered in several dBase files. For data reporting purposes, having all GPS data stored in a centralized database has the additional advantage of providing users with the capability to build queries and produce reports which involve aggregating or summarizing travel time and speed data by, say, time period or corridor. Having GPS data scattered in several dBase files would make the task of managing the data considerably more difficult.

Table 3
Description of geographic database table fields

Attribute	Key	Description
Table CORR_NAMES		
CorrCode	X	Corridor code
CorrName		Corridor name
FunctCode		Corridor function class code: 1 or 2
BeginLoc		Location where corridor begins
EndLoc		Location where corridor ends
CorrLength		Corridor length: in miles
Table FUNCT_TYPES		
FunctCode	X	Corridor functional class code: 1 or 2
Description		Corridor functional name
Table ROUTES		
RouteCode	X	Route identification code
RouteName		Route name
CorrCode		Corridor code to which route belongs
Table ROUTE_LINKS		
LinkCode	X	Link identification code
RouteCode		Route identification code
Table LINKS		
LinkCode	X	Link identification code
CorrCode		Corridor code to which segment belongs
LinkDirection		Link direction: EB, WB, NB, SB
LinkName		Link name
LinkTypeCode		Link type code
Length		Link length, measured with GIS: in miles
BeginMilePost		Milepost associated with beginning of link
EndMilePost		Milepost associated with end of link
CSecCode		DOTD control section code
ID		Link code generated by TransCAD
DIR		Directional code generated by TransCAD
Table LINK_TYPES		
LinkTypeCode	X	Link type code
Description		Link type name
Table POSTED_SPEED_LIMITS		
LinkCode	X	Link identification code
BeginMilePost	X	Posted speed limit beginning milepost
EndMilePost	X	Posted speed limit ending milepost
SpeedLimitDate	X	Effective date of posted speed limit
PostedSpeedLimit		Posted speed limit: in mph

Attribute	Key	Description
Table WEEK_DAYS		
GPSDate	X	Date GPS data was collected
DayOfWeek		Date of week GPS data was collected
Table DRIVERS		
DrivInitials	X	Probe vehicle driver initials
Name		Driver's name
DateOfBirth		Driver's date of birth
Occupation		Driver's occupation
Table GPS_FILE_STATUS		
Status	X	GPS file data reduction status code
Description		GPS file data reduction status description
Table GPS_FILE_PROPERTIES		
GPSFileCode	X	GPS file identification code
GPSFileName		GPS file name
GPSFilePath		Directory path associated with a GPS file
GPSDate		Date GPS data was collected
StartTime		First record time stamp: in seconds UTC
EndTime		Last record time stamp: in seconds UTC
UTCOffset		Local time - UTC time offset: in hours
NRecords		Number of records in GPS file
DrivInitials		Probe vehicle driver initials
Status		GPS file data reduction status code
Table DIF_COR_STATUS		
DGPSStatus	X	Differential correction status code
Description		Differential correction status description
Table GPS_DATA		
GPSFileCode	X	GPS file identification code
LocalTime	X	GPS point time stamp: in seconds local time
GPSTime		GPS point time stamp: in seconds UTC time
Longitude		GPS point longitude
Latitude		GPS point latitude
GPSSpeed		Probe vehicle speed: in mph
DGPSStatus		Differential correction status code
LinkCode		Link identification code
RouteCode		Route identification code
MilePost		GPS point milepost along route
Offset		GPS point - link transverse distance: in ft

We store all database tables in a single Access file. Tables LINKS, ROUTE_LINKS, and GPS_DATA contain records imported from TransCAD files. Table LINKS contains the same records as file *links.dbf*, which is a file generated by TransCAD for viewing attribute data associated with each link in the highway network. Table ROUTE_LINKS is the result of joining file *links.dbf* and all route link *.dbf* files in TransCAD. Table GPS_DATA contains linearly referenced GPS data which results from the linear referencing process in TransCAD. As explained with more detail in the following section, a typical work flow after collecting GPS data in the field begins and ends with Access. It involves using Access to generate a new entry in table GPS_FILE_PROPERTIES, then TransCAD to do the route assignment and linear referencing operations, and then Access again to append the linearly referenced GPS data to table GPS_DATA.

Table CORR_NAMES

CorrCode	CorrName	FunctCode	BeginLoc	EndLoc	CorrLength
6	Airline Hwy	2	La 1145	2 mi east of Highland Rd	22.87
7	Florida Blvd	2	River Front	Range Ave	14.69

Table FUNCT_TYPES

Funct Code	Description
1	Interstate
2	Principal arterial

Table ROUTES

RouteCode	RouteName	CorrCode
07FLBDE	Florida Blvd EB	7
07FLBDW	Florida Blvd WB	7

Table ROUTE_LINKS

LinkCode	RouteCode
1779	07FLBDE
1849	07FLBDW

Table LINKS

LinkCode	CorrCode	LinkDirection	LinkName	LinkTypeCode	Length (mi)	BeginMilePost	EndMilePost	CSecCode	ID	DIR
1779	7	EB	Florida Blvd	1	0.2613	4.2880	4.5493	13-04	28035	1
1780	7	EB	Florida Blvd	1	0.1153	4.5493	4.6645	13-04	28063	1
1849	7	WB	Florida Blvd	1	0.1153	9.9939	10.1092	13-04	28069	1
1850	7	WB	Florida Blvd	1	0.2611	10.1092	10.3703	13-04	28049	1

Table LINK_TYPES

LinkTypeCode	Description
1	main
2	interchange
3	on-ramp
4	off-ramp
5	service road
6	short link

Table POSTED_SPEED_LIMITS

LinkCode	BeginMilePost	EndMilePost	SpeedLimitDate	PostedSpeedLimit (mph)
1779	4.2880	4.5493	07/31/97	50
1780	4.5493	4.6645	07/31/97	50
1849	9.9939	10.1092	07/31/97	50
1850	10.1092	10.3703	07/31/97	50

Table WEEK_DAYS

GPSDate	DayOfWeek
10/19/95	TH
10/20/95	FR
10/21/95	SA
10/22/95	SU
10/23/95	MO

Table DRIVERS

DrivInitials	Name	DateOfBirth	Occupation
MS	Driver No. 1	01/01/73	LSU Student
BP	Driver No. 2	01/01/73	LSU Student

Table GPS_FILE_STATUS

Status	Description
0	GPS file entry generated
1	Point geographic file generated
2	MI/MO operation completed
3	Linear referencing completed
4	GPS data imported into Access

Table GPS_FILE_PROPERTIES

GPSFileCode	GPSFileName	GPSFilePath	GPSDate	StartTime	EndTime	UTCOffset	NRecords	DrivInitials	Status
178	10191129.txt	G:ttg\gpsdata\95fall\10191129	10/19/95	41,695.75	49,111.75	-5	6140	MS	4
179	10192210.txt	G:ttg\gpsdata\95fall\10192210	10/19/95	80,184.25	82,286.25	-5	1,719	BP	4

Table DIF_COR_STATUS

DGPSStatus	Description
0	No differential correction
1	No differential correction
2	2-D differential correction
3	3-D differential correction

Table GPS_DATA

GPSFileCode	LocalTime	GPSTime	Longitude	Latitude	GPSSpeed	DGPSStatus	LinkCode	RouteCode	MilePost	Offset
178	27,173.75	45,173.75	-91.1183809	30.4514496	44.2	3	1779	07FLDBE	4.2910	8.0
178	27,174.75	45,174.75	-91.1181741	30.4514556	44.4	3	1779	07FLDBE	4.3033	8.8
:	:	:	:	:	:	:	:	:	:	:
178	27,190.75	45,190.75	-91.1147336	30.4515751	41.6	3	1779	07FLDBE	4.5084	9.5
178	27,193.75	45,193.75	30.4515929	-91.1141925	37.1	3	1779	07FLDBE	4.5407	9.1
178	27,194.75	45194.75	30.4515993	-91.1140204	36.9	3	1780	07FLDBE	4.5510	8.8
178	27,195.75	45195.75	30.4516051	-91.1138483	37.1	3	1780	07FLDBE	4.5612	8.4
:	:	:	:	:	:	:	:	:	:	:
178	27,207.25	45207.25	30.4516539	-91.1123446	25.5	3	1780	07FLDBE	4.6508	7.3
178	27,208.75	45208.75	30.4516611	-91.1121580	27.5	3	1780	07FLDBE	4.6620	6.9

Figure 8
Sample of records from the database

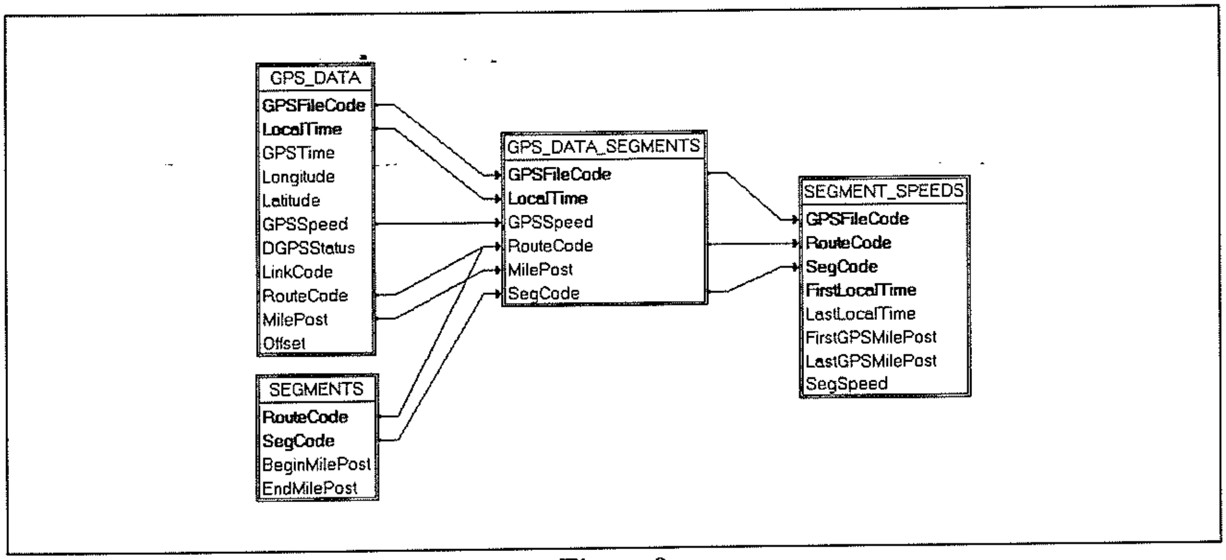


Figure 9
Schema for computation of segment speeds

While TransCAD is completing the linear referencing operation, it accesses table GPS_FILE_PROPERTIES in Access using an ODBC connection to determine which GPS files are to be processed and to update their status. However, because TransCAD does not allow users to append data to tables using ODBC connections (only update), we were unable to append table GPS_DATA in Access directly from TransCAD. As a result, we were forced to output all linearly referenced GPS data into *gpsdata.dbf* files (one for each GPS file), then import these *gpsdata.dbf* files into temporary tables in Access, and finally, append these temporary tables to table GPS_DATA.

Data Reduction Methodology

We developed an application called Travel Time with GPS (TTG) to automate the data reduction process needed to produce table GPS_DATA. TTG allows us to do the following:

- Filter the raw GPS data file so that we have a compact, comma-delimited American Standard Code for Information Interchange (ASCII) file that only contains coordinates, time, speed, and differential correction status. This comma-delimited file is the input data file for the rest of the process
- Import the filtered GPS data file into TransCAD to generate a point geographic file that we can overlay to the highway network vector map
- Specify when and where the vehicle entered and exited a study route using an animated GPS play back utility
- Calculate a milepost value for each GPS point along the route using the results from the GPS point marking process
- Import linearly referenced GPS data into temporary tables in Access and, then, append these temporary tables to table GPS_DATA.

Figure 10 shows a generic view of a typical work flow using TTG. For completeness, figure 10 shows both data reduction steps [boxes (1) through (19)] and data reporting steps [(20) through (23)]. In this section, we only discuss the data reduction steps (we discuss the data reporting steps in the next section). The first step involves preparing the raw GPS data file to make it ready for data reduction (1). First, we copy the raw GPS data file into an appropriate directory. It is convenient to define a separate directory for each GPS file because TTG always outputs the results from the linear referencing process for any GPS file to a file called *gpsdata.dbf* in the same directory as the GPS file. Second, we convert the raw GPS data into a compact, comma-delimited ASCII file that only contains coordinates, time, speed, and differential correction status. Third, we generate an entry for the GPS file in table GPS_FILE_PROPERTIES. At this point, the Status field in table GPS_FILE_PROPERTIES (figure 8) becomes 0.

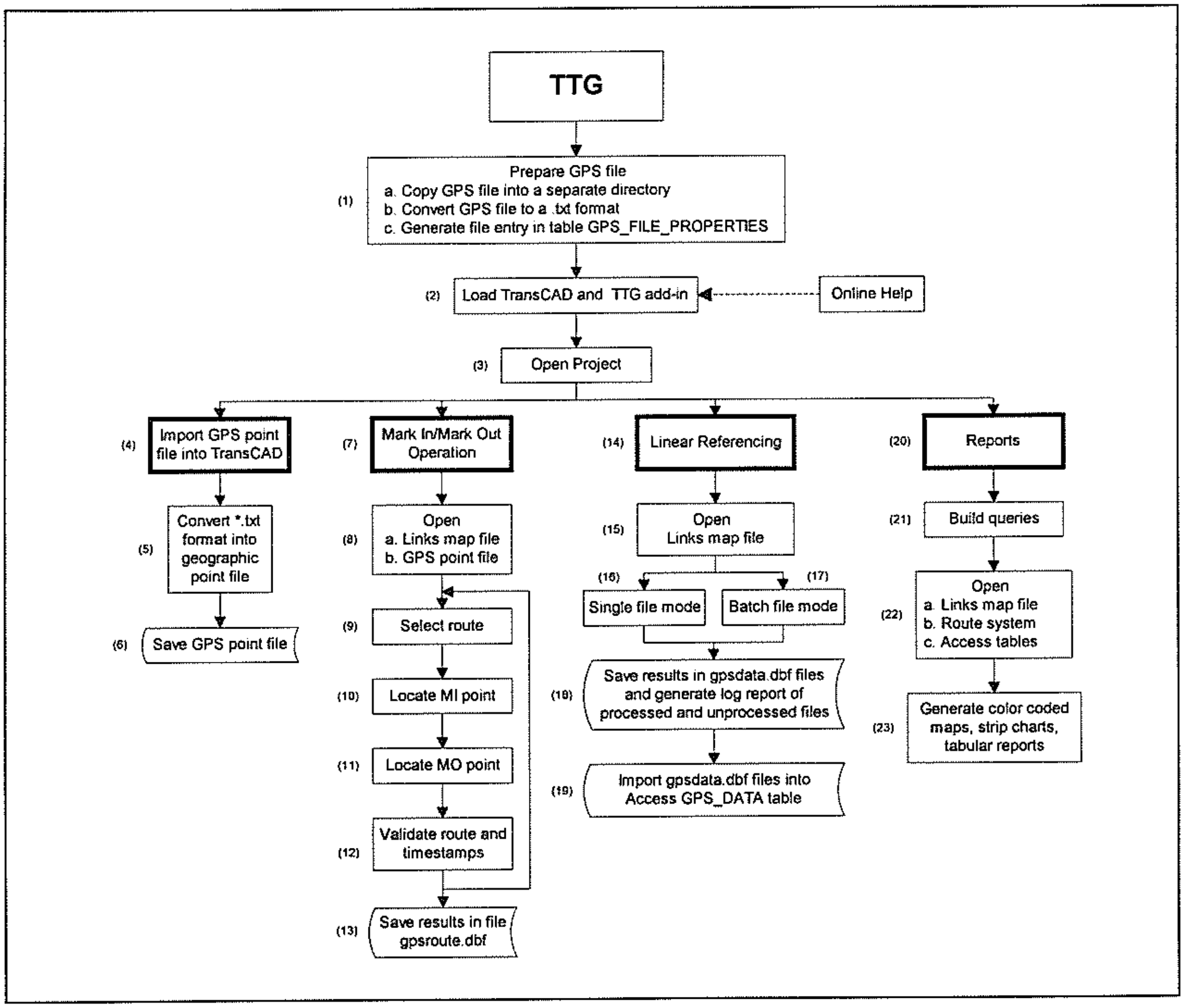


Figure 10
Data reduction workflow

The second step involves loading TransCAD and loading a TTG add-in (2). The TTG add-in is an application we developed in TransCAD's GIS Developer's Kit language (GISDK) to process GPS files in TransCAD. Like most applications written in GISDK, we load the TTG add-in from the Tools|Add-Ins menu option in TransCAD. This results in a new TransCAD menu called TTG from which we can run all TTG procedures. The companion CD contains the TTG add-in code and describes the process to install it and run it.

The third step involves opening a project file (3). This project file allows TTG to validate file paths and Access database tables (using an ODBC connection).

The fourth step involves generating a point geographic file in TransCAD based on the *.txt* GPS file (4) (5) (6). The result is a set of 11 TransCAD-generated files which, in addition to the *.txt* file, contain all the graphical and tabular information needed to display and process each GPS file. One of the those files, having a *.dbd* extension, is normally known as the point geographic file (or GPS point file) and can be opened in a TransCAD map at any time to view the location of all GPS points. Generating a point geographic file is a one-step operation which can be run at any time after opening a project by choosing the TTG|Point File|Create menu option. After generating a point geographic file, the corresponding Status field in table GPS_FILE_PROPERTIES automatically becomes one.

The fifth step involves assigning routes and specifying when and where the probe vehicle entered and exited study routes (7) (8) (9) (10) (11) (12) (13). For this, we use an animated GPS Player utility (figure 11). This utility allows us to play back a GPS point file and/or display any number of GPS points. First, we open the links map and a GPS point file (8). Second, we select a route covered by the probe vehicle (9) from the GPS Player route pull-down menu (figure 11). For example, we would select Florida Blvd EB for processing GPS points located along the Florida Boulevard EB route. Third, we click on the GPS Player MI button and then click on the first GPS point located on the route of interest (10). This causes the corresponding time stamp to appear on the GPS Player Start Time box (figure 11). Fourth, we click on the GPS Player MO button and then click on the last GPS point located on the route of interest (11). This causes the corresponding time stamp to appear on the GPS Player End Time box (figure 11). Fifth, we click on the GPS Player Accept button to validate the route and time stamps (12). TTG verifies that the route is correct by checking route codes associated with links in the vicinity of the MI and MO points, and also verifies that the time stamps are consistent (for example, that the MO time stamp is larger than the MI time stamp). If TTG does not detect any anomaly, it accepts the route and MI/MO time stamp pair and displays the corresponding results in the GPS Player Record History box (figure 11). At this point, we can either process other routes covered by the probe vehicle in the same GPS point file (9) (10) (11) (12) or save the results in file *gpsroute.dbf* (13) (table 4). After saving file *gpsroute.dbf*, TTG converts the Status field in table GPS_FILE_PROPERTIES (figure 8) to two.

For added flexibility, we developed the GPS Player utility so that we can define pairs of M<u>I</u> and M<u>O</u> points anywhere along a route and define more than one pair of M<u>I</u> and M<u>O</u> points per route. This is useful for partial route analysis and for filtering out spurious GPS points (for example, GPS points that result when the probe vehicle exits the study route momentarily). For example, in table 4, we could have two Florida Boulevard EB (07FLDBE) records between 44,534.25 and 45,565.25 instead of just one. A valid pair of entries would be, for instance, 44,534.25 - 45,208.75 and 45,315.25 - 45,565.25.

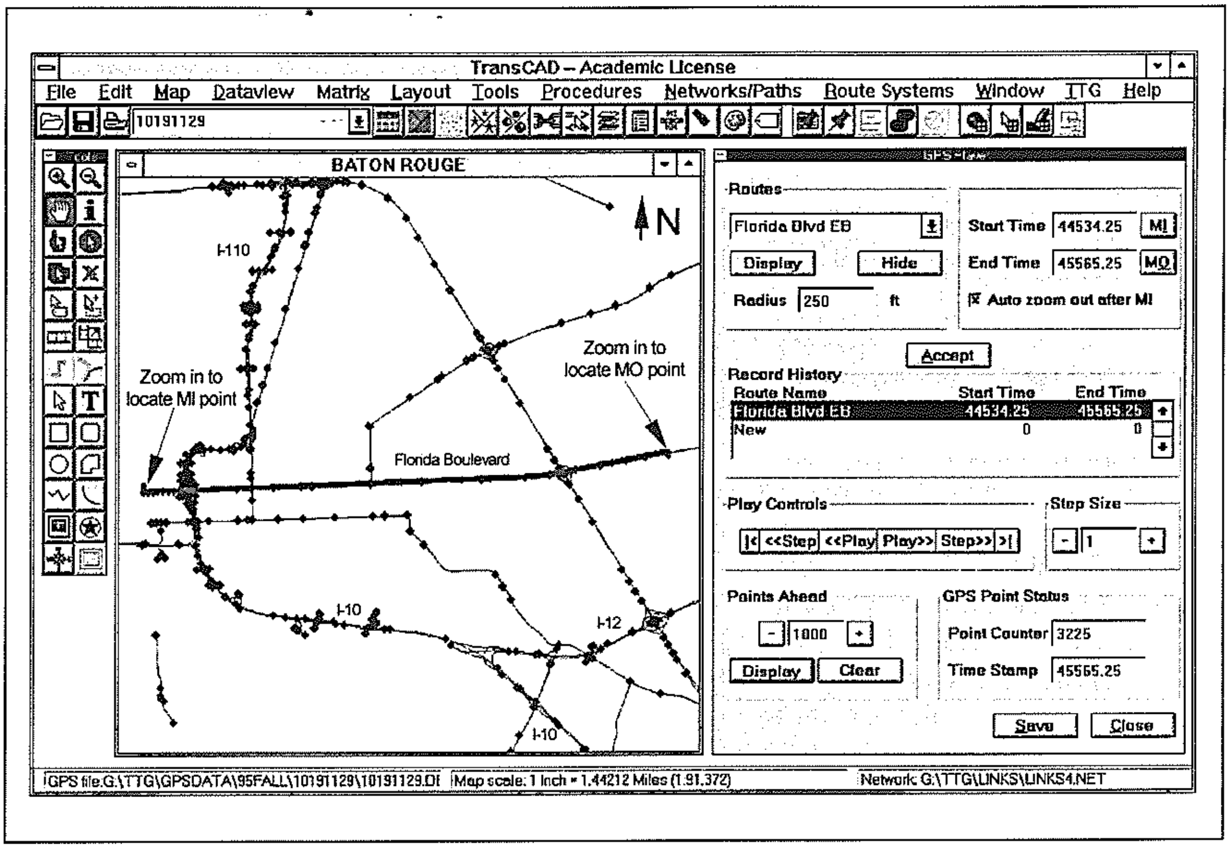


Figure 11
Links map, GPS point file, and TTG GPS Player utility

Table 4
***gpsroute.dbf* file associated with GPS file *10191129.txt* (figure 11)**

GPSFileCode	RouteCode	StartTime	EndTime
178	12GVSTW	42,068.25	42,107.75
178	07FLBDE	42,162.25	43,149.25
178	07FLBDW	43,377.25	44,280.25
178	07FLBDE	44,533.25	45,566.25
178	07FLBDW	45,631.25	46,622.25
178	07FLBDE	46,664.75	47,456.75
178	07FLBDW	47,579.75	48,476.75
178	18NCDRS	48,709.25	48,726.75

The sixth step involves linearly referencing all GPS points (14) (15) (16) (17) (18) (19). We do this by computing a milepost value for each GPS point along its assigned route by using the results from the GPS point marking process. For example, the milepost associated with the GPS point with a time stamp of 45,173.75 along Florida Boulevard EB is 4.2910 (first record in table GPS_DATA in figure 8). Likewise, the milepost associated with the GPS point with a time stamp of 45,208.75 along the same route is 4.6620 (last record in table GPS_DATA in figure 8). TTG calculates mileposts by mapping each GPS point to its assigned route. This process involves using the coordinates of each GPS point, finding the coordinates of the closest point on the route (i.e. mapping the GPS point to the route), and calculating the milepost associated with the mapped point. TTG also measures the horizontal distance between each GPS point and the corresponding mapped point on the route. This distance represents a link offset associated with each GPS point (Offset field in table GPS_DATA of figure 8) and provides an indication of the closeness of the GPS data to the highway network map. We use the Offset field to provide a verification of the positional accuracy of GPS points and to flag GPS points that may have been referenced to an incorrect route.

The linear referencing process is fully automated and only requires the user to open the links map (15). There are two modes of operation: single file mode (16) and batch file mode (17). In the single file mode, TTG calculates mileposts for any GPS point file manually opened by the user (this GPS point file must also have a corresponding *gpsroute.dbf* file). In the batch file mode, TTG automatically calculates mileposts for all GPS point files for which Status = 2 in table GPS_FILE_PROPERTIES. In either mode, TTG opens the *gpsroute.dbf* file associated with a GPS point file and the corresponding route system files. After processing a GPS point file, TTG saves the linearly referenced data in file *gpsdata.dbf* in the same directory as the GPS point file (18). When all GPS point files have been processed, TTG generates a log report of processed and unprocessed files. At this point, the user can exit TransCAD and load Access to append all *gpsdata.dbf* files to table GPS_DATA (19). This appending process is also fully automated and only requires the user to activate a macro. This macro is also included in the companion CD.

Data Reporting Methodology

After storing the linearly referenced GPS data in Access, as described in the previous section, we could build a variety of relational database queries to derive data needed for the production of graphical and tabular travel time and speed reports. In this section, we include a few database querying and data reporting examples to aid readers in identifying options which may be applicable to their own needs. For additional data reporting examples, readers are referred to other sources *[8]*.

Database queries

We can construct two types of queries: non-GIS queries and GIS queries. Non-GIS queries can be built and executed outside the GIS graphical environment. Examples of non-GIS queries are queries we use in Access to determine records associated with individual GPS files or average link/segment speed values per time period. In contrast, GIS queries require the use of GIS selection and location tools for their building and execution. Examples of GIS queries are some of the queries we use in TransCAD to generate color-coded maps.

While the number of queries that can be generated from the database is enormous, a few queries seem to be needed quite frequently. Some of the queries that fall into this category are the following:

- Selection of records for specific travel time runs or date ranges and time periods
- Computation of link/segment speed and travel time
- Computation of representative (maximum, minimum, average, median) link/segment speeds
- Computation of travel time delay
- Computation of speed and travel time at the section/route levels

Selection of records for specific travel time runs or date ranges and time periods

Suppose we need all records associated with GPS file *10191129.txt*. For this file, GPSFileCode = 178, as shown in table GPS_FILE_PROPERTIES of figure 8. The corresponding query, using Access notation, would be:

```
SELECT DISTINCTROW  GPS_DATA.*
              FROM  GPS_DATA
             WHERE  ((GPS_DATA.GPSFileCode = 178));
```

We could also use standard SQL notation. In this case, the query would be:

```
SELECT  *
  FROM  GPS_DATA
 WHERE  GPSFileCode = 178;
```

Access allows users to build queries either by typing the entire query text or by using a graphical query by example (QBE) interface (figure 12). These two options are interchangeable. For clarity, we will only include the graphic QBE version.

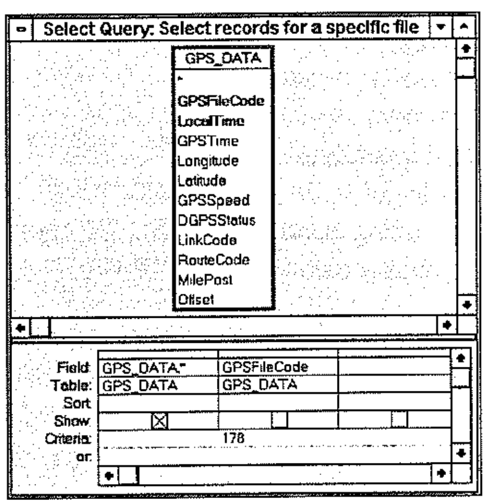


Figure 12
Query to select GPS records associated with a GPS file

Suppose now that we need all weekday GPS records from 7:00 - 8:00 a.m. during the academic year 1995-1996. Figure 13 shows the corresponding query.

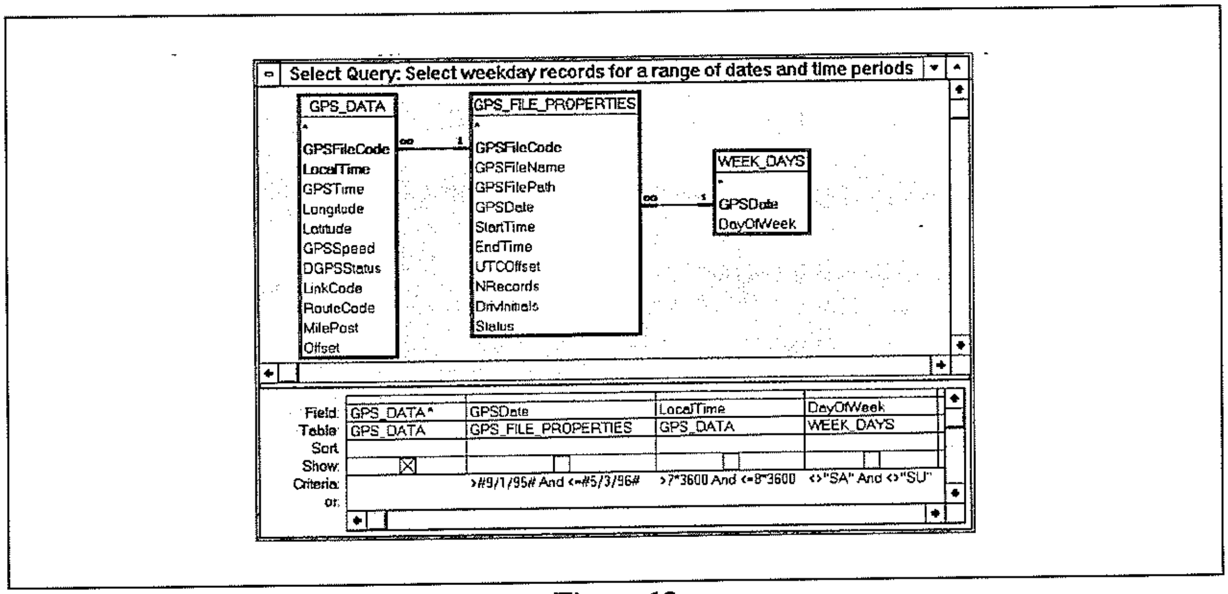


Figure 13

Query to select weekday records for a range of dates and time periods

Computation of link/segment speed and travel time

Suppose we need to compute link speeds and travel time based on the GPS data from file *10191129.txt* (table GPS_DATA in figure 8). This process involves defining the GPS points that are associated with each link and then using a mathematical procedure to compute link speeds and travel time. For each GPS point we have three pieces of information: milepost (or distance traveled), time stamp, and speed. Theoretically, we only need two of the three quantities to completely define the problem. However, because speed is computed independently of positional fixes by the GPS receiver and has its own associated errors, using two of the quantities, e.g. speed and time stamp, to estimate the third quantity–distance traveled–does not necessarily produce the same results as those included in table GPS_DATA. Defining which two quantities to use depends on their relative accuracy and the resulting effect on the computation of link speeds and travel time. Appendix A describes a generalized mathematical model that can be used for this purpose.

We developed a form in Access called SegmentSpeeds to assist users in this process (figure 14). To compute link speeds and travel times we input the GPS data table name, say GPS_DATA, choose one of the procedures listed, and select a time threshold. As described in appendix A, the first procedure, Link or Segment Length/Interpolated Travel Time, computes link speeds based on link lengths and interpolated time stamps at the link entrance and exit points. The second procedure, Speed Integration, computes link speeds based on the instantaneous speeds of all GPS points located within a link. The time threshold is used to check for discontinuities in the input GPS data table and allows form SegmentSpeeds to group GPS records that are associated with the same link and belong to the same time stamp sequence. With this information, form SegmentSpeeds computes link speeds and outputs the results to a table called

SEGMENT_SPEEDS_x, where x represents the number associated with the procedure chosen. For example, if we choose the first procedure, the output table would be called SEGMENT_SPEEDS_1. Grouping GPS data by link code and time stamp sequence is necessary so that GPS records that were collected on the same link but that belong to a different run (either a different GPS file or a different loop within the same GPS file) are not mixed together.

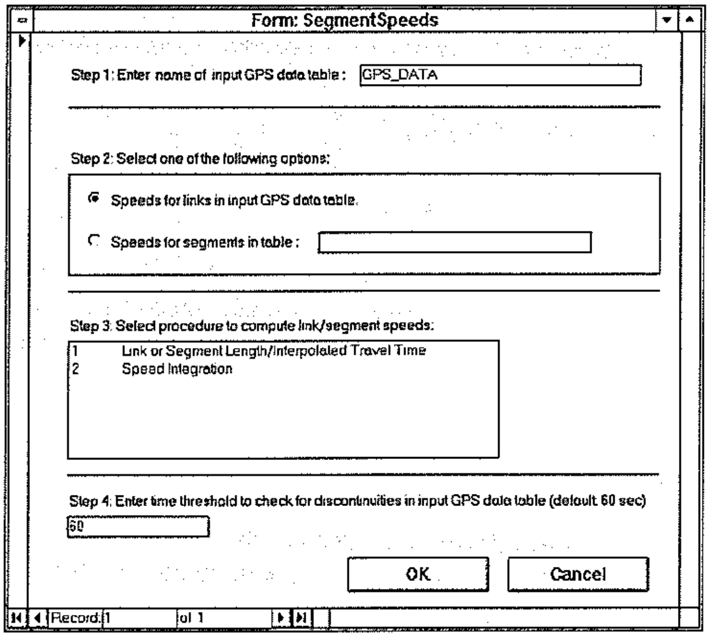


Figure 14
Form to compute link/segment speeds

Form SegmentSpeeds can accept any input table name (or query name) as long as the structure and data are of the same type as table GPS_DATA. This is particularly useful for computing link speeds for a subset of records from table GPS_DATA. For example, for computing link speeds for all GPS records included in file *10191129.txt* (GPSFileCode = 178), we could build a query or table called GPS_DATA_178. Form SegmentSpeeds would then use this name to compute only the link speeds for all GPS records associated with file *10191129.txt*. Table 5 shows a sample of resulting link speeds.

Table 5
Link speeds for GPS records associated with file *10191129.txt* (GPSFileCode = 178)

GPSFileCode	RouteCode	LinkCode	BeginLocalTime	LinkSpeed
178	07FLBDE	1779	27,173.51	44.63
178	07FLBDE	1780	27,194.58	28.62
178	07FLBDE	1781	27,209.08	15.24
178	07FLBDE	1782	27,320.83	35.66

Form SegmentSpeeds can also compute speeds and travel times for segments which follow any highway segmentation scheme. Following the process described in figure 9, in addition to the GPS data table, say GPS_DATA, we would need a table, say SEGMENTS, containing the beginning and ending mileposts associated with each highway segment. For example, table SEGMENTS in figure 15 shows a sample of segment milepost data for short segments located on Florida Boulevard (these segments are the same as those which we had in first phase of the CMS project). To enter this table in form SegmentSpeeds (figure 14), we would select the second option under Step 2 (to activate the text box), and then we would type in the table name. Based on tables GPS_DATA and SEGMENTS, the form would produce an intermediate table called GPS_DATA_SEGMENTS and, depending on the speed procedure chosen, a SEGMENT_SPEEDS_x table (figure 9 and figure 15).

Computation of representative link/segment speeds

Suppose we need to compute representative link/segment speeds (maximum, minimum, average, median). The input table for this process is a segment speed table, say SEGMENT_SPEEDS_1, that results from aggregating GPS point speed data into link or segment speed data.

Using standard query functions, we can compute minimum, average, and maximum link/segment speeds. Figure 16 shows the corresponding query. Note in the cell that computes average link/segment speed (AvgSpeed) that we use a harmonic mean speed formulation which is consistent with the speed aggregation process discussed in appendix A. Notice also that the query shown in figure 16 does not include a cell for computing median link/segment speeds. Like other database packages, Access lacks a function to compute median values directly. As a result, it became necessary to develop a utility to do it. For completeness, we developed a form in Access called MinMaxAvgMedianSegmentSpeeds that calculates not only median speeds but also minimum, maximum, and harmonic mean speeds (figure 17). In a typical application, we would input the name of a table or query containing link/segment speeds and the name of the desired output table (the structure of the input table has to be the same as that of any SEGMENT_SPEEDS_x table discussed previously). The form would then sort records by link/segment, calculate the number of records, minimum speed, maximum speed, average (harmonic mean) speed, and median speed for each link/segment, and store the results in the user-defined output table.

Computation of travel time delay

Suppose we need to compute average speed deficits and travel time delays per link/segment. Link/segment speed deficit is defined as the difference between a reference speed, say free flow speed, and a representative link/segment speed, say average speed. Likewise, link/segment travel time delay is defined as the difference between a representative link/segment travel time, say average travel time, and a reference travel time, say free flow travel time.

Table GPS_DATA

GPSFileCode	LocalTime	GPSTime	Longitude	Latitude	GPSSpeed	DGPSStatus	LinkCode	RouteCode	MilePost	Offset
178	27,173.75	45,173.75	-91.118381	30.451450	44.2	3	1779	07FLDBE	4.2910	8.0
178	27,174.75	45,174.75	-91.118174	30.451456	44.4	3	1779	07FLDBE	4.3033	8.8
:	:	:	:	:	:	:	:	:	:	:
178	27,182.25	45,182.25	-91.11657	30.451511	47.4	3	1779	07FLBDE	4.3989	10.2
178	27,183.25	45,183.25	-91.116349	30.451518	47.6	3	1779	07FLBDE	4.4121	10.6
178	27,184.75	45,184.75	-91.116014	30.451532	47.8	3	1779	07FLBDE	4.4321	10.2
178	27,185.25	45,185.25	-91.115902	30.451537	47.8	3	1779	07FLBDE	4.4388	9.9
178	27,186.75	45,186.75	-91.115568	30.451548	47.3	3	1779	07FLBDE	4.4587	10.2
178	27,187.25	45,187.25	-91.115458	30.451552	46.8	3	1779	07FLBDE	4.4653	10.2
:	:	:	:	:	:	:	:	:	:	:
178	27,190.75	45,190.75	-91.1147336	30.4515751	41.6	3	1779	07FLDBE	4.5084	9.5
178	27,193.75	45,193.75	30.4515929	-91.1141925	37.1	3	1779	07FLDBE	4.5407	9.1
178	27,194.75	45194.75	30.4515993	-91.1140204	36.9	3	1780	07FLDBE	4.5510	8.8
178	27,195.75	45195.75	30.4516051	-91.1138483	37.1	3	1780	07FLDBE	4.5612	8.4
:	:	:	:	:	:	:	:	:	:	:
178	27,207.25	45207.25	30.4516539	-91.1123446	25.5	3	1780	07FLDBE	4.6508	7.3
178	27,208.75	45208.75	30.4516611	-91.1121580	27.5	3	1780	07FLDBE	4.6620	6.9

Table SEGMENTS

RouteCode	SegCode	BeginMilePost	EndMilePost
07FLBDE	13144	3.7490	3.8888
07FLBDE	13145	3.8888	4.0884
07FLBDE	13146	4.0884	4.2880
07FLBDE	14129	4.2880	4.4185
07FLBDE	14130	4.4185	4.5493
07FLBDE	11908	4.5493	4.6645
07FLBDE	13150	4.6645	4.8013
07FLBDE	13151	4.8013	4.9380
07FLBDE	13152	4.9380	5.1376
07FLBDE	13153	5.1376	5.2687
07FLBDE	13164	5.2687	5.4683
07FLBDE	12041	5.4683	5.6212
07FLBDE	12042	5.6212	5.6627
07FLBDE	11793	5.6627	5.8352
07FLBDE	13186	5.8352	5.9605

Table GPS_DATA_SEGMENTS

GPSFileCode	LocalTime	GPSSpeed	RouteCode	MilePost	SegCode
178	27173.75	44.2	07FLBDE	4.2910	14129
178	27174.75	44.4	07FLBDE	4.3033	14129
:	:	:	:	:	:
178	27182.25	47.4	07FLBDE	4.3989	14129
178	27183.25	47.6	07FLBDE	4.4121	14129
178	27184.75	47.8	07FLBDE	4.4321	14130
178	27185.25	47.8	07FLBDE	4.4388	14130
:	:	:	:	:	:
178	27190.75	41.6	07FLBDE	4.5084	14130
178	27193.75	37.1	07FLBDE	4.5407	14130
178	27194.75	36.9	07FLBDE	4.5510	11908
178	27195.75	37.1	07FLBDE	4.5612	11908
:	:	:	:	:	:
178	27207.25	25.5	07FLBDE	4.6508	11908
178	27208.75	27.5	07FLBDE	4.6620	11908

Table SEGMENT_SPEEDS_1

GPSFileCode	RouteCode	SegCode	BeginLocalTime	SegSpeed
178	07FLBDE	14129	27173.75	45.97
178	07FLBDE	14130	27184.75	43.36
178	07FLBDE	11908	27194.75	28.62

Figure 15
Computation of segment speeds

Free flow speeds measurements may be based on posted speed limits, maximum observed speeds, or runs conducted specifically for this purpose. Posted speed limits are easy to use but tend to be unreliable because motorists routinely drive faster than the posted speed limit if traffic conditions permit it. Maximum observed speeds have the advantage that all other observed speeds are lower, but may produce unrealistic results if the maximum observed speeds are actually outliers. Speeds based on runs conducted specifically for measuring free flow speeds, say late at night or on Sunday mornings, are conceptually sound but, in some cases, are lower than weekday speeds. One possible explanation for this phenomenon is that drivers tend to drive less aggressively late at night or on Sunday mornings and are, therefore, more willing to drive at speeds closer to the posted speed limits.

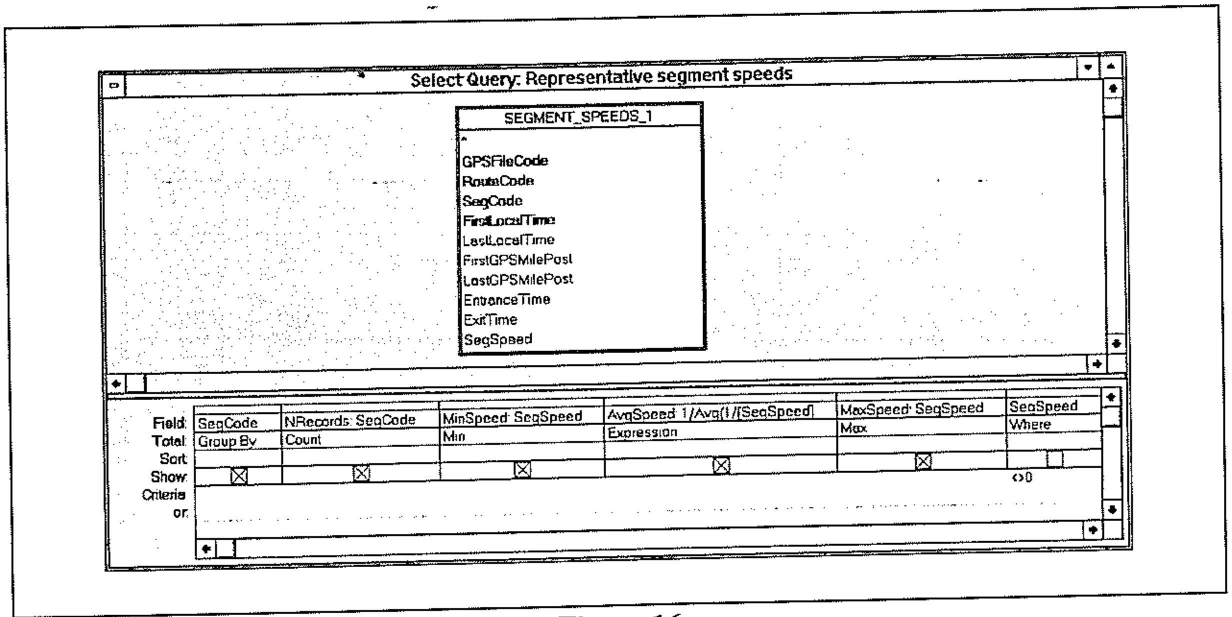


Figure 16
Query to compute minimum, average, and maximum link/segment speeds

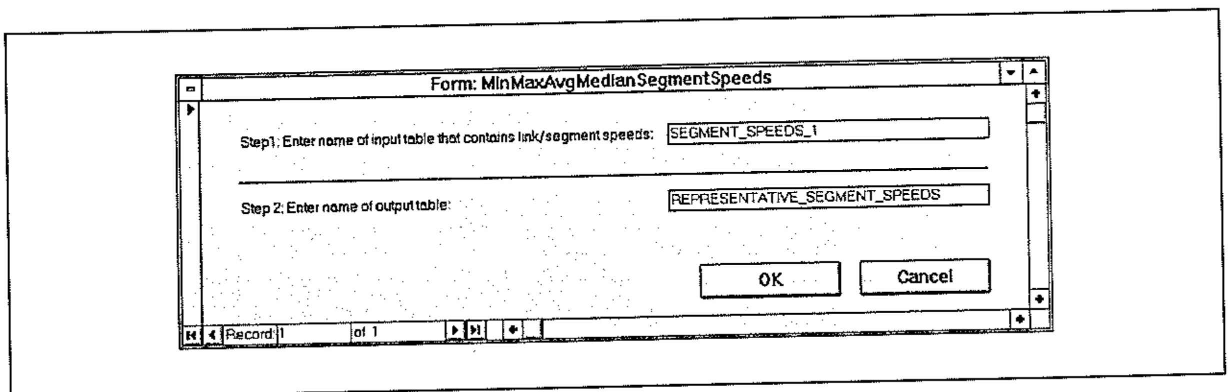


Figure 17
Form to compute minimum, maximum, average (harmonic mean), and median link/segment speeds

Assume, for the sake of simplicity, that Sunday morning runs are chosen to represent free flow speeds. Figure 18 shows the corresponding query. Notice that this query also computes free flow travel times. To estimate travel time we need a table, such as LINKS or SEGMENTS, which contains the beginning and ending mileposts of each linear element. The difference between the ending milepost and the beginning milepost for each link/segment gives us the link/segment length which, in turns, allows us to calculate link/segment free flow travel time (in seconds).

In order to compute link/segment speed deficit and travel time delay, we join the queries of figures 16 and 18 and calculate speed and travel time differences. Figure 19 shows the corresponding query.

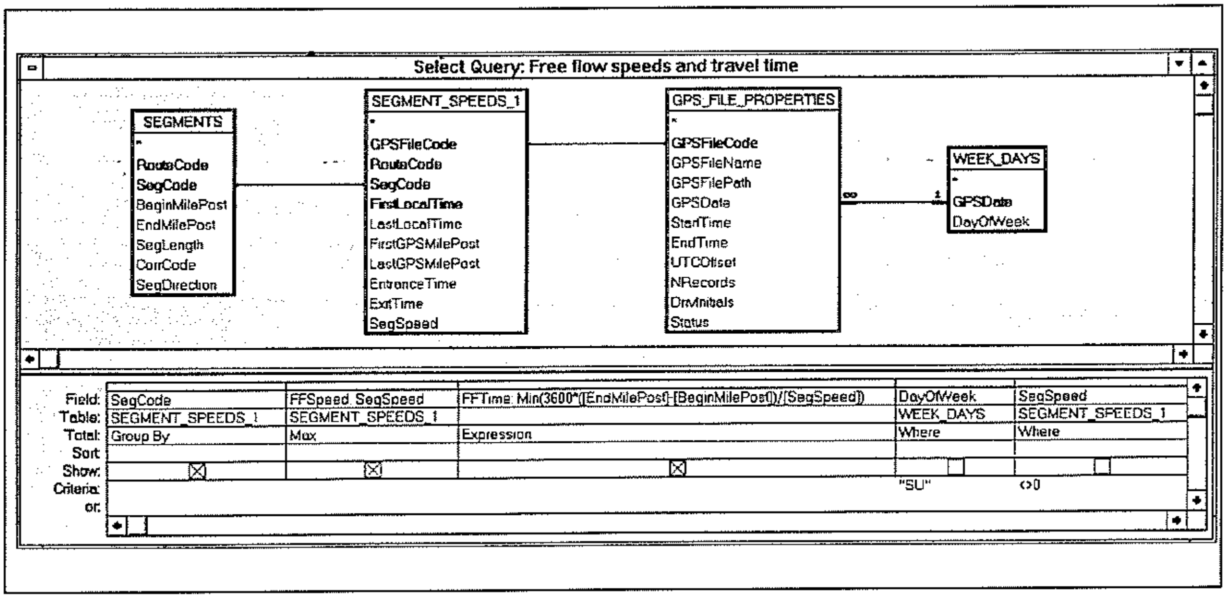


Figure 18
Query to compute free flow speed and travel time by link/segment

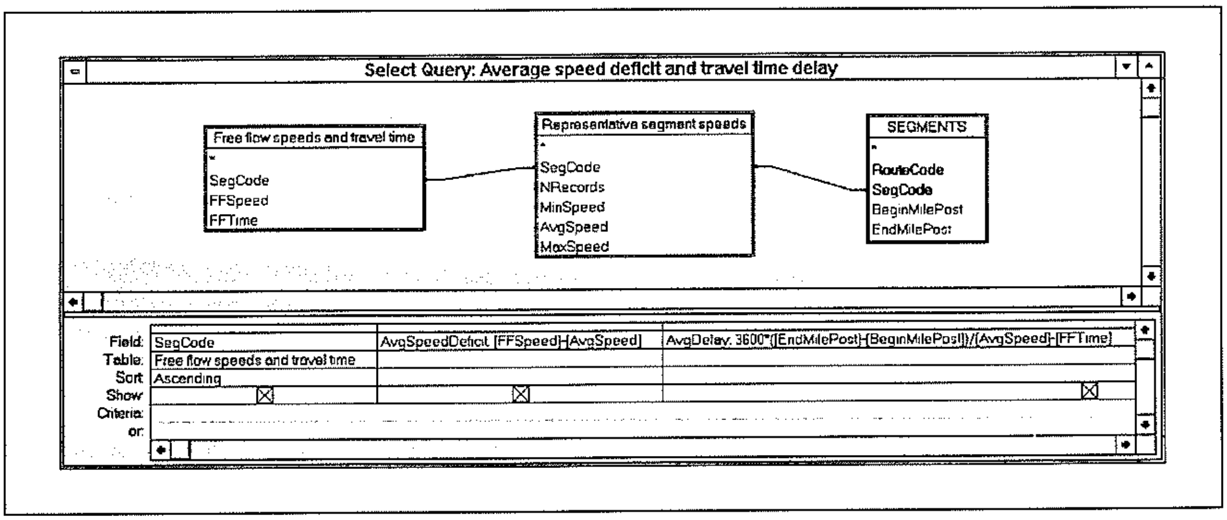


Figure 19
Query to compute average speed deficit and travel time delay by link/segment

Computation of speed and travel time at the section/route levels

Suppose we need to compute average speed and cumulative travel time for a section (or even for an entire route (figure 5). The procedure to follow depends on whether or not all GPS point files included in the analysis cover the entire section or route. If they do, we could actually construct tables similar to table SEGMENTS in figure 15, with "long" segments having the same beginning and ending mileposts as the sections or routes under consideration. Then we would use form SegmentSpeeds (figure 14) to compute "long" segment speeds, and, finally, we would compute average "long" segment speeds using the query shown in figure 16.

If not all GPS point files included in the analysis cover the entire section or route, we would need to do the following: (1) compute speeds for short segments (say 0.2 mi long) or links (figure 14), (2) compute average short segment/link speeds (figure 16), (3) calculate average short segment/link travel time, (4) compute a cumulative travel time for the section or route, and (5) calculate an overall average speed. As an illustration, figure 20 shows the query needed to complete steps (3) through (5), assuming we need to compute average speed (in mph) and cumulative travel time (in min) at the route level for all routes in the network.

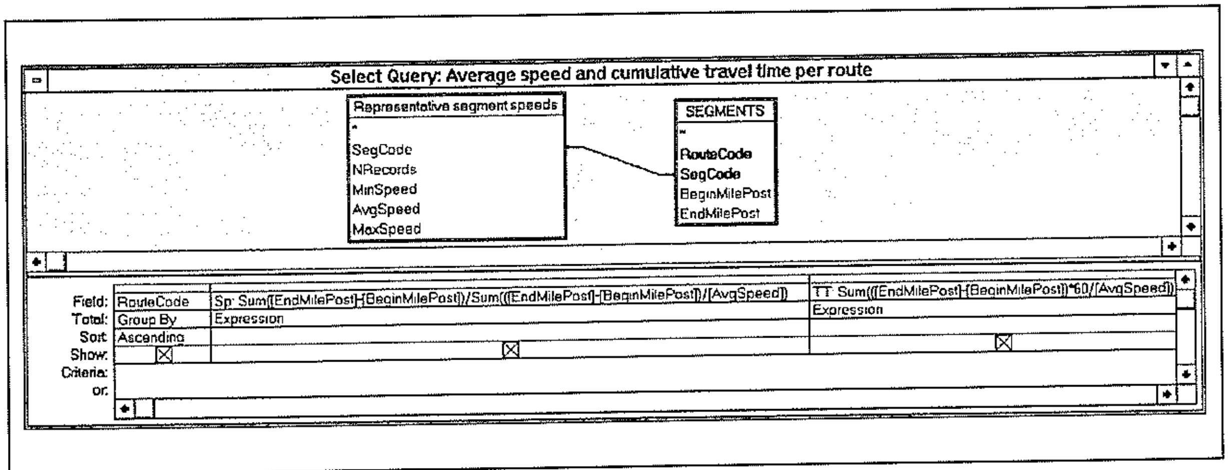


Figure 20
Query to compute average speed and travel time at the route level

Data reporting options

As mentioned previously, there are many travel time and speed data reporting possibilities available to analysts *[8]*. Here, we consider the following three reporting options: color-coded maps, archival tabular reports, and distance-time diagrams.

Color-coded maps

In the first phase of the CMS project, we used GIS tools to produce color-coded maps showing average speeds on all segments of the highway network (figure 21) *[2]*. For producing these maps, we built queries in Oracle (the database package in which we had all segment travel time data) to compute average segment speeds, and then we grouped these segment speeds by speed range (0-20 mph, 20-30 mph, and so on) to create map layers using MGE.

We can still build maps in MGE using segment speeds obtained with the procedures discussed in this report. This option provides maximum flexibility to MPOs because it allows them to produce and overlay speed maps in the same graphical environment they have used in recent years. As shown in figures 9, 14, 15, and 16, all that is required is a table like table SEGMENTS that contains the route and the beginning and ending mileposts associated with each segment on the network (figure 9, figure 15). Then we would use form SegmentSpeeds (figure 14) and the query to compute representative segment speeds (figure 16). Finally, we would export the resulting average segment speed table to Oracle.

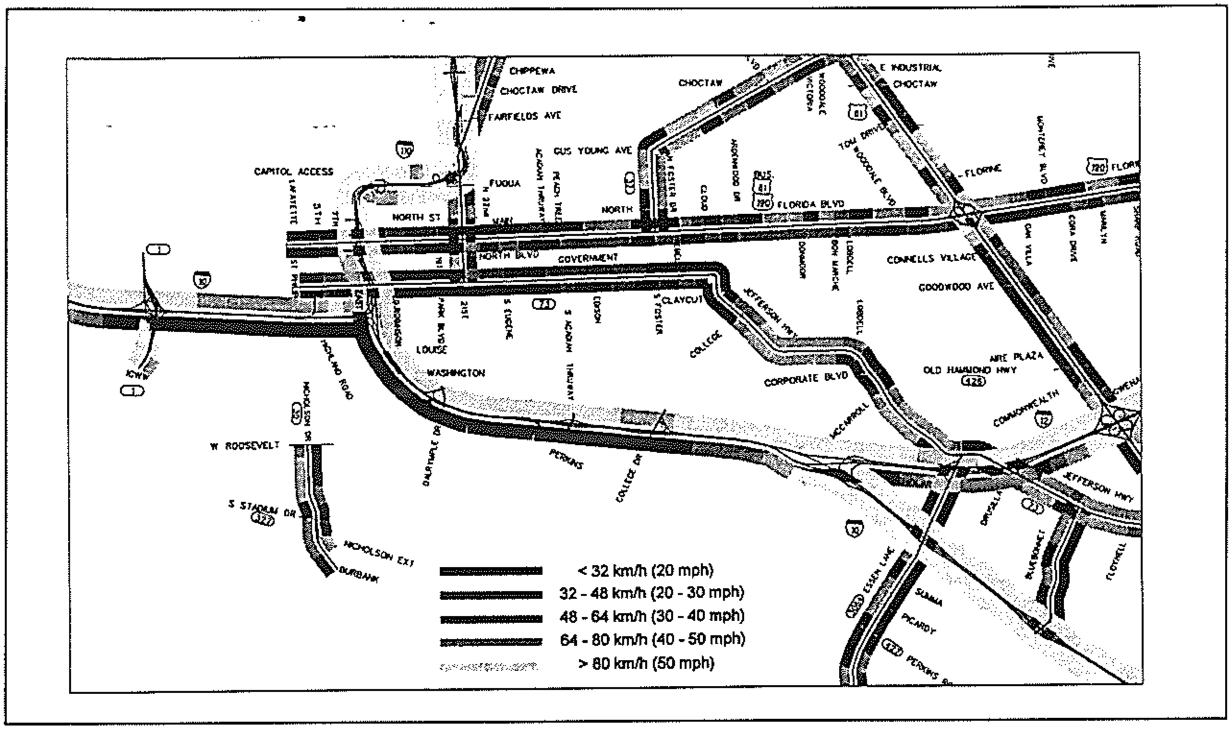


Figure 21
P.M. peak (4:30-5:30 p.m.) average speeds in Baton Rouge during the 1995-1996 academic year

With TransCAD, we can produce color coded maps and "strip charts." These charts show the variation of attributes along routes using profile-type diagrams. For example, we can produce both original GPS speed maps and aggregated speed maps for any segmentation scheme (segments, links, sections, or routes). As an illustration, figure 22 shows a sample map with original GPS speeds from the runs made on Florida Boulevard on October 19, 1995 (file *10191129.txt* in figure 8). The strip chart shown corresponds to the eastbound run. Figure 23 shows the corresponding maps and strip charts for segment aggregated speeds and link aggregated speeds.

Archival tabular reports

In the first phase of the CMS project, we developed a procedure in Access to document average speeds and cumulative travel time for all CMS segments in the highway network using a tabular report format *[2]*. Figure 24 shows a sample report page. We can still use this procedure using segment speeds obtained with the dynamic segmentation approach discussed in this report. All that is required is to export a table such as table SEGMENT_SPEEDS_1, which contains CMS segment speeds, to the Access database used to produce archival tabular reports.

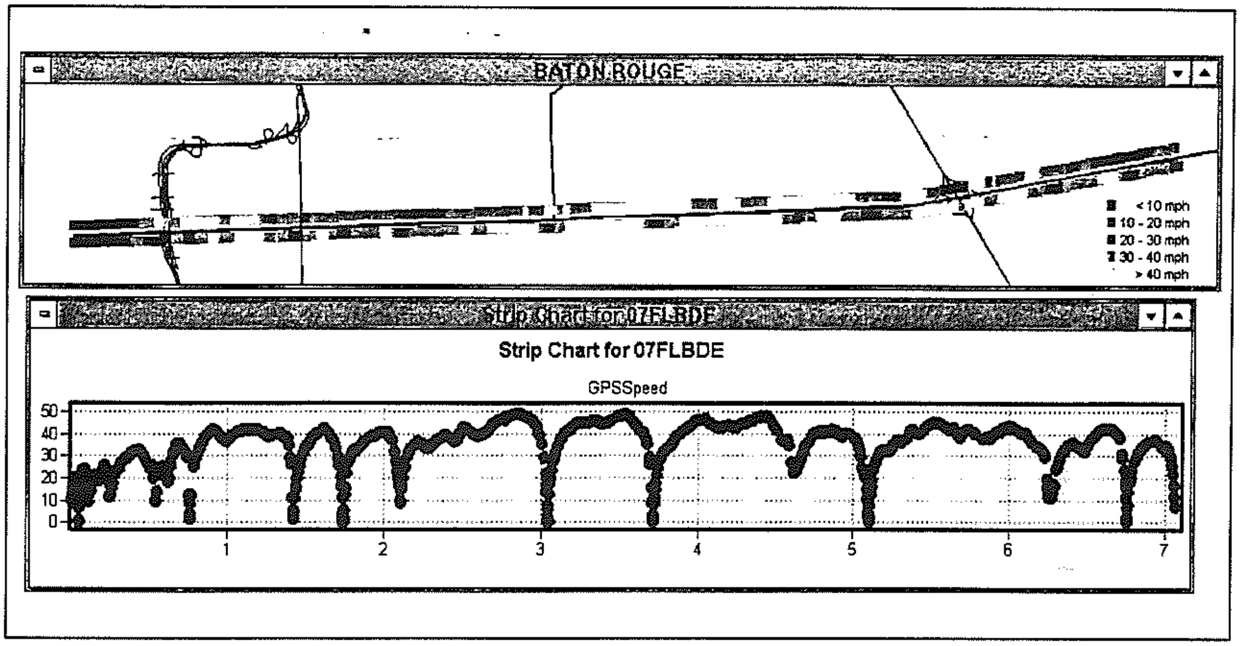


Figure 22
GPS speeds for runs made on Florida Boulevard in Baton Rouge on October 19, 1995

Distance-time diagrams

Because each GPS point is assigned a unique milepost value, we can construct distance-time diagrams like that shown in figure 25. Notice in figure 25 points a-l indicate the location and time of major changes in speed and acceleration. For example, points a, e, and i indicate when the vehicle started to decelerate to join the standing queues at the North Foster, Ardenwood, and Wooddale signalized intersections. Points b, f, and j indicate when the vehicle stopped. Points c, g, and k indicate when the vehicle started to accelerate. Finally, points d, h, and l indicate when the vehicle stopped accelerating, i.e. when the vehicle speed became essentially uniform.

Locating critical points such as points a-l accurately is essential for computing various performance measures like traffic delay, control delay, stopped delay, running speed, average speed, and speed deficit. However, current techniques that measure or compute these performance measures generally involve significant errors because of the difficulty in determining the actual location of the critical points (points a-l). To address this difficulty, we developed procedures that analyze acceleration-time diagrams and automatically determine the location and time associated with critical acceleration points. Details of the theory and algorithms behind the procedure are included in appendix B.

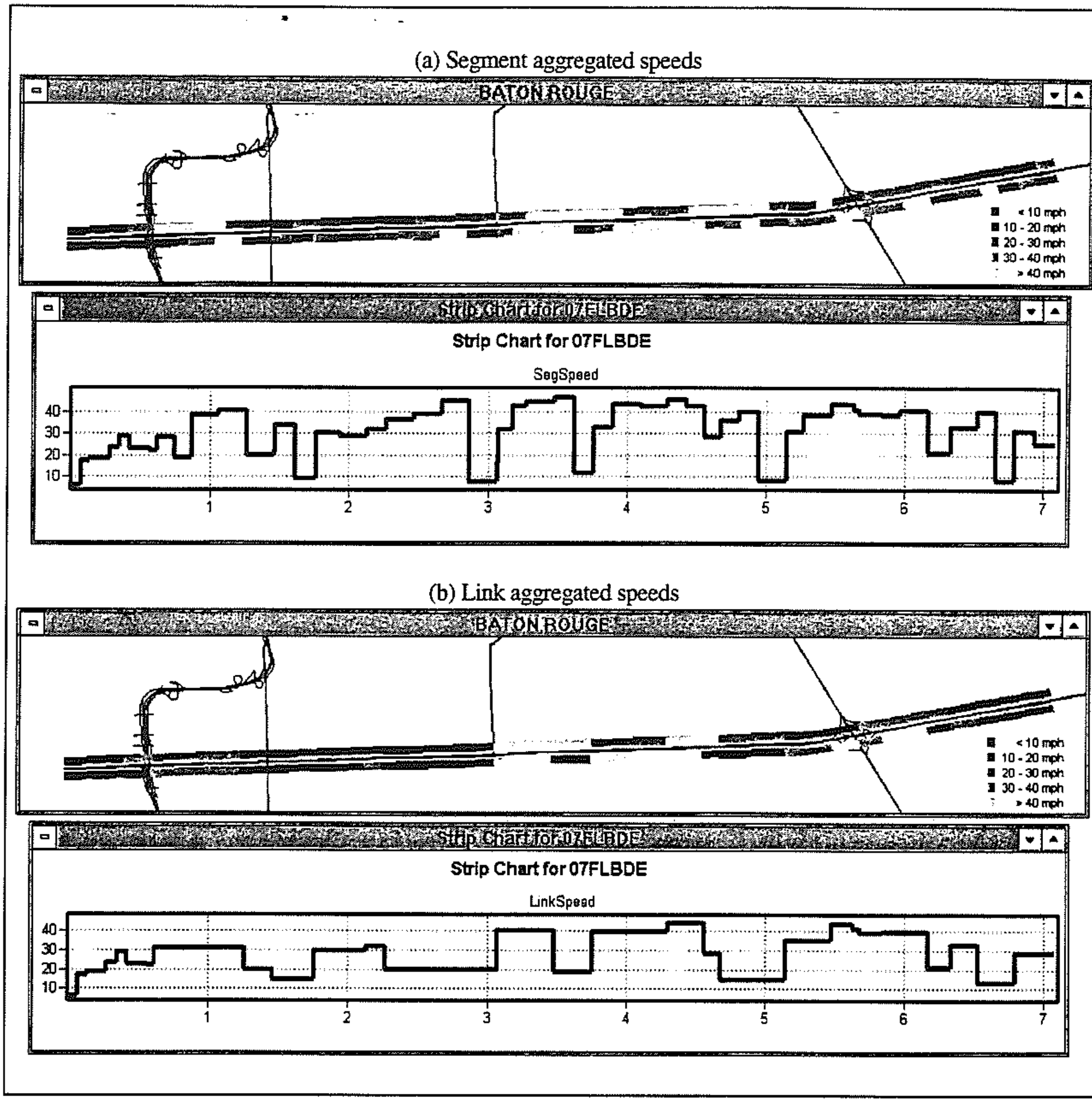
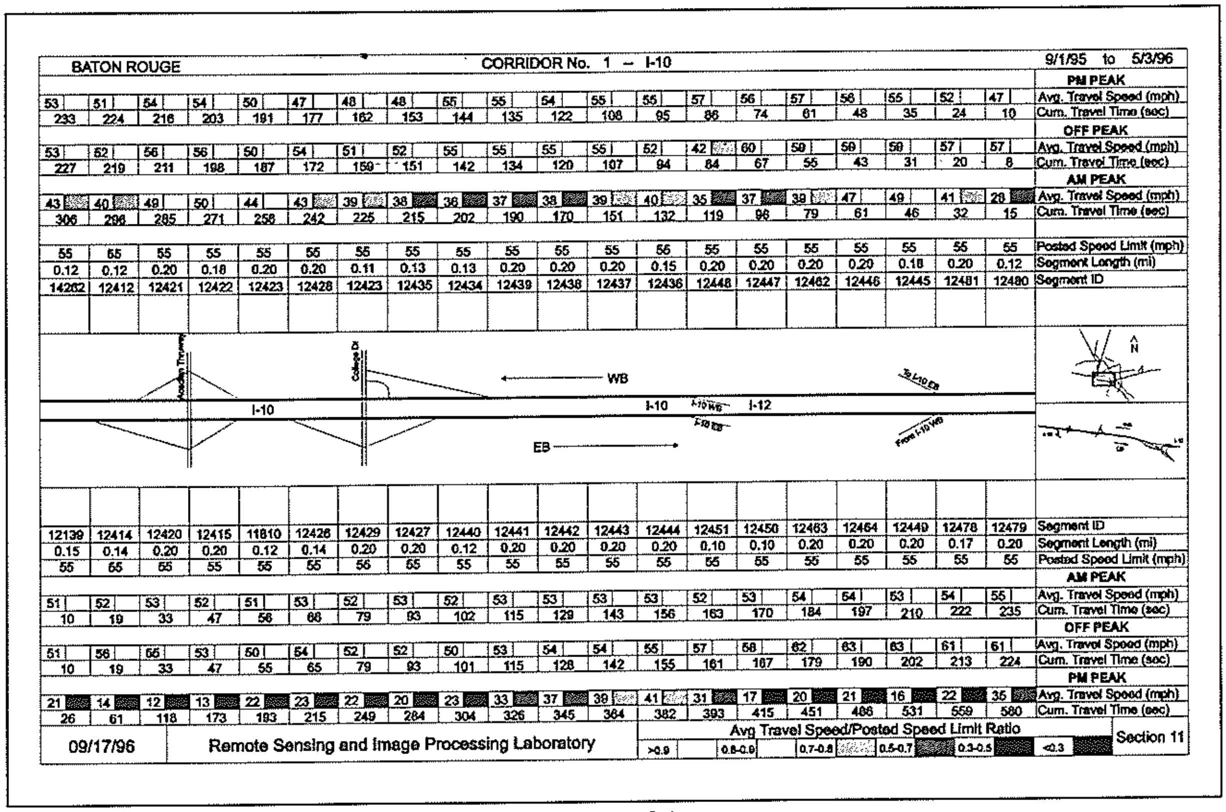


Figure 23
Segment speeds and link speeds for runs made on Florida Boulevard in Baton Rouge on October 19, 1995



BATON ROUGE										CORRIDOR No. 1 – I-10										9/1/95 to 5/3/96
																				PM PEAK
53	51	54	54	50	47	48	48	55	55	54	55	55	57	56	57	56	55	52	47	Avg. Travel Speed (mph)
233	224	216	203	191	177	162	153	144	135	122	108	95	86	74	61	48	35	24	10	Cum. Travel Time (sec)
																				OFF PEAK
53	52	56	56	50	54	51	52	55	55	55	55	52	42	60	59	59	59	57	57	Avg. Travel Speed (mph)
227	219	211	198	187	172	159	151	142	134	120	107	94	84	67	55	43	31	20	8	Cum. Travel Time (sec)
																				AM PEAK
43	40	49	50	44	43	39	38	36	37	38	39	40	35	37	39	47	49	41	28	Avg. Travel Speed (mph)
306	296	285	271	258	242	225	215	202	190	170	151	132	119	96	79	61	46	32	15	Cum. Travel Time (sec)
55	65	55	55	55	55	55	55	55	55	55	55	56	55	55	55	55	55	55	55	Posted Speed Limit (mph)
0.12	0.12	0.20	0.18	0.20	0.20	0.11	0.13	0.13	0.20	0.20	0.20	0.15	0.20	0.20	0.20	0.20	0.18	0.20	0.12	Segment Length (mi)
14262	12412	12421	12422	12423	12428	12423	12435	12434	12439	12438	12437	12436	12448	12447	12462	12446	12445	12481	12480	Segment ID

12139	12414	12420	12415	11810	12426	12429	12427	12440	12441	12442	12443	12444	12451	12450	12463	12464	12449	12478	12479	Segment ID
0.15	0.14	0.20	0.20	0.12	0.14	0.20	0.20	0.12	0.20	0.20	0.20	0.20	0.10	0.10	0.20	0.20	0.20	0.17	0.20	Segment Length (mi)
55	65	65	55	55	55	56	65	55	55	55	55	55	55	55	55	55	55	55	55	Posted Speed Limit (mph)
																				AM PEAK
51	52	53	52	51	53	52	53	52	53	53	53	53	52	53	54	54	53	54	55	Avg. Travel Speed (mph)
10	19	33	47	56	66	79	93	102	115	129	143	156	163	170	184	197	210	222	235	Cum. Travel Time (sec)
																				OFF PEAK
51	56	55	53	50	54	52	52	50	53	54	54	55	57	58	62	63	63	61	61	Avg. Travel Speed (mph)
10	19	33	47	55	65	79	93	101	115	128	142	155	161	167	179	190	202	213	224	Cum. Travel Time (sec)
																				PM PEAK
21	14	12	13	22	23	22	20	23	33	37	39	41	31	17	20	21	16	22	35	Avg. Travel Speed (mph)
26	61	118	173	193	215	249	284	304	326	345	364	382	393	415	451	486	531	559	580	Cum. Travel Time (sec)

09/17/96	Remote Sensing and Image Processing Laboratory	Avg Travel Speed/Posted Speed Limit Ratio: >0.9, 0.8-0.9, 0.7-0.8, 0.5-0.7, 0.3-0.5, <0.3	Section 11

Figure 24
Example archival report page

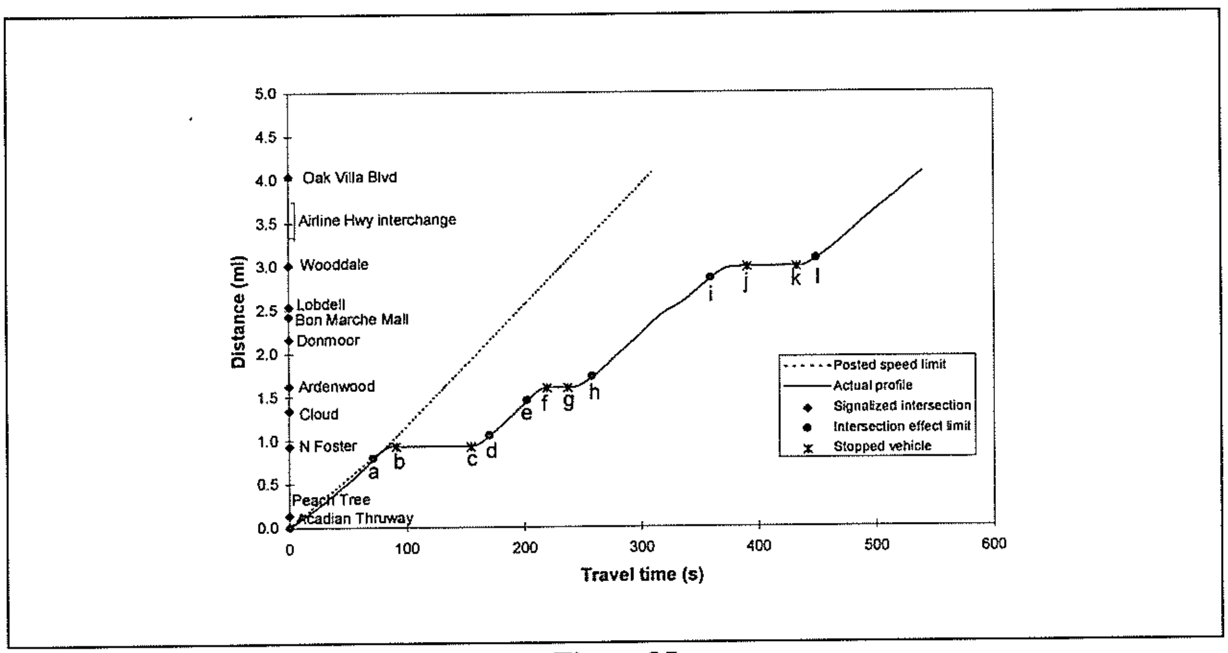


Figure 25
Distance-time diagram for a run made on Florida Boulevard EB in Baton Rouge, on October 19, 1995 (file *10191129.txt* in figure 8)

DISCUSSION OF RESULTS

We tested the GPS-GIS methodology described in the previous sections using data collected in Baton Rouge in 1995 and 1996, i.e. the same data we used for developing the CMS architecture in Baton Rouge *[2]*. Because we used similar architectures for New Orleans and Shreveport, the analysis and discussion of results included in this section can easily be applied to those metropolitan areas. For the most part, they can also be applied to other metropolitan areas because the procedures developed here are general.

The discussion of results includes the following topics: spatial model, data management and data reduction, data reporting, and evaluation of the TransCAD-Access environment.

Spatial Model

The Baton Rouge MPO and the Capital Region Planning Commission (CRPC) defined a congestion corridor network composed of 22 corridors covering 151 mi (figure 26). As shown in table 6, of the 22 corridors, three were located on the Interstate highway system and covered 45 mi. The remaining 19 corridors were located on principal arterials and covered 106 mi.

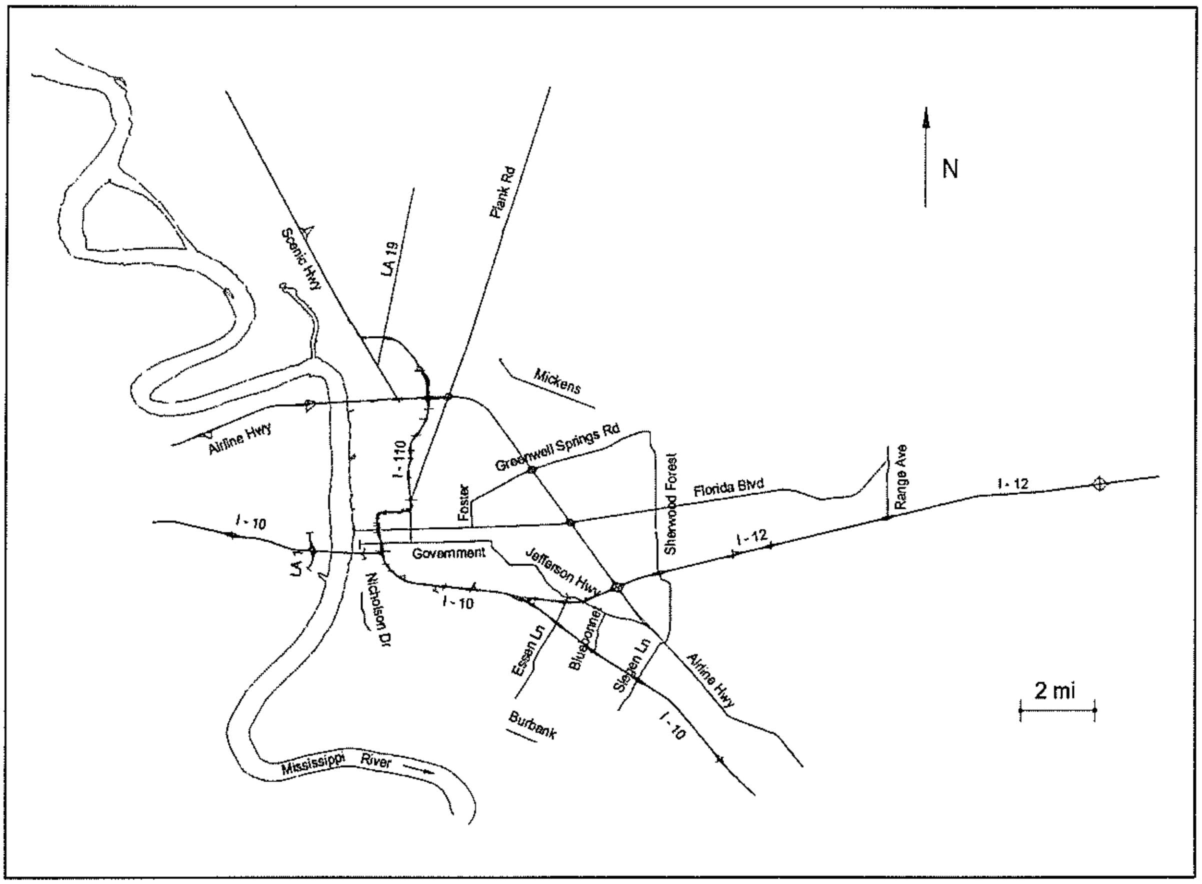


Figure 26
Congestion corridor network in Baton Rouge

Table 6
Congestion corridor network in Baton Rouge

ID	Name	From	To	Dir.	Functional class	Length (mi)	
1	I-10	2 mi West Of LA 415	Ascension Parish line	E-W	Interstate	19.0	
2	I-110	I-10, I-110 Split	Scenic Hwy	N-S	Interstate	8.7	
3	I-12	I-10 &-12 Split	1.5 mi East of LA 447	E-W	Interstate	17.6	
4	La 19	Scenic Hwy	Wimbush Lane	N-S	Princ. Arterial		5.6
5	Plank Rd	Government St	LA 64	N-S	Princ. Arterial		14.7
6	Airline Hwy	LA 1145	2 mi East of Highland Rd	N-S	Princ. Arterial		22.9
7	Florida Blvd	River Front	Range Ave	E-W	Princ. Arterial		14.7
8	Mickens Rd	Hooper Rd	Joor Rd	E-W	Princ. Arterial		3.0
9	Sherwood Forest Blvd	Greenwell Springs Rd	Airline Hwy	N-S	Princ. Arterial		6.8
10	Siegen Lane	Airline Hwy	Perkins Rd	N-S	Princ. Arterial		2.5
11	N Foster Dr	Florida Blvd	Greenwell Springs Rd	N-S	Princ. Arterial		0.8
12	Government St	River Front	Jefferson Hwy	E-W	Princ. Arterial		3.4
13	Jefferson Hwy	Airline Hwy	Government St	E-W	Princ. Arterial		5.4
14	Staring Lane	Perkins Rd	Highland Rd	N-S	Princ. Arterial		2.0
15	Essen Lane	Perkins Rd	Jefferson Hwy	N-S	Princ. Arterial		1.9
16	Bluebonnet Rd	Jefferson Hwy	I-10	N-S	Princ. Arterial		1.3
17	Burbank Dr	Gardere Ln	Bluebonnet Rd	E-W	Princ. Arterial		0.9
18	Nicholson Dr	Roosevelt St	Burbank Dr	N-S	Princ. Arterial		1.2
19	LA 1	I-10	ICWW	N-S	Princ. Arterial		1.1
20	Greenwell Springs Rd	N Foster Dr	Sherwood Forest Blvd	E-W	Princ. Arterial		5.2
21	Scenic Hwy	Airline Hwy	LA 64	N-S	Princ. Arterial		10.7
22	Range Ave	Florida Blvd	I-12	N-S	Princ. Arterial		2.3
			TOTAL			45.3	106.4

To generate the map of figure 26 in TransCAD, we imported the MGE *.dgn* file we had created in the first phase of the CMS project using Caliper's Intergraph Import Add-In. This add-in converted all MGE segments into TransCAD linear features and assigned them a unique identification number. The result was a line geographic file (and accompanying index and table files) which contained all basic data associated with the CMS segments. The conversion process worked reasonably well for most segments, i.e. TransCAD maintained their location, topology, and attribute information intact. At several locations, particularly interchanges and overpasses, TransCAD broke the original MGE segments. These segments had been the result of a joining operation in Microstation, for example, after they had been accidentally broken at points where segments appear to intersect on the screen. Conversely, segments which had not undergone any editing (mainly joining) in Microstation were not affected during the import operation into TransCAD.

For this project, we joined contiguous segments (and rejoined broken segments) to form links, i.e. linear features connecting major physical discontinuities. We also assigned a unique identification code to each link (LinkCode) by using a different field than that used by TransCAD to index graphical features automatically (ID). As mentioned previously, if for any reason a user needs to compact a geographic file in TransCAD (i.e. converting a standard geographic file into a compact file and then back into a standard geographic file), TransCAD redefines new ID values for all graphical features. By comparison, LinkCode values remain the same for all links.

For each corridor in table 6, we defined two routes: one for each direction. TransCAD allows users to define all routes in a single route system file. As we explain with greater detail in the evaluation of TransCAD and Access subsection, we found a bug in TransCAD that prevents users from locating linear features using neighborhood operations reliably. For this reason, we decided to create a separate route system file for each route in the network. This approach did not significantly complicate our procedure but did provide us with better control on the mapping operation. Table 7 summarizes the routes we defined in Baton Rouge.

Table 7
TransCAD routes in Baton Rouge

Code	Name	# links	Route length (mi)		
			TransCAD	MGE	Difference
01I10E	I-10 EB	26	19.0205	19.0314	0.0109
01I10W	I-10 WB	27	19.0673	19.0786	0.0112
02I110N	I-110 NB	26	8.8523	8.8381	-0.0142
02I110S	I-110 SB	27	8.5980	8.5839	-0.0141
03I12E	I-12 EB	18	17.5271	17.5562	0.0291
03I12W	I-12 WB	18	17.5920	17.6211	0.0291
04LA19N	LA 19 NB	12	5.6370	5.6213	-0.0157
04LA19S	LA 19 SB	13	5.6410	5.6128	-0.0282
05PLKN	Plank Rd NB	33	14.6778	14.6371	-0.0407
05PLKS	Plank Rd SB	34	14.6787	14.6379	-0.0407
06AIRN	Airline Hwy NB	53	22.8694	22.8669	-0.0025
06AIRS	Airline Hwy SB	53	22.8763	22.8737	-0.0025
07FLBDE	Florida Blvd EB	41	14.6642	14.6874	0.0232
07FLBDW	Florida Blvd WB	41	14.6632	14.6864	0.0231
08MKRDE	Mickens Rd EB	2	3.0107	3.0124	0.0017
08MKRDW	Mickens Rd WB	2	3.0072	3.0089	0.0017
09SHBDN	Sherwood Forest Blvd NB	15	6.7669	6.7598	-0.0070
09SHBDS	Sherwood Forest Blvd SB	15	6.7549	6.7392	-0.0157
10SGLNN	Siegen Lane NB	8	2.4806	2.4654	-0.0152
10SGLNS	Siegen Lane SB	8	2.4834	2.4685	-0.0149
11NFDRN	N Foster Dr NB	3	0.7735	0.7713	-0.0022
11NFDRS	N Foster Dr SB	3	0.7841	0.7818	-0.0023
12GVSTE	Government St EB	17	3.4263	3.4331	0.0068
12GVSTW	Government St WB	17	3.4384	3.4452	0.0068
13JFHYE	Jefferson Hwy EB	13	5.4074	5.4090	0.0016
13JFHYW	Jefferson Hwy WB	13	5.4157	5.4173	0.0016
14STLNN	Staring Lane NB	1	1.9907	1.9867	-0.0040
14STLNS	Staring Lane SB	1	1.9913	1.9872	-0.0041
15ESLNN	Essen Lane NB	9	1.8659	1.8554	-0.0105
15ESLNS	Essen Lane SB	9	1.8617	1.8512	-0.0105
16BBRDN	Bluebonnet Rd NB	4	1.3185	1.3108	-0.0077
16BBRDS	Bluebonnet Rd SB	4	1.3172	1.3095	-0.0078
17BKDRE	Burbank Dr EB	1	0.9193	0.9202	0.0008
17BKDRW	Burbank Dr WB	1	0.9195	0.9204	0.0009
18NCDRN	Nicholson Dr NB	3	1.2481	1.2448	-0.0032
18NCDRS	Nicholson Dr SB	3	1.2508	1.2476	-0.0032
19LA1N	LA 1 NB	2	1.1468[1]	0.3855	-0.7613
19LA1S	LA 1 SB	2	1.1383[1]	0.4487	-0.6896
20GSRDE	Greenwell Springs Rd EB	17	5.2324	5.2340	0.0016
20GSRDW	Greenwell Springs Rd WB	17	5.2373	5.2427	0.0054
21SCHYN	Scenic Hwy NB	16	10.7027	10.6812	-0.0215
21SCHYS	Scenic Hwy SB	14	10.6729	10.6512	-0.0217
22RNAVN	Range Ave NB	8	2.3005	2.2862	-0.0143
22RNAVS	Range Ave SB	8	2.2993	2.2850	-0.0143

[1] These routes were extended in this project to include links on both sides of I-10.

When TransCAD created the line geographic file containing highway links, it added a length field to the list of link attributes. This field is automatically generated and populated by TransCAD. At the link resolution level, lengths computed by TransCAD were very similar to lengths computed by MGE. However, when we added link lengths to compute route lengths, we did observe significant differences between the two GIS packages. This is shown in table 7. For example, the length of the Florida Boulevard eastbound route was 14.6642 mi according to TransCAD and 14.6874 mi according to MGE. The difference was 0.0232 mi (or 37 m), which was much larger than the positional accuracy of the GPS data (2-5 m). It may be worth noting that TransCAD, by default, displays lengths with only two decimal digits. To view lengths with more decimal digits, we had to use a formula field in which lengths were multiplied by a factor of 1,000.

The differences between the two GIS packages can be traced back to the way each package models the surface of the Earth (which is not explicitly described in either user's manual). According to Intergraph officials, MGE uses a variable radius ellipsoidal model developed by Pittman *[9]*. According to Caliper officials, TransCAD uses a constant radius spherical model. Because variable radius ellipsoidal models take into consideration the eccentricity of the surface of the Earth explicitly, they compute surface distances more accurately than constant radius spherical models. To better substantiate this point, we measured distances between two pairs of points using both GIS packages and compared the results with those obtained using two independent algorithms: Maling's algorithm and Vincenty's algorithm *[10]*, *[11]*. As table 8 shows, distances measured with MGE were much closer to those measured with Maling and Vincenty's algorithms.

Table 8
Evaluation of TransCAD and MGE distance measurements

Point	Longitude	Latitude	Distance (mi)			Differences (mi)	
			Maling/Vincenty (1)	TransCAD (2)	MGE (3)	(2) - (1)	(3) - (1)
A	-91.097603	30.455195					
			3.14234	3.14681	3.14171	0.00447	-0.00063
B	-91.069908	30.416412					
C	-91.154725	30.450297					
			3.01946	3.01285	3.01878	-0.00661	-0.00068
D	-91.104177	30.451918					

The TransCAD distance measuring problem is a relatively minor issue as long as we remain in the TransCAD environment when computing link speeds and travel times, i.e. as long as we use TransCAD-derived linearly referenced GPS data with TransCAD-derived checkpoint milepost data. The reason is that TransCAD shows the location of GPS points correctly with respect to links and highway checkpoints. Unfortunately, this is not true 100 percent of the time. A bug in TransCAD causes some GPS points to be mapped to the wrong relative position along the route; this, in turn, causes the assigned milepost and link code to be incorrect. We address this issue in the evaluation of TransCAD and Access subsection. Obviously, there is an implicit error in the computation of link speeds and travel times because of the error in the computation of distances. However, these errors are relatively small for individual links.

The TransCAD distance measuring problem becomes an issue when mixing TransCAD-derived linearly referenced data with MGE-derived checkpoint milepost data. A typical situation would be trying to compute segment speeds based on TransCAD-derived linearly referenced GPS data using MGE-derived CMS segment beginning and ending milepost data. As we move along the route under study, the offset between the two milepost referencing systems tends to grow, eventually causing some GPS points to be assigned to wrong segments. This would be particularly troublesome in the case of signalized intersections (a vehicle may appear to be stopped after a signalized intersection). To address this situation, at least until Caliper solves the TransCAD distance measuring problem, we recommend adjusting all MGE-derived segment beginning and ending milepost data using TransCAD derived checkpoint milepost data as anchor points. Table 9 illustrates the procedure. We could then use the adjusted MGE segment milepost data with the TransCAD-derived GPS milepost data to compute segment speeds and travel time. After Caliper solves the TransCAD distance measuring problem, we would recommend changing the value of the Status field in table GPS_FILE_PROPERTIES to two for all GPS files, running the linear referencing process again using the batch mode option, and re-importing all *gpsdata.dbf* files to Access. We would expect the overhead of this operation to be low because the batch linear referencing procedure and the Access data importing procedures are fully automated procedures.

Table 9
Adjustment of MGE-derived CMS segment mileposts

Unadjusted CMS segment data				TransCAD checkpoint data			Adjusted CMS segment data			
Segment code	Milepost Beginning	Milepost Ending	Length (mi)	Link code	Milepost Beginning	Milepost Ending	Segment code	Milepost Beginning	Milepost Ending	Length (mi)
				1777	3.4680	3.7490				
13144	3.7564	3.8965	0.1401				13144	3.7490	3.8888	0.1398
13145	3.8965	4.0965	0.2000				13145	3.8888	4.0884	0.1996
13146	4.0965	4.2965	0.2000	1778	3.7490	4.2880	13146	4.0884	4.2880	0.1996
14129	4.2965	4.4272	0.1308				14129	4.2880	4.4185	0.1305
14130	4.4272	4.5583	0.1310	1779	4.2880	4.5493	14130	4.4185	4.5493	0.1308
11908	4.5583	4.6738	0.1155	1780	4.5493	4.6645	11908	4.5493	4.6645	0.1152
13150	4.6738	4.8108	0.1370				13150	4.6645	4.8013	0.1368
13151	4.8108	4.9478	0.1370				13151	4.8013	4.9380	0.1367
13152	4.9478	5.1478	0.2000	1781	4.6645	5.1376	13152	4.9380	5.1376	0.1996

Data Management and Data Reduction

We tested the data management system and the data reduction application using GPS data from 428 files we collected during the first phase of the CMS project. Appendix C summarizes dates, file names, number of GPS point records, and percentage of differential correction for all GPS data files. In total, we collected 2.4 million GPS records on 25,000 miles of travel time runs on the 151-mi highway network. Of the 2.4 million GPS records collected, 1.8 million GPS records were located on the main routes. The remaining 0.6 million GPS records were located on other parts of the network including service roads, on-ramps, off-ramps, and intersecting streets. We stored all 1.8 million linearly referenced GPS point records in table GPS_DATA in the Access database file. The resulting size of this database file, including the other tables shown in figure 8, was 238 Mbytes. This number translates to approximately 132 bytes per linearly referenced GPS record.

Processing and storing 1.8 million linearly referenced GPS records in a relational database allowed us to improve and optimize the data reduction application, develop generalized data quality control checks, and detect limitations and bugs of the developing platform (for a discussion about this topic, see the evaluation of TransCAD and Access subsection).

Data reduction application

We developed the data reduction application trying to give users as much flexibility as possible without jeopardizing database integrity. Because the application is expected to be used in an environment in which tens or hundreds of travel time runs may be executed, it was important to design the application using a centralized database approach. This approach requires some extra effort on the part of users when setting up projects and at the beginning and end of the data reduction process (mainly to generate entries for the GPS data file in Access and to import and append *gpsdata.dbf* files to Access). However, we believe this extra effort is worth the benefits the system provides, particularly with respect to the ability to build comprehensive queries to generate travel time reports and the ability to conduct generalized data quality control checks.

For planning purposes, it is important to estimate processing times and data storage requirements for individual GPS data files. Processing times depend on human factors, hardware and software characteristics, number of routes covered, and whether files are processed individually from beginning to end or in batches. Data storage requirements depend on the number of GPS points collected and, to a lesser degree, on the number of routes covered. As an illustration, table 10 shows the processing times and data storage required for processing file *10191129.txt*. This file contains 6,140 GPS records (of which 4,702 are located on the main routes) collected during eight runs on four routes over a two-hour period. The processing times reflect the performance of an average operator and the use of a Windows NT 3.51 PC with a single 100 MHz Pentium processor and 32 Mbytes of RAM. The first 15-17 minutes (i.e. until completing the MI/MO operation) are operator dependent, meaning the operator needs to focus his/her attention on the data reduction process. After this initial period, the application runs automatically, and, consequently, operator dependence is either null or minimal. From a practical point of view, the clear distinction between high operator dependent activities and null/minimal operator dependent activities means it may be advisable to process as many files as possible up to the MI/MO operation during normal office hours, and then leave the remaining two activities to be processed unattended (for example, at night or on weekends).

A data reduction time of 15-17 minutes for two hours worth of GPS data indicates a data reduction speed of about eight minutes per hour of data collection. By comparison, traditional data reduction procedures based on manual data collection procedures require about one hour of data reduction per hour of data collection *[8]*. In other words, the GPS/GIS data reduction procedure described in this report is about 7.5 times faster than traditional manual data reduction procedures.

Table 10
Processing time and data storage requirements for file *10191129.txt*

Activity	Processing time (min)	Operator dependence	Data storage requirement (Kbytes)	Comment
Download GPS file from field laptop to network/local hard drive	1-2 min	High	270.2	
Generate *.txt* GPS data file	< 1 min	High	299.8	This file could be deleted at the end
Generate entry in table GPS_FILE_PROPERTIES	1-2 min	High	0.2	
Generate point geographic file	2 min	High	982.8	These files could be deleted at the end
Execute MI/MO operation	10 min	High	6.7	
Subtotal	**15-17 min**		**1,559.7**	
Execute linear referencing	35 min	Null/minimal	974.9	These files could be deleted at the end
Import table *gpsdata.dbf* to Access	10-30 min [1]	Null/minimal	507.9	
Subtotal	**45-65 min**		**1,482.8**	
TOTAL	**60-82 min**		**3,042.5**	

[1] Actual processing time depends on the number of files and records already imported into Access.

As shown in table 10, the original 6,140 record GPS data file required 270.2 Kbytes of disk space. This file was a regular text file in ASCII format. Because of TransCAD's file system configuration, TransCAD required 1.29 Mbytes of disk space more to generate the point geographic file (for a total of 1.56 Mbytes). The linear referencing process for 4,702 useful GPS points required 974.9 Kbytes (659.5 Kbytes for file *gpsdata.dbf* and 315.4 Kbytes for index file *gpsdata.mdx*). Importing and appending file *gpsdata.dbf* to table GPS_DATA in Access required 507.9 Kbytes. In total, processing the original 270.2 Kbyte GPS data file required about 3.0 Mbytes of disk space.

To save disk space, it may be possible to delete most intermediate files after finishing the data reduction process (actually, after verifying that all GPS data conforms to specified data quality requirements). However, at the very minimum, the following files should be kept: the original GPS data file and file *gpsroute.dbf* (which results from the MI/MO operation). If the need arises, all other files can be easily regenerated from these two files.

Data quality control checks

While processing the GPS data files in Baton Rouge, we detected some areas where errors tend to occur frequently. To assist readers in this process, we developed a set of procedures or checks for data quality control that cover the entire operation from data collection to data reporting.

Data collection checks

- GPS equipment: Make sure the laptop computer clock date and time are correct because GPS receivers automatically assign data file names based on this information. Likewise, make sure the differential correction unit is working properly before beginning the travel time run.

- Route planning: Break the highway network into driving sections which can be driven in relatively short periods of time. For example, a 20 mi, highly congested corridor may be broken into two or three smaller driving sections, each of which can be driven in less than half an hour. To ensure complete coverage, the sections must overlap.
- Driving technique: Instruct the drivers to drive consistently, for example, using the floating car technique.
- File processing: After completing the run, instruct technicians to download data from the field laptop computer to a desktop computer in the office. This way, chances of misplacing or loosing GPS data files are considerably reduced. Depending on the circumstances and level of expertise of the technicians, it may also be advisable to train technicians to complete the data reduction procedure up to the MI/MO operation immediately after returning to the office from the field.
- GPS data file: Verify the GPS point file does not contain duplicate records. As an aid to users, file *points.exe*, which converts CMS *.ndc* files to comma-delimited *.txt* files, automatically filters out duplicate records. Likewise, verify the GPS point file does not contain outlying GPS data. Some receivers occasionally output records with coordinate data that would place the affected GPS points miles away from their actual location. Users have two options to address this situation: (a) remove the defective records before generating the GPS point file, or (b) generate the GPS point file and remove the defective records after the data reduction process (checking for GPS records with extremely large offset values helps users to locate outlying GPS data).

Data reduction checks

- File system: Make sure each GPS data file is stored in a separate subdirectory (having the same name as the GPS data file) under the GPSDATA subdirectory. Likewise, make sure the file entry in table GPS_FILE_PROPERTIES, particularly the file name and path, is correct.
- Point geographic files: Verify that all TransCAD-generated point geographic files are good. On an apparently random basis, TransCAD generates corrupt point geographic files (one out of 20-30 files, based on our experience with 428 GPS files in Baton Rouge). To check for corrupt files, open the point geographic file in TransCAD as soon as the file is generated and display its associated table. A corrupt file might display GPS points correctly on a map screen, but display absurd entries in the associated table. To solve the problem, run the procedure to generate the point geographic file again. Usually, the second time the resulting point geographic file is good.
- MI/MO operation: Verify that route assignments and beginning and ending time stamps are correct. An effective way of doing this is by checking the contents of file *gpsroute.dbf*. As an aid to users, this file is automatically displayed at the end of every MI/MO session.
- Route system files: Before beginning with the formal linear referencing process, verify that all route system files are good, i.e. that routes only contain valid links and that the beginning and ending mileposts of individual links are correct. To check this information, open the link files associated with each route (**L.dbf*) in the ROUTES subdirectory. This step is crucial. Depending on how the network file is set up, using the shortest path option to generate routes in TransCAD may cause routes to include "invalid" shortcuts. One option to address this situation would be to define turn penalties to prevent TransCAD from choosing "invalid" links.

A second option would be to add links to the route one by one. In any case, users are advised against editing routes after creating them because TransCAD does not update the beginning and end mileposts of any existing link after the editing process (manually editing or copying milepost values in file **L.dbf* does not solve the problem). See the evaluation of TransCAD and Access subsection for more information on this topic.

- Linear referencing: Verify that link codes, mileposts, and offsets are correct/meaningful. Checking for large offsets (larger than 60 ft) is perhaps the most effective way of detecting incorrect link code and milepost assignments. Because TTG always maps GPS data to links based on the routes defined in the MI/MO operation, wrong route assignments, which involve wrong link and milepost assignments, usually result in extremely large offsets (unless the route chosen is parallel and very close to the correct route). Conversely, correct route assignments, which normally involve correct link and milepost assignments, usually result in small offsets. To check for large offsets, open table *gpsdata.dbf* in TransCAD and sort its contents by Offset.

Because the links map remains constant during the linear referencing operation, offsets can be used to provide an indirect measurement of GPS data positional errors. To illustrate this point, figure 27 shows the offset distribution of differentially corrected GPS points and the offset distribution of uncorrected GPS points in Baton Rouge. Differentially corrected GPS points had much lower offsets than uncorrected GPS points, clearly indicating the benefit of using GPS equipment with differential correction. It may be worth noting that almost 70 percent of differentially corrected GPS points had an offset of less than 18 ft (5.5 m). This number is consistent with the positional accuracy of the GPS equipment we used in Baton Rouge (2-5 m spherical error probability (SEP) *[2]*.

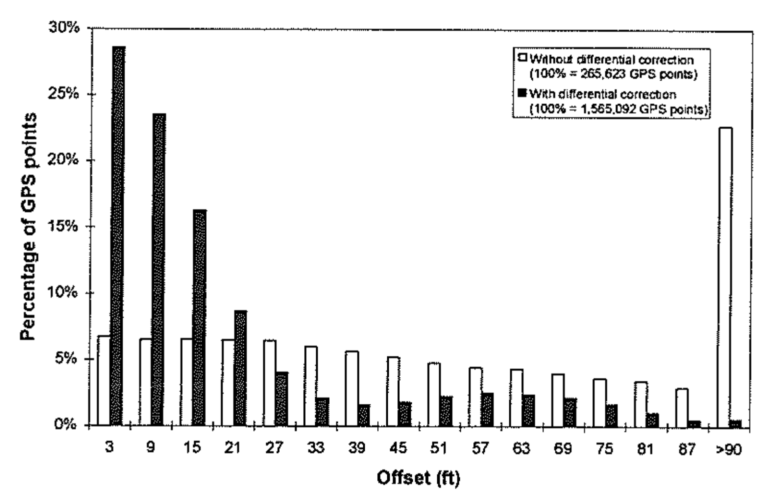


Figure 27

Offset distribution of differentially corrected and uncorrected GPS points in Baton Rouge

Data reporting checks

- Link/segment speeds: Verify that link/segment speeds are meaningful. As an aid to users, we developed several procedures in Access to compute link/segment speeds and a query to display speed differences between these procedures. It may be worth noting that by using this query we detected a bug in the TransCAD function that maps points to linear features along routes (see the evaluation of TransCAD and Access subsection for a more detailed description). This bug causes some isolated GPS points to be mapped to the wrong relative position along the route which, in turn, causes the corresponding assigned mileposts and link codes to be incorrect. As a result, the affected link/segment speeds are also incorrect. Because each implemented procedure to compute link/segment speeds uses different pieces of information, the resulting link/segment speeds are also different. To detect large speed differences between procedures more efficiently, it is advisable to sort speed differences in descending order.
- Average link/segment speeds: Verify that representative link/segment speeds are meaningful. We included two procedures to calculate representative link/segment speeds: using a harmonic mean formulation, and using a median formulation. As discussed previously, harmonic mean speeds are based on arithmetic mean travel times. However, harmonic mean speeds are very sensitive to low outlying speeds which tend to occur under atypically adverse traffic conditions. By comparison, median speeds are not sensitive to outliers and, therefore, they tend to provide more robust estimates of central tendency than harmonic mean speeds.

Data Reporting

To aid readers in identifying travel time and speed data reporting options, we developed several database querying and data reporting alternatives. In this section, we analyze these alternatives.

Database queries

Storing all linearly referenced GPS data in Access gave us a graphical, user-friendly environment in which to design the database schema, populate tables, design queries, and develop macros and forms. To test this environment, we stored all data associated with 1.8 million linearly referenced GPS data in an Access database file. As mentioned previously, the size of this file was about 238 Mbytes.

In general, Access handled this database file reasonably well, meaning that it did not crash, corrupt records, or generate strange results. However, we did notice a considerable decrease in performance as the number of records in table GPS_DATA increased. The most obvious effect was the additional time required to append new records to table GPS_DATA and to run queries involving table GPS_DATA. For example, appending data for the first GPS data files took less than one minute per file. In contrast, appending data for the last GPS data files took more than 30 minutes per file (even if the number of records per file was approximately the same). Likewise, running queries involving table GPS_DATA took much longer once all 1.8 million records were appended to table GPS_DATA than when this table only had a few records.

Appending data to table GPS_DATA is a one-time operation that runs automatically. Because user interaction is null to minimal, the extra time required to append records as the number of appended records increases is not really an issue. Furthermore, as computers become faster and databases more efficient, this issue should go away very soon. For building and running queries, however, a significant decrease in performance is an issue. To address this situation, users have several options:

1. To limit the size of the database based on date ranges or seasons. For example, users could define separate *ttgxxxx.mdb* files to store all linearly GPS data associated with the runs in year *xxxx*.
2. To run make table queries to store subsets of records from table GPS_DATA in different tables. For example, users could run a make table query to define a table called GPS_DATA_1997_430_530 to retrieve all records from GPS_DATA that were collected between 4:30 and 5:30 p.m. during 1997. This option has the obvious disadvantage of increasing the size of the database file and allows for the possibility of potential data integrity problems.
3. Similar to the previous option, except that the tables containing the subset of records are exported to a different database file. This alternative allows users to maintain the original database file as a master file based on which smaller, more manageable database files containing subsets of records are generated.

Data reporting options

In general, we found no difficulties in generating color coded maps and strip charts using TransCAD. This is, after all, one of TransCAD's strong characteristics. Other data reporting options like color coded maps in MGE and tabular archival reports in Access were essentially the same as those which we had already developed in the first phase of the CMS project, and, therefore, we will not extend the discussion here *[2]*. It may be of interest, however, to discuss our experience in the use of distance-time diagrams for computing delay at signalized intersections.

Computation of delay at signalized intersections

As mentioned previously, we developed an automated procedure for computing delay performance measures like control delay and stopped delay at signalized intersections using GPS data. Appendix B describes the underlying theory and algorithms and includes an example of application of the procedure using GPS data from a run made on Florida Boulevard eastbound on October 19, 1995 (file *10191129.txt* in figure 8). To extend the analysis, we used GPS data from additional runs made on Florida Boulevard and Airline Highway in Baton Rouge. Florida Boulevard is a four- and six-lane signalized arterial with posted speed limits between 40 mph and 50 mph. The traffic signal system is a fixed time system with a 150-second cycle length from 5:00 a.m. to 9:00 a.m. (the Florida Boulevard runs used in this report were made between 7:00 a.m. and 8:00 a.m.). Airline Highway is a four-lane signalized arterial with a 50 mph posted speed limit. The traffic signal system is a coordinated actuated system with a background 150-second cycle length.

We considered four groups of runs, one for each direction of travel: Florida Boulevard eastbound (EB), Florida Boulevard westbound (WB), Airline Highway northbound (NB), and Airline Highway southbound (SB). Figure 28 shows a sample distance-time diagram for each of the four groups of runs considered, with marks indicating the location of the delay control points (start of deceleration, end of deceleration, start of acceleration, and end of acceleration) at each signalized intersection the vehicle stopped. Table 11 shows a tabulation of measured control delay (d_c), approach delay (d_{ap}), and stopped delay (d_s) values, along with a code associated with the driver who performed the travel time run. As described in appendix B, control delay is the total delay experienced by a vehicle due to the influence of a signalized intersection and includes initial deceleration delay, stopped delay, and final acceleration delay. Approach delay is the amount of delay up to the signalized intersection stop bar. Stopped delay is the time a vehicle is stopped at a signalized intersection.

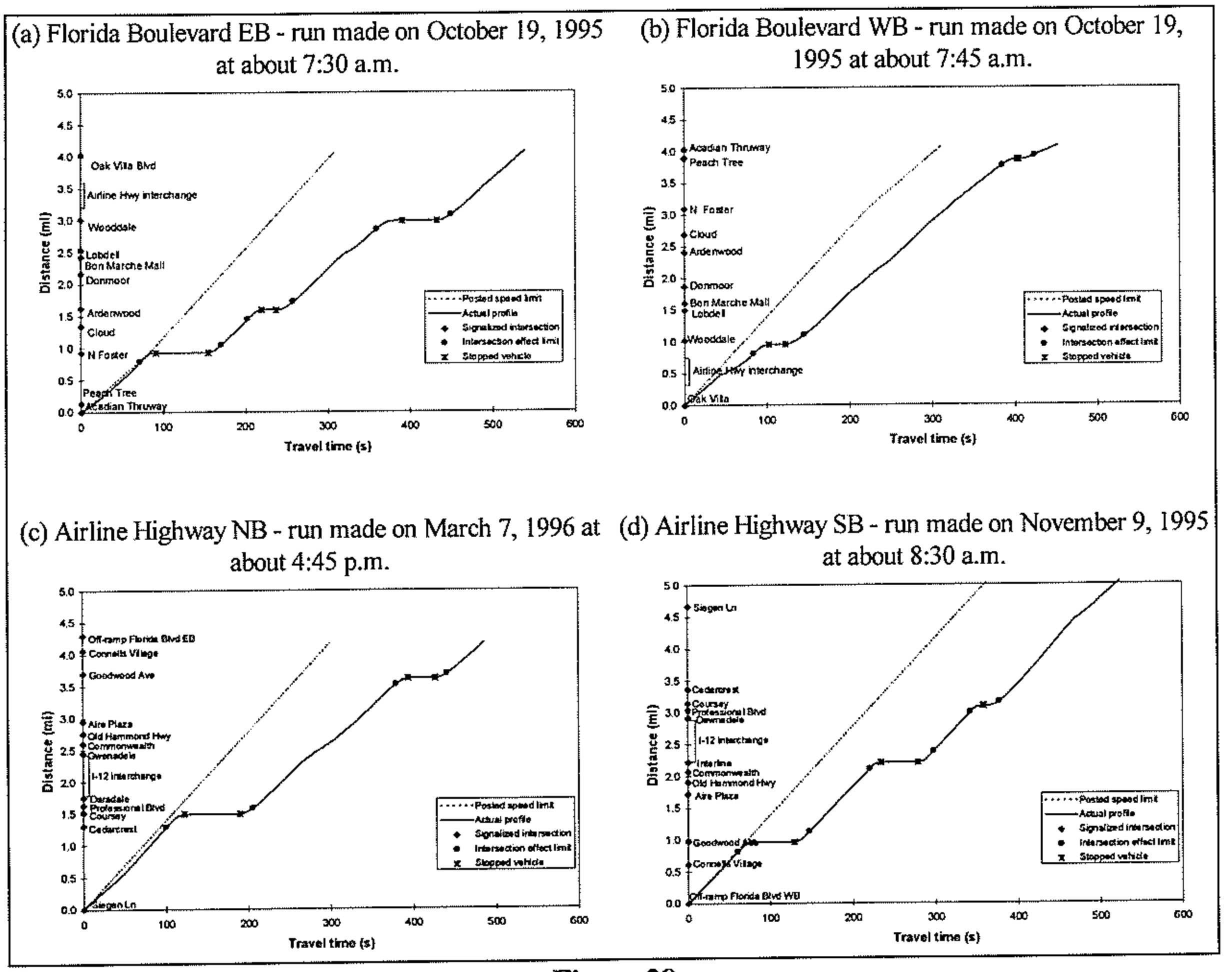


Figure 28

Sample distance-time diagrams in Florida Boulevard and Airline Highway in Baton Rouge

Table 11

Summary of control delay, approach delay, and stopped delay for the runs made on Florida Boulevard and Airline Highway in Baton Rouge

Florida Boulevard EB (24 observations based on 9 morning runs)				Florida Boulevard WB (9 observations based on 9 morning runs)				Airline Highway NB (27 observations based on 13 afternoon runs)				Airline Highway SB (49 observations based on 21 morning runs)							
d_c (s)	d_{ap} (s)	d_s (s)	Driver ID	d_c (s)	d_{ap} (s)	d_s (s)	Driver ID	d_c (s)	d_{ap} (s)	d_s (s)	Driver ID	d_c (s)	d_{ap} (s)	d_s (s)	Driver ID	d_c (s)	d_{ap} (s)	d_s (s)	Driver ID
28.9	26.2	6.0	ms	24.9	21.0	2.0	ms	23.3	19.4	10.0	bp	19.6	19.9	8.5	bm	63.8	60.3	51.5	bm
28.9	27.7	15.5	bm	26.8	26.1	1.5	ms	29.0	25.8	12.5	bp	22.9	19.1	1.0	bm	65.0	61.1	36.5	ms
29.9	23.6	12.0	bm	27.2	27.2	3.5	bm	29.6	29.6	2.5	bp	23.0	22.3	1.0	bm	65.2	62.4	44.0	bm
30.0	28.2	4.0	bm	40.2	36.8	20.0	ms	31.0	29.5	8.5	pt	24.7	19.9	4.0	bm	70.6	67.8	51.5	bm
30.1	25.4	7.0	ms	84.2	85.6	55.0	ms	33.5	29.5	14.0	bp	24.8	20.6	2.5	bm	71.2	71.0	57.4	ms
34.1	31.5	8.0	bm	84.8	81.6	69.0	ms	34.4	31.0	8.0	pt	24.9	23.8	3.0	bm	71.4	67.8	56.5	bm
34.3	25.5	11.5	ms	85.9	78.0	69.0	bm	43.3	41.6	5.5	ma	25.4	19.8	6.0	ms	71.4	68.5	50.5	bm
36.1	31.0	17.5	ms	86.0	85.2	65.5	ms	44.2	43.4	20.5	pt	33.2	30.8	12.5	ms	72.3	68.7	50.0	bm
39.4	36.2	18.0	bm	91.9	86.6	72.0	ms	44.6	39.8	7.5	pt	35.9	29.5	20.0	bm	74.5	70.6	64.0	bm
42.1	38.7	26.5	ms					46.1	46.9	25.5	ms	38.2	36.2	14.0	bm	76.6	68.1	57.5	bm
45.3	41.3	32.0	ms					46.2	45.2	23.5	pt	38.6	34.1	24.0	bm	77.5	74.7	51.5	bm
48.0	47.0	34.5	bm					47.9	50.3	24.0	ms	43.3	40.6	31.0	bm	78.1	70.7	36.5	ms
63.3	59.7	41.5	ms					49.6	49.9	32.5	bp	43.4	40.6	19.5	bm	80.0	74.8	47.0	bm
68.5	66.7	40.5	bm					53.7	49.9	38.0	bp	43.7	41.0	24.5	bm	80.0	82.9	58.0	bm
69.1	66.6	49.0	ms					54.0	50.4	31.5	bp	45.8	43.9	27.5	bm	80.2	74.2	59.0	bm
70.1	66.5	51.5	bm					54.8	52.1	36.9	pt	47.0	45.7	16.5	bm	83.3	78.9	63.5	bm
71.1	67.4	47.0	ms					56.3	53.2	40.0	bp	47.1	47.9	23.0	bm	83.4	80.3	53.5	bm
71.5	69.6	50.0	ms					59.1	58.4	36.0	pt	48.4	48.4	26.5	bm	89.2	84.9	71.0	bm
73.9	69.6	42.0	ms					63.3	60.4	40.0	bp	53.8	50.8	34.5	bm	89.4	84.9	66.5	bm
74.0	69.3	47.0	ms					64.0	60.4	40.5	pt	54.3	52.6	40.5	bm	90.4	84.0	63.0	bm
74.5	70.6	40.5	ms					69.3	71.7	56.0	bp	54.6	54.6	30.0	bm	91.1	87.7	55.0	bm
75.3	71.7	63.5	ms					71.2	69.7	34.5	pt	57.5	55.0	45.0	bm	91.6	93.4	52.5	bm
76.6	73.4	56.5	bm					72.2	67.9	50.5	bp	60.7	57.9	37.0	bm	93.4	86.8	74.0	bm
84.1	81.2	69.5	ms					74.0	72.2	46.0	pt	61.1	59.5	43.5	bm	99.4	90.0	73.0	bm
								84.2	79.6	67.5	bp	63.2	60.0	46.5	bm				
								90.3	85.5	76.5	ma								
								90.9	87.0	73.5	bp								

An analysis of the data in table 11 reveals that most d_s/d_c ratios are not close to the 0.76 conversion factor included in the Highway Capacity Manual (HCM) to convert control delay to stopped delay *[12]*. This finding empirically supports the work by others in the sense that such a factor is not constant *[13]*, *[14]*. Because we had more than 100 delay observations, we decided to generalize our results by constructing plots to determine the relationship between d_c, d_{ap}, and d_s for each of the four groups of records shown in table 11. The reader should note that by combining a large number of delay observations from several signalized intersections, we are measuring a "composite" delay which is not conceptually the same as the delay measured at an isolated signalized intersection using delay values from several vehicles. However, it is reasonable to assume that both measures of delay should yield similar results because (1) our GPS data represented 'typical' driving conditions (the drivers used the floating car technique to monitor traffic flow) and (2) the drivers did not have any restriction with respect to their position in a queue at any signalized intersection, which means that the GPS data and corresponding delay measurements could be considered as random samples.

The relationship between d_c, d_{ap}, and d_s for each of the four groups of records shown in table 11 turned out to be linear in all cases. Table 12 summarizes these trends. Although each driver had unique acceleration and deceleration characteristics, there were not significant variations in the relationship between d_c, d_{ap}, and d_s for these drivers. Because all trends were similar, we combined all records and determined the corresponding relationships between d_c, d_{ap}, and d_s (table 12 and figure 29). An analysis of these trends indicates the following:

Table 12

Relationship between control delay d_c, approach delay d_{ap}, and stopped delay d_s

Route	d_s vs. d_c		d_s vs. d_{ap}		d_{ap} vs. d_c		d_{ap} vs. d_c (zero intercept)	
	Equation	R^2	Equation	R^2	Equation	R^2	Equation	R^2
Florida Boulevard EB	$d_s = 0.952d_c - 18.6$	0.914	$d_s = 0.937d_{ap} - 14.5$	0.915	$d_{ap} = 1.013d_c - 4.21$	0.993	$d_{ap} = 0.944d_c$	0.988
Florida Boulevard WB	$d_s = 1.051d_c - 24.7$	0.985	$d_s = 1.058d_{ap} - 22.4$	0.960	$d_{ap} = 0.976d_c - 1.18$	0.991	$d_{ap} = 0.960d_c$	0.991
Airline Highway NB	$d_s = 1.050d_c - 24.8$	0.894	$d_s = 1.058d_{ap} - 22.4$	0.880	$d_{ap} = 0.978d_c - 1.05$	0.988	$d_{ap} = 0.961d_c$	0.988
Airline Highway SB	$d_s = 0.899d_c - 15.5$	0.911	$d_s = 0.916d_{ap} - 13.6$	0.898	$d_{ap} = 0.969\,d_c - 1.30$	0.989	$d_{ap} = 0.950d_c$	0.988
All combined	$d_s = 0.959d_c - 19.3$	0.917	$d_s = 0.967d_{ap} - 16.9$	0.904	$d_{ap} = 0.979d_c - 1.78$	0.989	$d_{ap} = 0.952d_c$	0.988

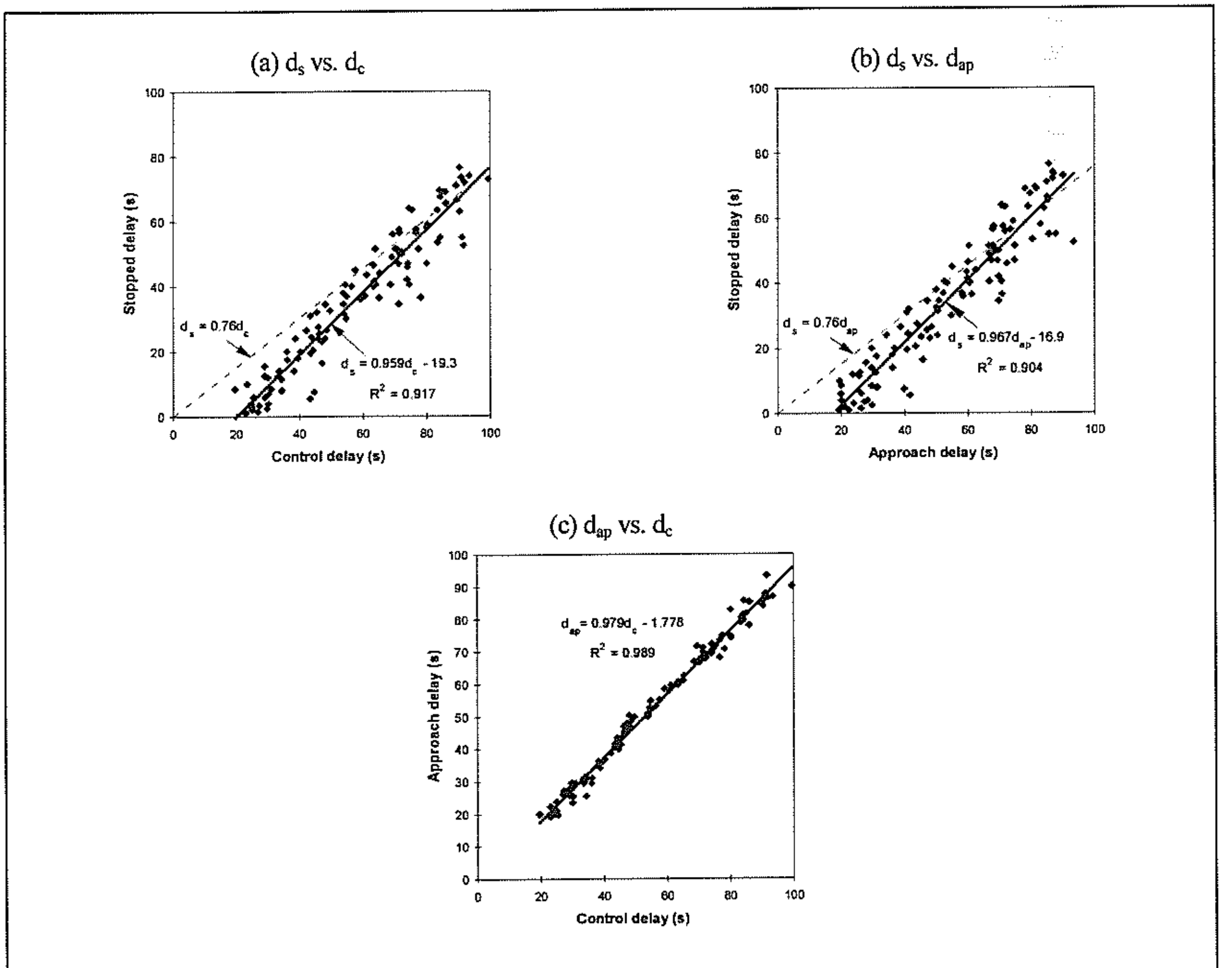


Figure 29

Comparison between control delay, approach delay, and stopped delay

1. The linear regression between d_s and d_c in figure 29a crosses the horizontal axis at a value of d_c of about 20 seconds. This value represents the minimum deceleration and acceleration delay that needs to take place before any stopped delay can occur. This is reasonable because for a

vehicle to stop (even for just a fraction of a second), the vehicle has to incur deceleration and acceleration delays. This is an indication that the relationship between d_c and d_s cannot be properly represented by a single factor, e.g. the 0.76 conversion factor included in the 1994 HCM *[12]*. It also raises questions with respect to the applicability of models such as those developed by Olszewski or Teply which do not include a deceleration-acceleration delay component that must be added to d_s in order to obtain d_c *[13]*, *[14]*. However, as Olszewski and Teply observed, we also believe cycle length, particularly the red interval length, should be included in the relationship between d_s and d_c.

2. The relationship between d_s and d_{ap} shown in figure 29b is very similar to that between d_s and d_c in figure 29a. The linear regression equation crosses the horizontal axis at a value of d_{ap} of about 18 seconds.
3. The relationship between d_{ap} and d_c shown in figure 29c confirms that d_{ap} is lower than d_c. Interestingly, the linear fit between d_{ap} and d_c is much higher (R^2 about 0.99) than that between d_s and d_c or that between d_s and d_{ap} (R^2 about 0.90). The intercept of the d_{ap} vs. d_c best fit linear equation is very low (less than two seconds). In fact, it could be eliminated without seriously affecting the quality of the regression. As shown in table 12, R^2 lowers from 0.989 to 0.988 when the linear trend is forced to pass through the origin. In this case, the slope of the d_{ap} vs. d_c linear equation becomes 0.95, suggesting that about five percent of the control delay at a signalized intersection takes place after the signalized intersection.
4. In figures 29a and 29b, the slopes of the best fit linear equations are 0.959 and 0.967, respectively. These slopes are very close to one and indicate that the deceleration and acceleration components of the control delay varied little throughout the range of d_c and d_{ap} values considered. This observation empirically supports the notion that relating stopped delays to control delay values through a numerical additive constant may be reasonable *[15]*. For this approach to be valid, however, the additive numerical constant must reflect typical deceleration and acceleration delays for the facility under consideration. For example, for the signalized arterials considered here, the data collected suggests a numerical additive constant of around 20 seconds. This value is significantly larger than the nine-second figure measured by Olszewski on three signalized intersections in Singapore *[13]*.
5. Figures 30a and 30b show the deceleration and acceleration frequency distributions associated with the control delay and stopped delay values shown in table 11. For each record, we computed an average deceleration as the ratio of the speed before the vehicle began to decelerate over the time needed to stop the vehicle at point 1 (see figure 36 in appendix B). Likewise, we computed an average acceleration as the ratio of the speed after the vehicle resumed uniform speed conditions over the time needed to achieve that uniform speed. As shown in figures 30a and 30b, deceleration and acceleration values ranged from 1.0 to 3.5 mph/s. The average deceleration was 2.1 mph/s, indicating comfortable decelerating conditions (2.1 mph/s is about 1/3 of the deceleration associated with coefficients of friction used for stopping sight distance analysis) *[16]*. The average acceleration was 2.0 mph/s, which is close to normal accelerating conditions on flat grade intersections *[16]*. Readers should note that the deceleration and acceleration values of figure 30 represent average values and that they only provide a general indication of driving behavior. Deceleration and acceleration rates are not really constant *[17]*. For example, figure 37 in appendix B shows that at the N. Foster intersection the deceleration range was -5.5 - 0 mph/s and the acceleration range was 0 - 5 mph/s.

6. Total deceleration lengths ranged from 0.02 mi to 0.26 mi, for an average of 0.14 mi (figure 31a). Likewise, total acceleration lengths ranged from 0.02 mi to 0.30 mi, for an average of 0.12 mi (figure 31b). To measure delay in the field, however, it is more important to know the distance from the point the vehicle begins to decelererate to the signalized intersection stop bar (or upstream section length) and the distance from the signalized intersection stop bar to the point the vehicle resumes uniform speed conditions (or downstream section length). Upstream section lengths ranged from 0.06 mi to 0.34 mi, for an average of 0.18 mi (figure 32a). Likewise, downstream section lengths ranged from slightly less than zero to 0.28 mi, for an average of 0.08 mi (figure 32b). In general, these values are much larger than others reported in the literature *[13], [18]*. One reason that explains our longer deceleration sections is that we did not constrain any deceleration section to be within a fixed distance upstream of the intersection. Another reason could be that long 150-second cycles usually result in long queues.

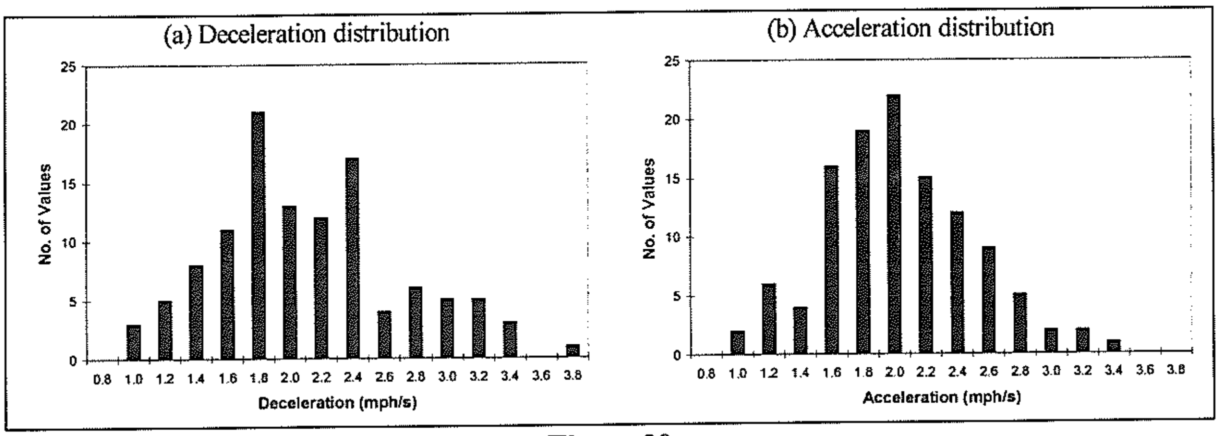


Figure 30
Frequency distribution of observed decelerations and accelerations

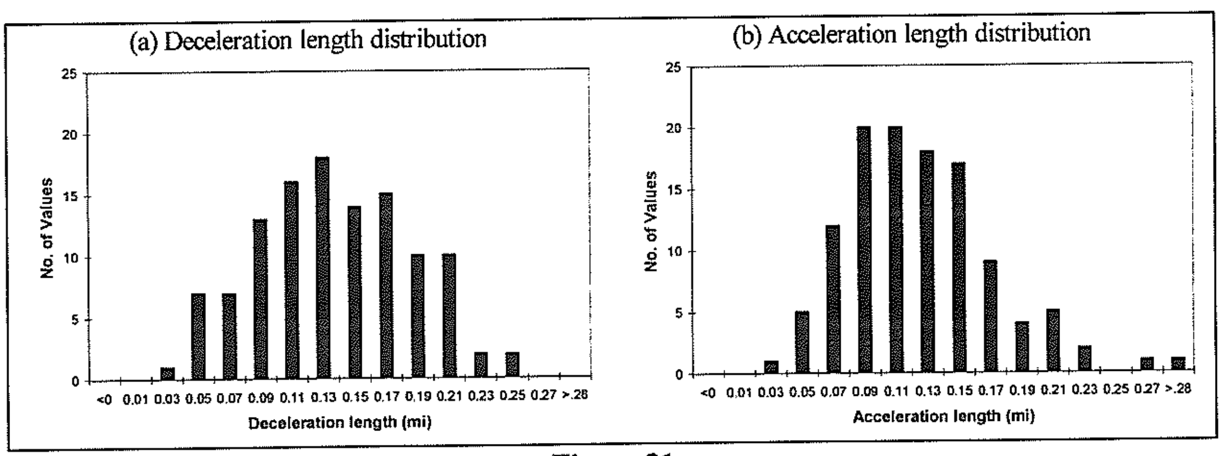


Figure 31
Frequency distribution of observed deceleration and acceleration lengths

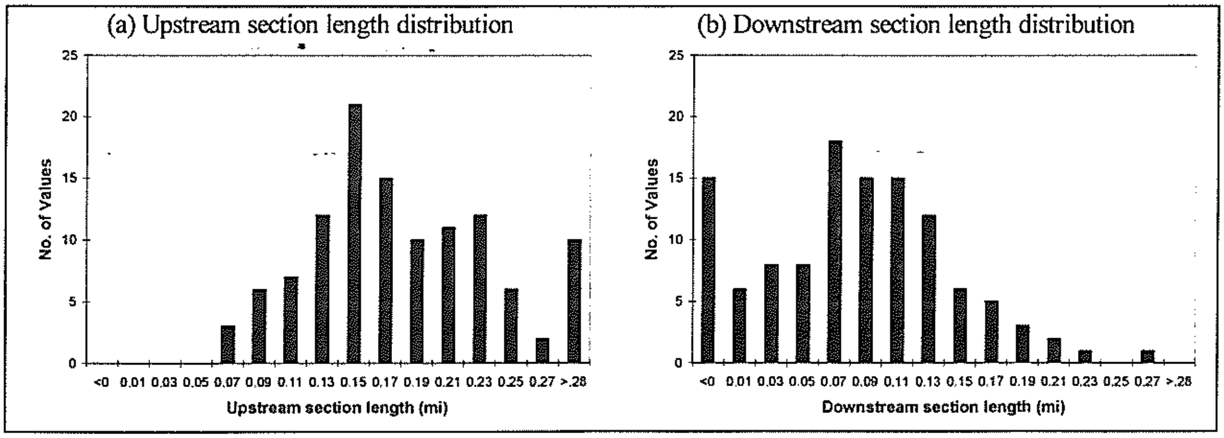


Figure 32
Frequency distribution of upstream and downstream section lengths

Evaluation of the TransCAD - Access Environment

Interspersed throughout this report are comments addressing specific issues about TransCAD and Access. In this subsection, we provide a more comprehensive analysis of the TransCAD - Access environment to give readers a broader perspective of its potential and performance.

In general, the TransCAD - Access environment provides users with a PC-based, relatively inexpensive environment to process vast amounts of GPS data and produce speed and travel time reports quickly and cost-effectively. This environment works reasonably well in most cases, particularly if one compares its price tag (around $3,500) with that of the MGE - Oracle environment (around $15,000). A much lower price tag involves trade-offs. In the case of Access, the trade-off with Oracle involves lower processing speeds. However, dramatic improvements in computer technology (faster processors and larger data storage capabilities at lower prices) should considerably reduce this effect in the near future.

In the case of TransCAD, the situation is more complicated. On one hand, TransCAD contains more transportation-tailored functions than any other GIS package in the market. This makes TransCAD an interesting option. On the other hand, probably because TransCAD is used mostly for transportation planning, some of those transportation-tailored functions do not seem to work properly for handling "fine" resolution data such as GPS data. Throughout this project, we have corresponded with Caliper officials to inform them about most of the problems/limitations we have found with TransCAD. We understand, however, that it takes time for developers to solve software problems and to include the resulting solutions in new software releases. In the mean time, readers are advised to pay extra attention to the accuracy of results obtained with the application.

A summary of problems we found in TransCAD follows. We believe this summary will be useful both for understanding software performance problems and for developing specifications that DOTD or the MPOs can use when procuring new GIS software.

Problems affecting results

- TransCAD computes surface distances based on a constant radius spherical model of the Earth's surface. This approach causes errors in the computation of horizontal distances which are larger than other GIS packages. Depending on the total length considered, those errors can be much larger than the positional accuracy of GPS data. The situation becomes apparent when mixing TransCAD-derived linearly referenced GPS data with checkpoint milepost data derived using other means, including other GIS packages. Table 8 illustrated this situation by comparing distances between two points computed with TransCAD with distances computed using Intergraph's MGE and two independent algorithms.
- TransCAD does not update beginning and ending mileposts associated with links in a route after editing the route alignment (for example, when adding links at the beginning of the route or realigning the route). For instance, as shown in table 13, if a link is added at the beginning of an already existing route, two links will end up having a beginning milepost of zero. Route files generated with TransCAD version 3.0 do not show the beginning and ending mileposts associated with route links, therefore hiding potential problems completely. The effect of this problem is in the computation of milepost values for linearly referenced GPS data. To avoid the problem, users must make sure that the beginning milepost associated with a link is the same as the ending milepost associated with the immediately preceding link. If the numbers do not match, users should regenerate the route file from scratch. As mentioned previously, simply copying and pasting milepost values in table **L.dbf* does not solve the problem.
- TransCAD functions that select linear features using neighborhood operations are not reliable. For example, clicking on a linear feature does not necessarily mean that that linear feature will be selected. Instead, TransCAD may select a nearby linear feature (which is evidenced by its change in color from blue to red and by the attribute information displayed in the feature information box), or select no linear feature at all. Similar situations occur when trying to select linear features using a specific radius. Sometimes, TransCAD does not select any linear feature even if the radius is large enough to cover several linear features completely. This problem becomes apparent in TTG when users click on the GPS Player Accept button to validate routes and are not able to do it because no links are selected regardless of how large the search radius is. In these cases, users must be more careful than usual in accepting their initial route selection.
- Occasionally, TransCAD generates corrupt point geographic files. A corrupt point geographic file might display GPS points correctly on a map screen but display absurd entries in the associated table. To solve the problem, users must run the procedure to generate the point geographic file again. Usually, the second time the resulting point geographic file is good.
- Occasionally, TransCAD maps GPS points to a wrong relative position along the route under consideration (even if the GPS points are displayed correctly on a map window in TransCAD, and even if the route system file is good). This results in incorrect milepost, link code, and offset assignments for the affected GPS points. As an illustration, consider the GPS data shown in figure 33. We collected this data on November 16, 1995, on Scenic Highway NB in Baton Rouge (file code = 223, file name = *11161237.txt*, as shown in appendix C). Notice that three GPS points that were supposed to be assigned to link 2996 were instead assigned to link 2997. Their corresponding mileposts and offsets were also wrong. Detecting cases such as

these by hand can be difficult due to the large number of records normally associated with typical GPS files. However, we found that by computing link speeds using some of the procedures discussed previously (figure 14) and by comparing differences among them we can determine locations where link code and milepost assignment anomalies have occurred.

Table 13
Incorrect mileposts vs. correct mileposts for links on Plank Road NB in Baton Rouge

Link code	Incorrect mileposts after adding links 1544-1445 at the beginning of the route		Correct mileposts by defining route from link 1544 to link 1471 directly	
	From	To	From	To
1544	0.0000	0.0028	0.0000	0.0028
1440	0.0000	0.2633	0.0028	0.2661
1441	0.2633	0.4178	0.2661	0.4206
1442	0.4178	0.4223	0.4206	0.4251
1443	0.0000	0.1301	0.4251	0.5552
1444	0.1301	0.2031	0.5552	0.6282
1445	0.2031	0.5142	0.6282	0.9393
1446	0.0000	0.0347	0.9393	0.9739
1447	0.0000	0.6218	0.9739	1.5957
1448	0.6218	0.8549	1.5957	1.8287
1449	0.8549	1.0881	1.8287	2.0619
1450	1.0881	1.4540	2.0619	2.4279
1451	1.4540	1.9287	2.4279	2.9026
1452	1.9287	2.4960	2.9026	3.4699
1453	2.4960	2.7981	3.4699	3.7720
1454	2.7981	2.8311	3.7720	3.8050
1455	2.8311	2.9865	3.8050	3.9604
1456	2.9865	3.2754	3.9604	4.2493
1457	3.2754	3.3499	4.2493	4.3237
1458	3.3499	3.5173	4.3237	4.4912
1459	3.5173	3.5650	4.4912	4.5389
1460	3.5650	3.6631	4.5389	4.6370
1461	3.6631	3.7314	4.6370	4.7052
1462	3.7314	3.8519	4.7052	4.8257
1463	3.8519	3.9567	4.8257	4.9306
1464	3.9567	4.1390	4.9306	5.1128
1465	4.1390	4.5952	5.1128	5.5690
1466	4.5952	6.7867	5.5690	7.7606
1467	6.7867	7.2534	7.7606	8.2272
1468	7.2534	8.4602	8.2272	9.4340
1469	8.4602	8.7836	9.4340	9.7575
1470	8.7836	9.4122	9.7575	10.3860
1471	9.4122	13.7038	10.3860	14.6776

Problems affecting performance but not results

- The ODBC connection allows users to update tables in remote databases but not create tables or append records to tables in remote databases. Because of this, we were forced to output all linearly referenced GPS data into temporary *gpsdata.dbf* files (one for each GPS file) and then use Access to append these *gpsdata.dbf* files to table GPS_DATA.

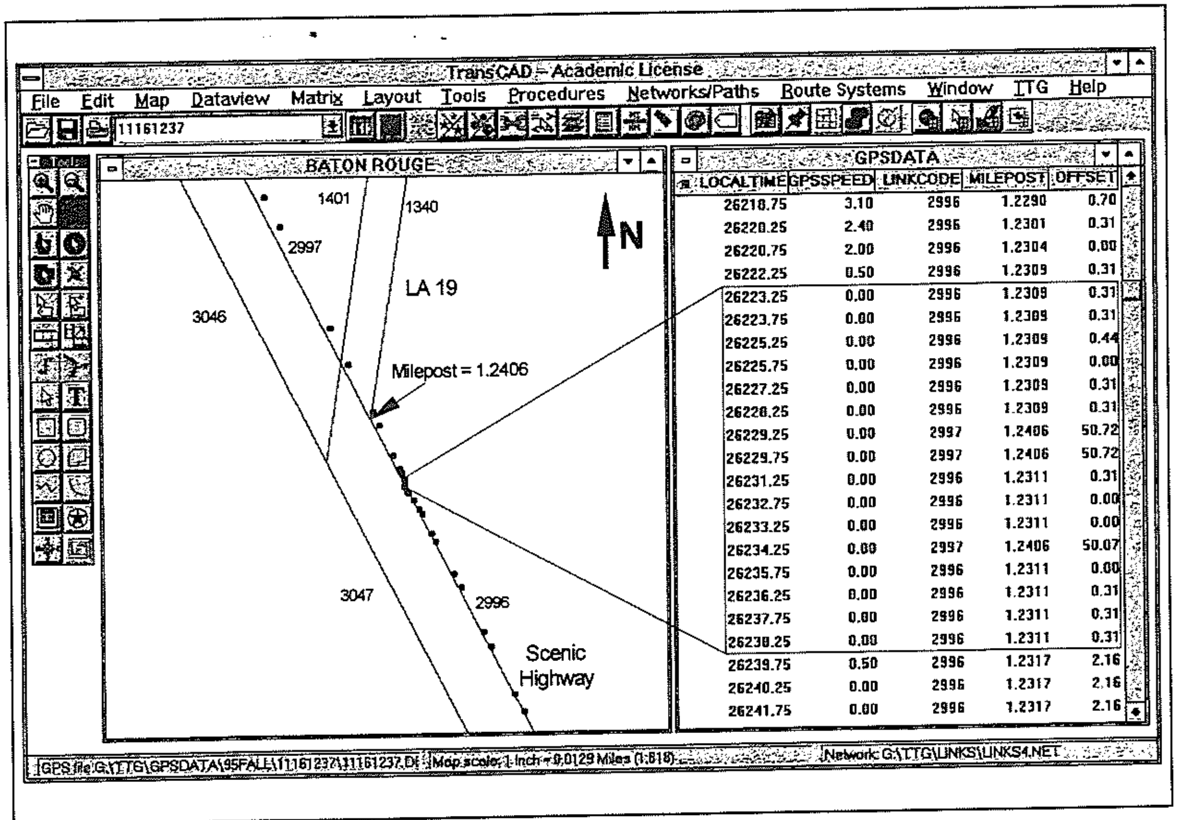


Figure 33
GPS data on Scenic Highway in Baton Rouge on November 16, 1995

- When importing Intergraph's MGE *.dgn* files, TransCAD breaks linear features that had been joined in Microstation previously. As a result, we were forced to rejoin those linear features in the resulting line geographic file in TransCAD.
- TransCAD is a 16-bit application. As a result, it only supports 16-bit ODBC connections. This means, for example, that Access version 2.0 is supported, but Access version 7.0 is not. This may be a problem to districts or MPOs that have already migrated completely to a 32-bit computing environment. However, it may be worth noting that TransCAD plans to release a 32-bit version of TransCAD before the end of 1998.
- TransCAD hard codes the path of the line geographic file based on which route files are generated. This means that routes cannot be opened in TransCAD (or added as layers) if the underlying line geographic file is not located in its original directory. If users wanted to use already existing line geographic files and route files in TTG, they would have to copy the line geographic file from its original location to the LINKS subdirectory, copy the route files from their original locations to the ROUTES subdirectory, and then use the TransCAD function ModifyRouteSystem to change the underlying path associated with the line geographic file. As an aid to users, TTG automatically executes this function every time a TTG project file is opened so that route files correctly point to the *links.dbd* file located in the LINKS subdirectory.
- The name of a GPS point file does not always appear in the list of layers after opening a GPS point file using TTG. To make the file name visible in the layer list, it is necessary to select the layer tool and then close it without making any changes. Not having the GPS point file visible

in the list of layers may be troublesome because in some platforms the GPS points are not initially visible in green after opening the GPS point file using TTG.

- TransCAD does not display milepost tick marks correctly. Depending on the scale, the location and values associated with milepost tick marks may change. This problem is particularly evident at small scales, i.e. when we zoom in on a particular area to view objects with greater detail.
- TransCAD OLE functionality does not work properly. As a result, TransCAD objects which are pasted into other Windows applications such as Word may appear distorted or display non-existing line work.

CONCLUSIONS

This report describes a set of CMS monitoring procedures which can be implemented on low cost PCs that are commonly available. The main contribution is a GPS-GIS methodology which can be used by district engineers, city traffic officials, and MPO planners for conducting travel time studies and monitoring congestion at various levels of resolution (from system to local). The set of procedures, collectively called Travel Time with GPS (TTG), include the following deliverables:

- A procedure for producing GPS-based network maps suitable for travel time studies
- A procedure for estimating sample size requirements during the data collection process
- A procedure and accompanying software based on dynamic segmentation for generating a linear referencing system to all GPS travel time and speed data along arterial streets and highway sections
- A procedure for reporting travel time and speed data both in graphical and tabular formats.

TTG is built around a general data model that includes a spatial model, a geographic relational database, and a procedure for linearly referencing GPS data. The spatial model uses a GPS-based bi-directional vector representation of the highway network. In this vector representation of the network, routes are partitioned into links and links are assigned unique identification numbers. The geographic relational database is composed of a series of tables that store information about links, routes, posted speed limits, GPS file descriptors, and linearly referenced GPS data. The procedure for linearly reference GPS data uses GIS dynamic segmentation tools. To automate this process, we developed an application that allows users to determine when a vehicle enters and exits a route and to automatically calculate mileposts for all GPS points along the routes of interest.

TTG addresses one of the most important technical issues faced by agencies using GPS to perform travel time studies: that of properly indexing, storing, and retrieving the data in an environment where millions of GPS records may be collected. By using a geographic relational database approach and dynamic segmentation tools, we provide a systematic procedure that enables agencies 1) to process GPS travel time data in a coherent, controlled way; 2) to store the data in a relational database configuration, which is critical for data archival purposes and the development of client-server applications; and 3) to efficiently retrieve and use the data for analysis and reports.

The report includes an updated methodology for estimating required sample sizes. This methodology uses t-distribution parameters and a more reliable procedure to estimate average sample ranges than that suggested in current ITE guidelines. Realizing that budgetary capabilities of agencies statewide vary widely, the new methodology includes tables summarizing sample size requirements for a wide range of confidence levels (75 percent to 99.73 percent) instead of just one (95 percent) that is included in current ITE guidelines. Such a wide range of confidence levels should help traffic engineers to make decisions regarding sample sizes which are consistent with their needs and constraints.

The report also includes a new methodology to measure delay at signalized intersections based on linearly referenced GPS data. This methodology is compatible with new Highway Capacity Manual procedures that focus on the measurement of control delay (which includes deceleration delay, stopped delay, and acceleration delay) as opposed to just stopped delay. Because it uses automated position and speed data collection devices (GPS receivers), the methodology can be used in situations where automatic and/or real-time delay measurements are needed. The methodology evaluates the main components of delay by analyzing distance-time, speed-time, and acceleration-time diagrams from a travel time run. A forward average acceleration algorithm detects when the vehicle starts to accelerate or decelerate. A backward average acceleration algorithm detects when the vehicle stops accelerating or decelerating. We tested these algorithms using linearly referenced GPS data from runs made on Florida Boulevard and Airline Highway in Baton Rouge. Based on more than 100 delay records on these two corridors, we observed the following (readers should note, however, that further studies on isolated intersections, corridors with lower cycle lengths, and corridors with lower posted speed limits are needed to generalize the results obtained):

- The relationship between stopped delay d_s and control delay d_c (and also that between d_s and approach delay d_{ap}) tends to be linear, but it does not pass through the origin. A relatively constant deceleration-deceleration delay value must be added to d_s in order to obtain d_c or d_{ap}. For modeling, in addition to the initial deceleration-acceleration value, other parameters such as red interval length should be included in the relationship between d_s and d_c.
- A significant percentage of the control delay takes place after the vehicle crosses the stop bar at the signalized intersection. For the signalized intersections considered in this report, that percentage was about five percent. Likewise, deceleration and acceleration lengths are much longer than others found in the literature. For the signalized intersections analyzed in this report, the average distance from the point where the vehicles started to decelerate to the intersection stop bars was 0.18 mi. Likewise, the average distance from the stop bars to the point where the vehicles had acquired a uniform speed was about 0.08 mi.

We implemented TTG using a TransCAD - Access environment. This environment provides users with a PC-based, relatively inexpensive environment to process vast amounts of GPS data and produce speed and travel time reports quickly and cost-effectively. This environment works well in most cases. However, there are problems, particularly with the current version of TransCAD (3.1), that may have an impact on results and/or performance of the application we developed. Probably because TransCAD is used mostly for transportation planning, some key TransCAD functions do not seem to work properly for handling "fine" resolution data such as GPS data. The details of these problems are documented in the Evaluation of the TransCAD - Access Environment section. Throughout this project, we have corresponded with Caliper officials to inform them about most of the problems/limitations we have found with TransCAD. We understand, however, that it takes time for developers to solve software problems and to include the resulting solutions in new software releases. In the mean time, readers are advised to pay extra attention to the accuracy of results obtained with the application.

LIST OF ACRONYMS AND ABBREVIATIONS

ASCII	American Standard Code for Information Interchange
CMS	Congestion management system
CRPC	East Baton Rouge Capital Region Planning Commission
DCI	Differential Corrections Inc.
DDE	Windows Dynamic Data Exchange
DGPS	Differentially corrected GPS (data)
DLL	Dynamic Link Libraries
EB	Eastbound
FHWA	Federal Highway Administration
ft	foot (1 ft = 0.3048 m)
GIS	Geographic information system
GISDK	GIS Developing Kit (TransCAD's developing language)
GPS	Global positioning system
HCM	Highway Capacity Manual
ISTEA	Intermodal Surface Transportation Efficiency Act
ITE	Institute of Transportation Engineers
Kbytes	Kilobytes
DOTD	Louisiana Department of Transportation and Development
MGE	Intergraph's Modular GIS Environment
m	Meter (1 m = 3.281 ft)
Mbytes	Megabytes
MHz	MegaHertz (1 MHz = 10^6 Hertz)
mi	Mile (1 mi = 1.609 km)
min	Minute
mph	Miles per hour (1 mph = 1.609 km/h)
MPO	Metropolitan planning organization
NB	Northbound
ODBC	Open database connectivity
OLE	Object linking and embedding
PC	Personal computer
QBE	Query by example
RDS	DCI's Remote Data System
SA	Selective availability
SB	Southbound
SEP	Spherical error probability
SQL	Structured query language
TIGER	Topologically integrating geographic encoding and referencing
UTC	Universal coordinated time
WB	Westbound

REFERENCES

1. *Intermodal Surface Transportation Efficiency Act*. U.S. code, Title 23, Chapter 3, Section 303, 1991.
2. Bullock, D., and Quiroga, C.A., *Development of a Congestion Management System Using GPS Technology*, Louisiana Transportation Research Center, Baton Rouge, LA, April 1997.
3. Leick, A., *GPS Satellite Surveying*. John Wiley & Sons, New York, 1995, 560 p.
4. Sandlin, A., McDonald, K., and Donahue, A., "Selective Availability: To Be or Not to Be," *GPS World*, Vol. 6, No. 9, 1995, pp. 44-51.
5. Robertson, H.D., "Travel Time and Delay Studies," *Manual of Transportation Engineering Studies*, Robertson, H.D. (Ed), Institute of Transportation Engineers, 4th Edition, Washington DC, 1994, pp. 52-68.
6. Quiroga, C.A., and Bullock, D., "Determination of Sample Sizes for Travel Time Studies," to be published in *ITE Journal* in early 1998.
7. May, A.D., *Traffic Flow Fundamentals*, Prentice-Hall, Englewood, New Jersey, 1990, 464 p.
8. Turner, S.M., Eisele, W.L., Benz, R.J., and Holdener, D.J., *Travel Time Data Collection Handbook*, Federal Highway Administration, Texas Transportation Institute, 1998.
9. Pittman, M.E., "Precision Direct and Inverse Solutions of the Geodesic," *Surveying and Mapping*, Vol. 46, No. 1, 1986, pp. 47-54.
10. Maling, D.H., *Coordinate Systems and Map Projections*, 2nd Edition, Pergamon Press, Oxford, England, 1992, 476 p.
11. Vincenty, T., "Application of the Chord Method to Solutions of Geodetic Lines," *Surveying and Mapping*, Vol. 46, No. 4, 1986, pp. 287-292.
12. TRB, *Highway Capacity Manual*, Special Report 209, 3rd Edition Update, National Research Council, Washington, DC, 1994.
13. Olszewski, P., "Overall Delay, Stopped Delay, and Stops at Signalized Intersections," *Journal of Transportation Engineering*, American Society of Civil Engineers, Vol. 119, No. 6, 1993, pp. 835-852.
14. Teply, S., *Accuracy of Delay Surveys at Signalized Intersections*, Transportation Research Record 1225, TRB, National Research Council, Washington, DC, 1989, pp. 24-32.
15. Rouphail, N.M., Anwar, M., Fambro, D.B., Sloup, P., and Perez, C.E., *Validation of a Generalized Delay Model for Vehicle-Actuated Traffic Signals*, 76th Annual Meeting Transportation Research Board, Preprint 970837, Washington, DC, 1997.
16. AASHTO, *A Policy on Geometric Design of Highways and Streets*, Washington, DC, 1990.
17. Bonneson, J.A., *Modeling Queued Driver Behavior at Signalized Junctions*, Transportation Research Record 1365, Transportation Research Board, Washington D.C., 1992, pp. 99-107.
18. Reilly, W.R., Gardner, C.C., and Kell, J.H., *A Technique for Measurement of Delay at Intersections. Vol. I Technical Report*, Report No. FHWA-RD-76-135, Washington, DC, 1976.
19. Gans, P., *Data Fitting in the Chemical Sciences*, John Wiley & Sons, Chichester, England, 1992, 258 p.
20. Quiroga, C.A., *An Integrated GPS-GIS Methodology for Performing Travel Time Studies*, Louisiana State University Ph.D. Dissertation, Baton Rouge, Louisiana, 1997, 171 p.
21. Trimble Navigation Ltd, *Placer GPS 400 Installation and Operator's Manual*, Sunnyvale, California, 1993.

22. Duncan, A.J., *Quality Control and Industrial Statistics*, Irwin, Homewood, Illinois, 5th Edition, 1986, 1123 p.
23. Hummer, J.E., "Intersection and Driveway Studies," *Manual of Transportation Engineering Studies*, Robertson, H.D. (Ed), Institute of Transportation Engineers, 4th Edition, Washington DC, 1994, pp. 69-87.
24. Fambro, D.B., and Rouphail, N.M., *Generalized Delay Model for Signalized Intersections and Arterial Streets*, 76th Annual Meeting Transportation Research Board, Preprint 970823, Washington, DC, 1997.
25. Millman, J., and Halkias, C.C., *Integrated Electronics: Analog and Digital Circuits and Systems*, McGraw-Hill, New York, NY, 1972.

APPENDIX A: MATHEMATICAL MODEL FOR COMPUTING SEGMENT SPEED AND TRAVEL TIME

Figure 34 shows the time-distance diagram of a probe vehicle as it traverses a segment [or any other linear feature like a link, section, or route (figure 5)] of length L. The dots represent GPS points collected with the GPS equipment carried on board. Because each GPS point has a milepost associated with it and we know the segment entrance milepost and the segment exit milepost, we can very easily determine (1) which points can be associated with the segment and (2) the relative position of each GPS point along the segment. This configuration gives us total flexibility regarding the definition of linear features for which we would like to compute speed and travel time. The only requirement is that we have adequate GPS coverage on the linear feature. Obviously, the definition of adequate GPS coverage depends on the procedure used to compute speed and travel time on the linear feature.

In order to compute link travel time and speed, we could interpolate the time stamps of the two GPS points located immediately before and after the segment entrance and the time stamps of the two GPS points located immediately before and after the segment exit (for example, points P_p and P_{p+1} in figure 34). For this, we would also need the mileposts of the GPS points involved and the segment entrance milepost and the segment exit milepost. Because each GPS point has an associated positional error ε, the interpolated time also has an associated error ε_t. This error translates into an error ε_u in the computation of the segment speed u.

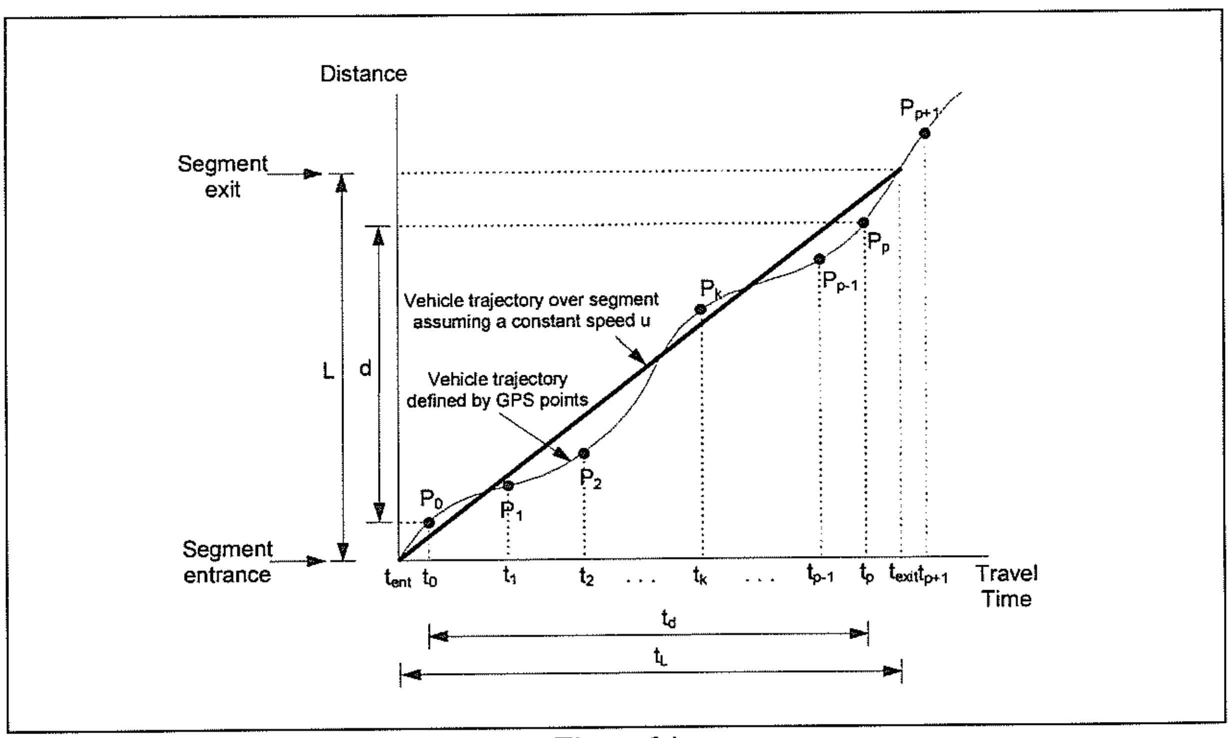


Figure 34
Time-distance diagram for GPS points on a segment

Let the interpolated times associated with the segment entrance and exit be t_{ent} and t_{exit}. The speed u associated with the segment of length L is (figure 34)

$$u \approx \frac{L}{t_{exit} - t_{ent}} = \frac{L}{t_L} \tag{4}$$

where t_L is segment travel time. Using error propagation theory we can show that ε_u can be expressed as *[19], [20]*

$$\varepsilon_u \approx \frac{\sqrt{2}\varepsilon}{t_{exit} - t_{ent}} = \frac{\sqrt{2}\varepsilon}{t_L} \tag{5}$$

Equation (5) shows that ε_u is a function of the positional accuracy of individual GPS points and the segment travel time. It provides an upper bound for ε_u because the errors associated with t_{ent} and t_{exit} are assumed to be independent. In practice, ε_u values may be lower if the errors associated with t_{ent} and t_{exit} are both of the same sign and order of magnitude. However, because typical L values are relatively large, say 0.2 mi or larger, we assume such errors to be independent for all practical purposes. It must be noted that equation (5) is based on the assumption that a linear interpolation scheme is appropriate for computing t_{ent} and t_{exit}. This assumption is reasonable in an environment in which GPS data is collected every one second and distances between contiguous GPS points are much smaller than segment lengths.

In our research we used GPS receivers which had the capability to record speeds in addition to coordinates and time. For speed computations, the receivers poll pseudorange data (distance from satellite to receiver) and pseudorange rate data for a fraction of a second. As a result, the computed speeds are almost instantaneous and essentially independent of position fixes. This allowed us to use GPS speeds to calculate distance driven which, in turn, allowed us to calculate segment speeds. In figure 34, the total distance d covered by the probe vehicle between t_o and t_p is

$$d = \int_{t_0}^{t_p} v\,dt \approx v_0\left(\frac{t_1 - t_0}{2}\right) + \left[\sum_{k=1}^{p-1} v_k\left(\frac{t_{k+1} - t_{k-1}}{2}\right)\right] + v_p\left(\frac{t_p - t_{p-1}}{2}\right)$$

or

$$d = \int_{t_0}^{t_p} v\,dt \approx \frac{1}{2}\sum_{k=0}^{p-1}(v_k + v_{k+1})(t_{k+1} - t_k) \tag{6}$$

assuming a special case of the trapezoidal rule in which time intervals do not remain constant (to account for variations in the GPS sampling rate). As before, the trapezoidal approximation is reasonable in an environment in which GPS data is collected every one second and distances between contiguous GPS points are much smaller than segment lengths.

The speed u_d associated with d is

$$u_d = \frac{d}{t_d} = \frac{d}{t_p - t_0} \tag{7}$$

Normally, the segment length L is known. If the initial and final GPS points associated with the segment are close to the entrance and exit points, respectively, d and L should be very similar. In this case, u_d can be assumed to apply over the entire segment length and can be used to estimate u. Equation (7) then becomes

$$u \approx \frac{1}{2(t_p - t_0)} \sum_{k=0}^{p-1} (v_k + v_{k+1})(t_{k+1} - t_{k-1}) \tag{8}$$

The resulting travel time t_L along the segment is then

$$t_L \approx \frac{L}{u} \tag{9}$$

It is possible to derive an expression for ε_u when using equation (8) to estimate u. As before, using error propagation theory, and assuming, for simplicity, that the sampling period or time interval Δt between contiguous GPS points is constant, we can show that

$$\varepsilon_u \approx \frac{\sqrt{p-0.5}}{p} \varepsilon_v \approx \frac{\sqrt{\Delta t(t_L - 0.5\Delta t)}}{t_L} \varepsilon_v \tag{10}$$

where ε_v is the error in speed associated with individual GPS points and p is the number of GPS points that can be associated with a segment (in addition to P_0) *[19]*.

Equations (5) and (10) are comparable in the sense that they only measure errors in segment speed due to errors associated with individual GPS data measurements (ε and ε_v, respectively). They do not measure errors due to other factors like GIS distance measuring accuracy or GPS sampling rates. Because equations (5) and (10) are comparable, we can combine them to determine a threshold beyond which equation (8) would be preferable for computing segment speeds. Such a threshold is given by

$$\varepsilon_v \le \frac{\sqrt{2}\varepsilon}{\sqrt{\Delta t(t_L - 0.5\Delta t)}} \tag{11}$$

Table 14 shows a few sample values of ε_v for several combinations of ε and t_L values, assuming $\Delta t = 1$ second. Note that ε_v is directly proportional to ε. This means that as ε decreases, i.e. as the positional accuracy of the GPS equipment increases, ε_v must also decrease in order for equation (8) to maintain its relative advantage. Otherwise, equation (4) will become relatively more accurate than equation (8). Note also that ε_v is inversely proportional to t_L. This means that as the segment travel time increases (either because speeds are lower or because segments are longer), the effect of positional errors for individual GPS points will be lower, and, as a result, equation (4) will become relatively more accurate than equation (8). For example, assume a GPS equipment with a positional accuracy of 1 m and a speed accuracy of 0.5 mph. Assume the probe vehicle is traveling at 50 mph and collecting GPS data every 1 second. For a segment length of 0.2 mi, the corresponding travel time t_L is 14.4 seconds. From table 14, the limiting error ε_v is approximately 0.85 mph, which is larger than the speed accuracy of the GPS equipment. Under these conditions, equation (8) would be more accurate than equation (4). By comparison, for a segment length of 1.0 mi, the corresponding travel time t_L is 72 seconds. From table 14, the limiting error ε_v is approximately 0.37 mph, which is smaller than the speed accuracy of the GPS equipment. Under these conditions, equation (4) would be more accurate than equation (8). In Baton Rouge, we used GPS equipment (including differential correction) with a positional accuracy of 2-5 m spherical error probability (SEP) and a speed accuracy of 0.1 mph (1 sigma) *[21]*. From Table 14, it is evident that the limiting error ε_v is much larger than 0.1 mph. This result suggests that equation (8) would be more accurate than equation (4) for Baton Rouge.

Table 14
Limiting error ε_v as a function of t_L and ε, assuming $\Delta t = 1$ second [equation (11)]

t_L (sec)	Limiting error ε_v in GPS speed measurements (mph)							
	$\varepsilon=0.1$ m	$\varepsilon=0.2$ m	$\varepsilon=0.5$ m	$\varepsilon=1$ m	$\varepsilon=2$ m	$\varepsilon=5$ m	$\varepsilon=10$ m	$\varepsilon=20$ m
5	0.15	0.30	0.75	1.49	2.98	7.46	14.92	29.83
10	0.10	0.21	0.51	1.03	2.05	5.13	10.27	20.53
20	0.07	0.14	0.36	0.72	1.43	3.58	7.17	14.33
30	0.06	0.12	0.29	0.58	1.17	2.91	5.83	11.65
60	0.04	0.08	0.21	0.41	0.82	2.05	4.10	8.20
120	0.03	0.06	0.14	0.29	0.58	1.45	2.89	5.79
240	0.02	0.04	0.10	0.20	0.41	1.02	2.04	4.09
480	0.01	0.03	0.07	0.14	0.29	0.72	1.44	2.89
960	0.01	0.02	0.05	0.10	0.20	0.51	1.02	2.04
1920	0.01	0.01	0.04	0.07	0.14	0.36	0.72	1.44

Equation (4) or equation (8) is valid for a single run on a single segment and provides the necessary tools to transform a set of GPS point time stamp and speed values into a single pair of travel time and speed values. In general, however, several runs involving several contiguous segments are typically made. In this case, it may be of interest to compute not only speed and travel time for individual segments due to individual runs, but also representative speed and travel time values for each segment and for all segments combined. To keep the discussion general, the number of runs per segment is assumed to be different (figure 35). This is particularly important if there are interchanges or intersections along the route and some segments have more records than other segments. Because of this, the procedure to compute representative speed and travel time values involves aggregating the data at the segment level first.

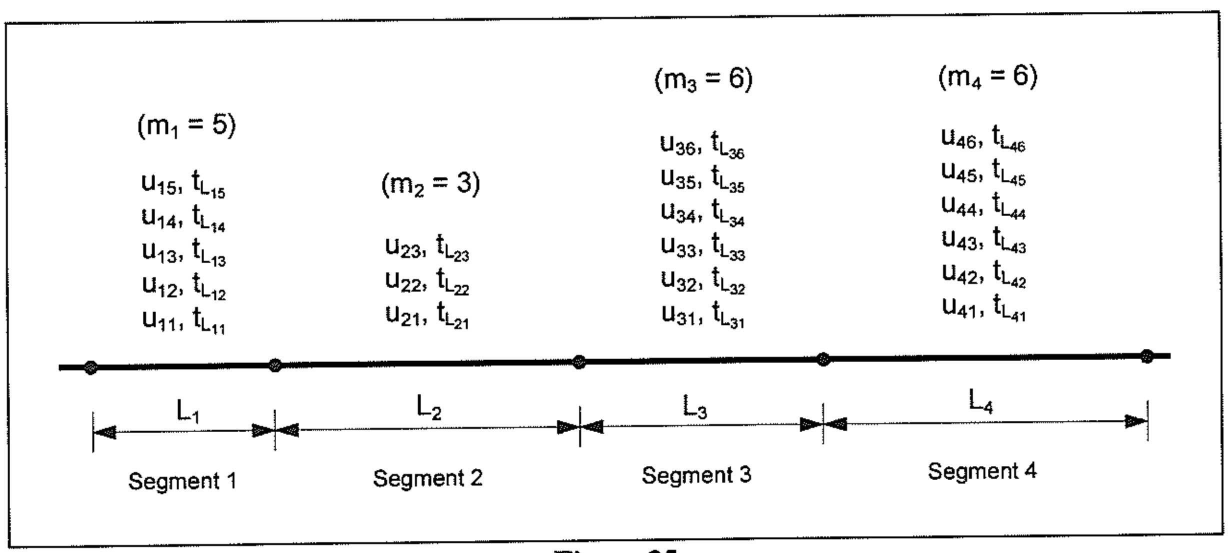


Figure 35
Sample segment speeds and travel times

Let the number of runs (or sample size) per segment be m_i. The average travel time per segment is

$$\bar{t}_i = \frac{1}{m_i}\sum_{j=1}^{m_i} t_{L_{ij}} = \frac{1}{m_i}\sum_{j=1}^{m_i}\frac{L_i}{u_{ij}} = L_i \frac{1}{m_i}\sum_{j=1}^{m_i}\frac{1}{u_{ij}} \tag{12}$$

where $t_{L_{ij}}$ is the j^{th} travel time record associated with segment i, L_i is the length of each segment, and u_{ij} is the j^{th} speed record associated with segment i. Notice that the m_i travel time records (or m_i speed records) for the segment can be obtained either with equation (4) or equation (8). Let the number of segments be n. The total distance L_T is

$$L_T = \sum_{i=1}^{n} L_i \tag{13}$$

The total travel time t_{T_L} over L_T is

$$t_{T_L} = \sum_{i=1}^{n}\bar{t}_i = \sum_{i=1}^{n}\left[L_i \frac{1}{m_i}\sum_{j=1}^{m_i}\frac{1}{u_{ij}}\right] \tag{14}$$

The average speed for all runs over L_T is

$$\bar{u}_L = \frac{L_T}{t_{T_L}} = \frac{\sum_{i=1}^{n} L_i}{\sum_{i=1}^{n}\bar{t}_i} = \frac{1}{\sum_{i=1}^{n}\left[\frac{L_i}{L_T}\frac{1}{m_i}\sum_{j=1}^{m_i}\frac{1}{u_{ij}}\right]} \tag{15}$$

The last part of this equation represents a weighted harmonic mean on segment speeds, where the weight is the ratio of the length of each segment to the total length considered. If all segments have the same length, the weight associated with each segment will be the same.

One disadvantage of using equation (15) to estimate average segment speeds is that the harmonic mean, by definition, is very sensitive to individual values which are much lower than the rest in the series. As a result, outlying low speeds (or outlying large travel times), which tend to occur on atypically adverse traffic conditions, could result in very small average speeds. Since the objective of using a formulation like equation (15) is to obtain a solid, robust estimator of central tendency, the question then is how appropriate the harmonic mean speed formulation is for estimating central tendency. After all, the central tendency estimator does not necessarily have to be the harmonic mean speed.

One possible modification to the harmonic mean formulation would be to manually filter out records that are "atypical." However, this does not appear to be feasible in an environment in which large amounts of data are being collected. A second possible modification would be to convert equation (12) so that median segment travel times are calculated instead of arithmetic mean segment travel times. The median is known for not being seriously affected by outliers and in many cases it is preferred by statisticians as a measurement of central tendency *[22]*. The end result would be a formulation similar to equation (12) that uses median segment travel times instead of arithmetic mean segment travel times. The median speed formulation would be

$$\bar{u}_L = \frac{L_T}{\sum_{i=1}^{n} t_{m_i}} = \frac{1}{\sum_{i=1}^{n}\left[\frac{L_i}{L_T}\frac{1}{u_{m_i}}\right]} \quad (16)$$

where t_{m_i} and u_{m_i} are median travel time and median speed associated with segment i during the time period being studied.

The equations described previously are general in the sense that they can be used either for one or several runs, and either for one or several contiguous segments. As an illustration, consider the GPS data from the run made on October 19, 1995, on Florida Boulevard EB (figure 4 and figure 8). For completeness, table 15 shows the GPS data associated with links 1779 and 1780 and summarizes the computational procedure. For example, for link 1779, using equation (4),

$$u \approx 3600 * \frac{4.5493 - 4.2880}{27{,}194.58 - 27{,}173.51} = \frac{0.2613}{21.07} = 44.63 \text{ mph}$$

where 21.07 s is the difference between the interpolated entrance time and interpolated exit times, or travel time t_L. Table 15 summarizes these values, as well as those corresponding to link 1780. Now, using equations (13), (14), and (15) to compute total travel time and speed on links 1779 and 1780,

$$L = 0.2613 + 0.1153 = 0.3766\text{mi}$$
$$t_{T_L} = 21.07 + 14.50 = 35.57 \text{ s}$$
$$\bar{u}_L = \frac{0.3766 * 3600}{35.57} = 38.11 \text{ mph}$$

By comparison, using equation (8) to compute speed on link 1779,

$$u \approx \frac{1}{2(27{,}193.75 - 27{,}173.75)}\left[(44.2 + 44.4) * 1 + \ldots + (41.6 + 37.1) * 3\right] = 45.01 \text{ mph}$$

Using equation (9),

$$t_L \approx 3600\left[\frac{0.2613}{45.01}\right] = 20.90 \text{ s}$$

Now, using equations (13), (14), and (15) to compute total travel time and speed on links 1779 and 1780,

$$L = 0.2613 + 0.1153 = 0.3766\text{mi}$$
$$t_{T_L} = 20.90 + 14.49 = 35.39 \text{ s}$$
$$\bar{u}_L = \frac{0.3766 * 3600}{35.39} = 38.31 \text{ mph}$$

Notice that the results based on equation (4) are very similar to those based on equation (8). This is a common situation, as long as the positional accuracy of the GPS data is acceptable and there is adequate GPS coverage on the segment. In the case of equation (4), adequate GPS coverage means no gaps in GPS data coverage in the vicinity of the segment entrance and in the vicinity of the segment exit. Sometimes, gaps in GPS data coverage in the vicinity of the segment entrance (or segment exit) may be acceptable if the speeds associated with the time stamps that would be used for time interpolation remain more or less constant (so that the linear interpolation approach is still valid). In the case of equation (8), adequate GPS coverage means no gaps in GPS data coverage throughout the segment. However, gaps in GPS data coverage may be acceptable if the variation of speeds along the segment is relatively low or if the gap occurs where the vehicle speed is relatively low (so that the effect of the gap in the speed integration process is minor). Conversely, gaps in the GPS data coverage may not be acceptable if the gaps occur where the vehicle speed is relatively high.

Table 15
Computation of link travel time and speed for run of October 19, 1995

GPS data			Link data			Equation (4) procedure				Eqn (8) procedure	
Time	Milepost	Speed	Code	Milepost		Interpolated time		u	t_L	u	t_L
				Entrance	Exit	Entrance	Exit				
(sec)	(mi)	(mph)		(mi)	(mi)	(sec)	(sec)	(mph)	(sec)	(mph)	(sec)
27,173.25	4.2848	44.2									
27,173.75	4.2910	44.2									
27,174.75	4.3033	44.4									
27,175.75	4.3158	45.1									
27,177.25	4.3347	45.7									
27,177.75	4.3410	45.8									
27,179.25	4.3600	46.0									
27,180.25	4.3729	46.7									
27,181.25	4.3859	47.1									
27,182.25	4.3989	47.4	1779	4.2880	4.5493	27,173.51	27,194.58	44.63	21.07	45.01	20.90
27,183.25	4.4121	47.6									
27,184.75	4.4321	47.8									
27,185.25	4.4388	47.8									
27,186.75	4.4587	47.3									
27,187.25	4.4653	46.8									
27,189.25	4.4906	44.0									
27,190.75	4.5084	41.6									
27,193.75	4.5407	37.1									
27,194.75	4.5510	36.9									
27,195.75	4.5612	37.1									
27,197.25	4.5767	37.1									
27,197.75	4.5817	36.7									
27,198.75	4.5916	34.5									
27,200.75	4.6086	27.0									
27,202.25	4.6188	22.8	1780	4.5493	4.6645	27,194.58	27,209.08	28.62	14.50	28.63	14.49
27,203.25	4.6251	22.4									
27,204.25	4.6312	22.3									
27,205.25	4.6375	22.7									
27,206.75	4.6473	24.7									
27,207.25	4.6508	25.5									
27,208.75	4.6620	27.5									
27,209.25	4.6658	28.3									

Because gaps in GPS coverage tend to be stochastic, it is not possible to determine beforehand on which formulation [equation (4) or equation (8)] the gap effect is more noticeable. Therefore, it may be advisable to compute segment speeds both with equation (4) and equation (8) and compare results. As an illustration, consider the link speed and travel time data shown in Table 16. This data corresponds to the run made on Florida Boulevard EB on October 19, 1995 (GPSFileCode = 178 in figure 8). In most cases, speeds based on equation (4) were very similar to those based on equation (8), and, as a result, the corresponding link travel times were also very similar. In some cases, differences in speed were large (-1.84 mph for link 1761, 1.48 mph for link 1777, and 1.22 mph for link 1787), and the corresponding differences in link travel time were also large (1.31 seconds, -4.37 seconds, and -1.64 seconds). In the case of link 1761, there were gaps in GPS coverage in the vicinity of the link exit. At the same time, speeds near the link exit were considerably lower than on the rest of the link. As a result, the gap effect on link speed was larger when using equation (4) than when using equation (8). Consequently, the link speed based on equation (8) could be considered more reliable. In the case of links 1777 and 1787, there were gaps in GPS coverage in the vicinity of the link entrances. At the same time, speeds near the link entrances were considerably larger than on the rest of each link. As a result, the gap effect on link speed was larger when using equation (8) (actually, much larger in the case of link 1777) than when using equation (4. Consequently, the link speeds based on equation (4) could be considered more reliable.

Table 16
Comparison between equation (4) and equation (8)

Link data		Equation (4) procedure		Equation (8) procedure		Differences	
Link code	Length (mi)	Speed (mph)	t_L (sec)	Speed (mph)	t_L (sec)	Speed (mph)	t_L (sec)
1760	0.0559	6.73	29.90	6.53	30.82	0.20	-0.92
1761	0.0677	17.65	13.81	19.49	12.50	-1.84	1.31
1762	0.0710	19.07	13.40	19.39	13.18	-0.32	0.22
1763	0.0736	19.04	13.91	19.32	13.71	-0.28	0.20
1764	0.0714	23.93	10.73	24.07	10.67	-0.13	0.06
1765	0.0727	29.44	8.89	30.33	8.63	-0.89	0.26
1766	0.1443	23.30	22.29	23.34	22.26	-0.03	0.03
1767	0.0581	22.31	9.37	23.90	8.74	-1.59	0.62
1768	0.6412	31.36	73.61	31.58	73.09	-0.22	0.52
1769	0.1963	20.43	34.58	20.03	35.28	0.41	-0.70
1770	0.0060	22.12	0.97	22.60	0.95	-0.48	0.02
1771	0.2956	15.05	70.71	14.88	71.53	0.17	-0.82
1772	0.3711	29.96	44.60	30.08	44.41	-0.13	0.19
1773	0.1372	32.18	15.35	32.89	15.02	-0.71	0.33
1774	0.7938	20.15	141.83	20.13	141.95	0.02	-0.12
1775	0.0061	16.65	1.32	17.20	1.27	-0.55	0.04
1776	0.4061	40.51	36.09	40.66	35.96	-0.15	0.13
1777	0.2810	19.28	52.47	17.80	56.84	1.48	-4.37
1778	0.5390	40.33	48.12	40.43	47.99	-0.11	0.13
1779	0.2613	44.63	21.07	45.01	20.90	-0.38	0.18
1780	0.1153	28.62	14.50	28.63	14.49	-0.02	0.01
1781	0.4731	15.24	111.75	15.23	111.83	0.01	-0.08
1782	0.3307	35.66	33.39	35.51	33.53	0.15	-0.14
1783	0.1528	43.92	12.53	44.10	12.48	-0.18	0.05
1784	0.0415	41.53	3.60	41.83	3.57	-0.30	0.03
1785	0.1726	39.66	15.66	39.95	15.55	-0.29	0.12
1786	0.3249	40.21	29.09	40.29	29.03	-0.08	0.06
1787	0.1630	21.49	27.30	20.27	28.95	1.22	-1.64
1788	0.1952	33.49	20.98	33.59	20.92	-0.10	0.06
1789	0.2662	13.80	69.44	13.50	70.98	0.30	-1.54
1790	0.2912	30.39	34.50	30.82	34.01	-0.43	0.49
Total	**7.0757**		**1038.08**		**1041.04**		**-2.96**

APPENDIX B: METHODOLOGY FOR COMPUTING DELAY AT SIGNALIZED INTERSECTIONS

An accurate field measurement of delay is important for the design and operation of traffic control systems. As a performance measure, delay plays a critical role in the determination of levels of service at signalized and unsignalized intersections. Delay is also included, explicitly or implicitly, in the calculation of average speeds used to determine levels of service on arterial streets.

Delay is defined and used in many different ways. For example, two definitions of delay commonly used for signalized intersections are approach delay and stopped delay. Following Reilly, Gardner, and Kell, approach delay is defined as the difference between the time used by any vehicle to traverse an approach delay section and the free flow time associated with that approach delay section *[18]*. Due to difficulties in determining the precise upstream and downstream limits of the deceleration and acceleration sections, Reilly, Gardner, and Kell defined the approach delay section as a section of roadway of fixed length beginning at an assigned distance upstream of the intersection and ending at the intersection stop bar. Their implicit assumption was that the delay outside this approach delay section, particularly that due to acceleration after the intersection, was negligible. Strictly speaking, however, such an assumption ignored two key characteristics of traffic flow at signalized intersections: (1) deceleration does not necessarily begin at a fixed distance upstream of the intersection and (2) a significant amount of acceleration occurs after the intersection stop bar. Furthermore, precisely tracing individual vehicle trajectories has been, until now, very laborious and time consuming *[13]*. As a result of these limitations, the approach delay concept is rarely used in practice.

Stopped delay is the time a vehicle is stopped at an intersection. There are several techniques to measure stopped delay, including the "stopped-vehicle count" technique, the "arrival and departure volume" technique, and the "test car" technique *[12]*. Of these, the "stopped-vehicle count" technique is the most common approach to measure stopped delay *[12]*. With this technique, stopped delay is measured indirectly based on the number of stopped vehicles recorded at specific time intervals and the volume of vehicles crossing the stop bar during the total duration of the survey. The technique is relatively simple to execute, although it may be quite labor intensive if long queues are present *[23]*. In addition to this, questions remain with respect to the accuracy of the technique, particularly in situations of low volumes and short red intervals *[14]*.

Strictly speaking, the speed of a vehicle must be zero in order to be considered stopped. In practice, a vehicle which is crawling forward in a queue or that has a gap of less than three car lengths with respect to the vehicle in front of it may also be considered as being stopped *[18]*, *[23]*. While conceptually sound, this approach has several problems including the difficulty to measure vehicle positions and speeds accurately and consistently. As a result, it is extremely difficult to determine whether a vehicle's speed is below a certain threshold to be considered crawling, or whether its gap with respect to the vehicle in front of it is less than three car lengths. When using GPS receivers to automatically track the motion of probe vehicles, the question reduces to that of finding a consensus regarding appropriate threshold values. Here, we assume that a vehicle is stopped when (a) the speed of the vehicle is zero or (b) the vehicle had already stopped and is crawling within a standing queue such that the average acceleration of a series of 10 adjacent GPS points is less than a prescribed threshold. In our work, we typically used threshold values in the neighborhood of 1.0-1.5 mph/s.

Despite the limitations of the stopped delay approach, the consensus has been that stopped delay is more reliable and easier to measure in the field than approach delay. Because of this, stopped delay has been the traditional performance measure used to determine levels of service at signalized intersections. To convert from stopped delay to approach delay, conversion factors such as the 0.76 factor included in equation (9-24) of the 1994 HCM are normally used *[12]*. However, several investigators have questioned the use of such factors on the basis that the relationship between approach delay and stopped delay is not constant *[13]*, *[14]*. In addition to this, an increasing body of evidence indicates that procedures such as those of the 1994 HCM do not properly address many real-world situations like oversaturated conditions, platoon dispersion, and queue spillback *[18]*. This has prompted the development of alternative procedures, based on signal (or control) delay as opposed to stopped delay, to evaluate the performance of arterial streets *[24]*. In the alternative procedures, signal (or control) delay includes initial deceleration delay, stopped delay, and final acceleration delay.

The new alternative procedures based on control delay have also prompted a need for improved procedures to measure delay in the field. Because GPS data can be linearly referenced as shown in figure 8, we can build detailed distance-time diagrams like that shown in figure 36. For completeness, figure 36 also shows the main delay terms at a signalized intersection. Before point 1 on the time-distance diagram, the vehicle is moving at a uniform speed. From point 1 to point 2, the vehicle decelerates until it stops at point 2 to join the standing queue before the signalized intersection. The vehicle remains stopped between points 2 and 3. Between points 3 and 5, the vehicle accelerates until it reaches a uniform speed at point 5. Notice that when the vehicle crosses the stop bar (point 4), the vehicle is still accelerating.

Following the definitions described previously, d_s represents stopped delay and is given by

$$d_s = t_3 - t_2 \tag{17}$$

Similarly, d_c represents control delay and is given by

$$d_c = (t_5 - t_1) - \frac{L_5 - L_1}{s_f} \tag{18}$$

where s_f is the free flow speed. For comparison purposes, figure 36 also shows a generic approach delay d_{ap}, which is not constrained by a fixed distance upstream of the signalized intersection. This delay is given by

$$d_{ap} = (t_4 - t_1) - \frac{L_4 - L_1}{s_f} \tag{19}$$

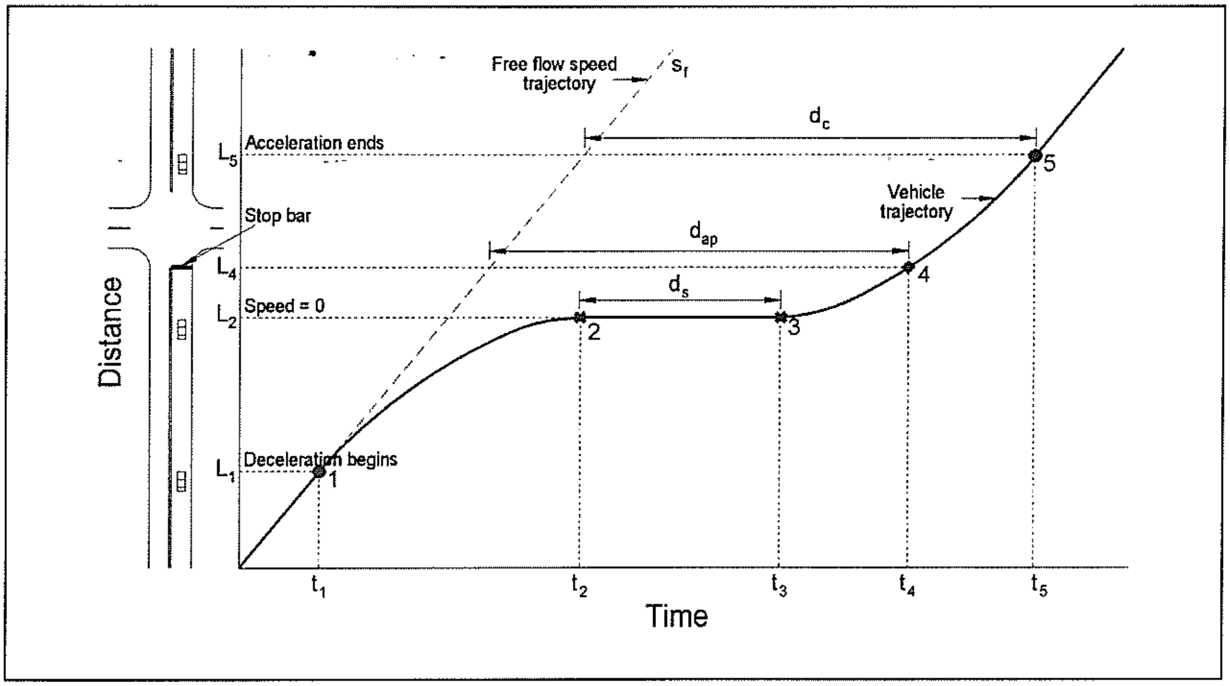


Figure 36
Schematic distance-time diagram depicting all delay terms at a signalized intersection

A diagram such as that shown in figure 36 is useful for defining the general shape of the distance-time relationship associated with a specific vehicle. However, it is not particularly suitable for precisely locating the delay control points for each signalized intersection (points 1, 2, 3, and 5) because the deceleration and acceleration components of the distance-time diagram tend to be asymptotic to the uniform speed components in the neighborhood of those points. To accurately locate these control points, we need to analyze speeds and their rate of change. In other words, we must apply a filter to the original distance-time diagram to make major changes in the distance-time diagram clearly visible.

To illustrate this point, consider the distance-time diagram shown in figure 37a. This diagram is based on a run made on Florida Boulevard EB in Baton Rouge on October 19, 1995 (file *10191129.txt* in figure 8). As shown in figure 37a, the vehicle stopped at three signalized intersections: North Foster, Ardenwood, and Wooddale. For comparison purposes, figure 37a also shows the distance-time diagram associated with a vehicle traveling at the posted speed limit. For this particular application, we assumed the posted speed limit to be the same as the free flow speed. Figure 37b shows the corresponding speed-time diagram. Visually, this speed-time diagram appears to be very efficient for detecting when the speed of the vehicle is zero and when the vehicle begins to decelerate but somewhat less efficient for detecting when the vehicle stops accelerating. From the point of view of implementing an automated procedure to detect these points, the fact that the vehicle does not always begin to decelerate from the same speed, or does not always accelerate until it returns to the same speed, imposes an additional complexity to the algorithm because of the difficulty in setting up a reference speed value against which speed variations are computed.

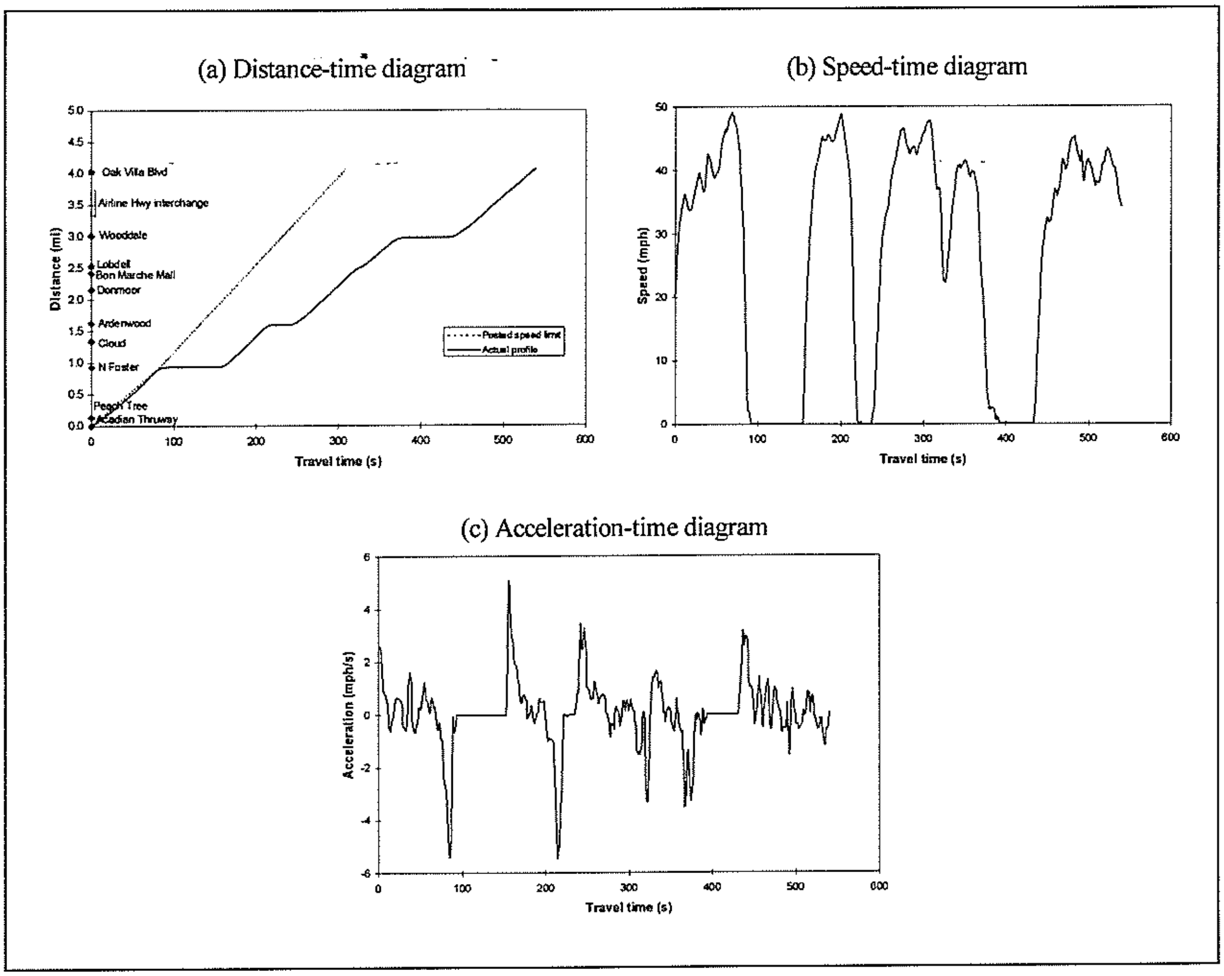


Figure 37

Distance-time, speed-time, and acceleration-time diagrams for a run made on Florida Boulevard EB in Baton Rouge, on October 19, 1995

One way to simplify the problem of varying uniform speeds is by locating major changes in speed using acceleration values. This way, the problem reduces to that of determining when the acceleration is zero, significantly larger than zero, or significantly lower than zero. Because the accuracy of individual GPS speed measurements in our research was higher than that associated with computing speed based on the location of two contiguous GPS points, we compute acceleration directly as the first derivative of speed with respect to time. For each GPS point, we compute acceleration using a central difference scheme as follows:

$$a_i = \frac{v_{i+1} - v_{i-1}}{t_{i+1} - t_{i-1}} \tag{20}$$

where

a_i = acceleration associated with GPS point i
v_{i+1}, v_{i-1} = speeds associated with GPS points i+1 and i-1, respectively
t_{i+1}, t_{i-1} = time stamps associated with GPS points i+1 and i-1, respectively.

Figure 37c shows the acceleration-time diagram for the test vehicle run on Florida Boulevard. In order to determine when the acceleration was significantly different from zero, we considered the use of a low pass filter such as the Butterworth filter *[25]*. To do this, we plotted the power spectral density of the acceleration profile to determine meaningful frequencies in the original data. With this information, we designed a Butterworth filter to filter out frequencies that were higher than the ones we considered meaningful. The result was an acceleration profile that was smoother than the original profile. Unfortunately, the smoothed acceleration profile also smoothed edges that were important for accurately defining when the vehicle began and stopped accelerating or decelerating.

To avoid the problem of having smoothed edges at critical points, we considered a different filtering scheme. To detect when the vehicle started to accelerate or decelerate (i.e. when the vehicle acceleration began to be significantly different from zero), we used a 10-point forward average acceleration algorithm. With this algorithm, we computed the average of acceleration values associated with 10 points (the point in question and nine points ahead) and compared the average value with a threshold value (typically between 0.5 mph/s and 1.5 mph/s). If the average value was larger than the threshold, the filtered acceleration a_{if} was made the same as the original acceleration value. Otherwise, it was made zero. In other words,

$$a_{if} = \begin{cases} a_i, \text{ if } 0.1\sum_{i}^{i+9} a_i > \text{threshold} \\ 0, \text{ if } 0.1\sum_{i}^{i+9} a_i \leq \text{threshold} \end{cases} \tag{21}$$

Likewise, to detect when the vehicle stopped accelerating or decelerating (i.e. when the vehicle acceleration became essentially zero), we used a 10-point backward average acceleration algorithm. With this algorithm, we computed the average of acceleration values associated with 10 points (the point in question and nine points backward) and compared the average value with a threshold value (typically between 0.5 mph/s and 1.5 mph/s). If the average value was larger than the threshold, the filtered acceleration a_{ib} was made the same as the original acceleration value. Otherwise, it was made zero. In other words,

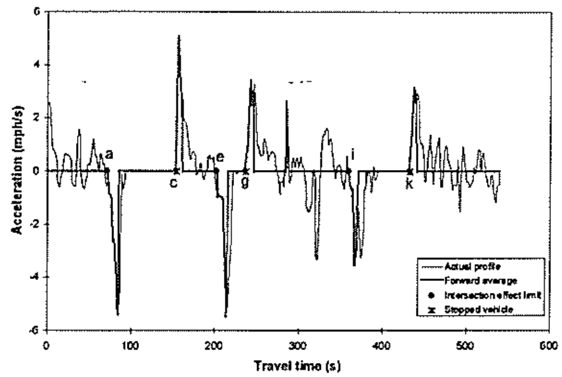
$$a_{ib} = \begin{cases} a_i, \text{ if } 0.1\sum_{i-9}^{i} a_i > \text{threshold} \\ 0, \text{ if } 0.1\sum_{i-9}^{i} a_i \leq \text{threshold} \end{cases} \tag{22}$$

We implemented these algorithms and the equations for calculating control delay, approach delay, and stopped delay using Visual Basic macros in MS Excel (the accompanying CD has a sample Excel spreadsheet application). In a typical session, the user enters the linearly referenced GPS data and defines an acceleration threshold, for example 1 mph/s. With this information, the application calculates forward average and backward average accelerations for each GPS point, checks for points where the speed is zero, and executes a macro that automatically marks locations where the forward average accelerations stop being zero and where the backward average accelerations become zero. The application then prints a summary report which includes the distance-time, speed-time, and acceleration-time diagrams, as well as a table containing control delay, approach delay, and stopped delay at each signalized intersection. As mentioned previously, the user can interactively modify the acceleration threshold to increase or decrease the sensitivity of the algorithm to individual acceleration values.

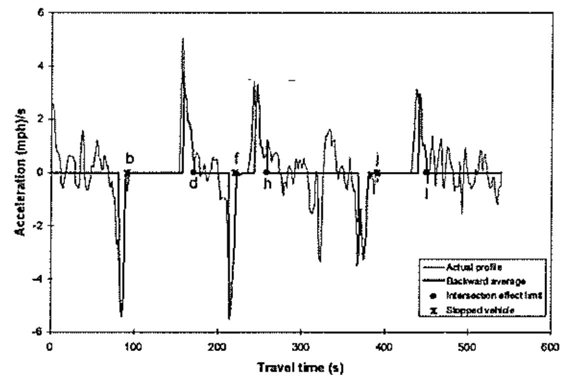
For illustration purposes, figure 38 summarizes the process of locating the delay control points for a run made on Florida Boulevard EB on October 19, 1995 (figure 37). Figure 37c shows the original acceleration-time diagram. Figure 38a shows the results of applying the 10-point forward average acceleration algorithm [equation (21)]. This algorithm effectively detected when the vehicle began to decelerate (acceleration became negative) at the North Foster, Ardenwood, and Wooddale signalized intersections (points a, e, and i in figure 37a) while filtering out minor deceleration noise. Likewise, the algorithm effectively detected when the vehicle started to accelerate (acceleration became positive) at the same signalized intersections (points c, g, and k in figure 37a). However, the algorithm performed poorly when detecting when the vehicle stopped decelerating or accelerating. By comparison, figure 38b shows the results of applying the 10-point backward average acceleration algorithm [equation (22)]. In this case, the backward average algorithm effectively detected when the vehicle stopped decelerating (because the vehicle speed became zero) (points b and f in figure 38b), or when the vehicle stopped accelerating (because the vehicle speed became essentially uniform) (points d, h, and l in figure 38b). In the case of point j, the algorithm computed a few backward average acceleration values as zero even though the vehicle had not stopped yet. As shown in figure 37b, the vehicle was essentially crawling at about 2 mph for about 10 seconds before it stopped. In this case, some manual editing was required to locate point j as the point where the vehicle was effectively stopped with locked wheels. As expected, the backward average algorithm performed poorly when detecting when the vehicle started decelerating or accelerating. For completeness, figure 38c shows the distance-time diagram, with a clear indication of the location and time associated with the significant changes in acceleration delineating the limits of the intersection effect on delay. Figure 38 d shows the control delay, approach delay, and stopped delay at the signalized intersections where the probe vehicle stopped.

In the example shown in figure 38, we used an acceleration threshold of 1.5 mph/s. If we had used 1 mph/s, the algorithm would have detected changes in acceleration in more locations, for example at the Lobdell intersection, even though the vehicle did not stop there. Because each travel time run is different (different time, location, driver, and traffic conditions), the user of the spreadsheet application may need to interactively modify the acceleration threshold to define a suitable value that will enable the detection of only major changes in acceleration.

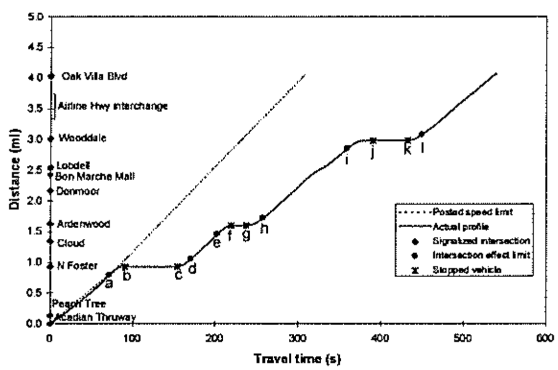
(a) 10-point forward average acceleration-time diagram for locating start of acceleration or deceleration



(b) 10-point backward average acceleration-time diagram for locating end of acceleration or deceleration



(c) Complete distance-time diagram with start and end points of acceleration and deceleration



(d) Control delay, approach delay, and stopped delay

Intersection	Delay (sec)		
	Control	Approach	Stopped
Acadian Thruway			
Peach Tree Dr			
North Foster Dr	75.3	71.7	63.5
Cloud Dr			
Ardenwood Dr	36.1	31.0	17.5
Donmoor Ave			
Bon Marche Mall			
Lobdell Ave			
Wooddale Ave	73.9	69.6	42.0
Oak Villa Blvd			

Figure 38
Computation of delay for a run made on Florida Boulevard EB in Baton Rouge on October 19, 1995

APPENDIX C: DATA COLLECTION SUMMARY IN BATON ROUGE

Notes:

1. TT (Traffic type): a.m. (a.m. peak); p.m. (p.m. peak); o.p. (off peak).
2. DI (Driver Initial): code associated with student driver.
3. Code: Unique identification code associated with a GPS data file.
4. Name: GPS data file name.
5. # Pts: Number of records contained in the file (when applied to raw GPS data files); or number of records in a file that were effectively associated with routes, processed, and imported into the Access database (when applied to the reduced data).
6. percent DC (Differential Correction): measure of the proportion of GPS points that had differential correction while being collected in the field.
7. Dupl: Number of records in a GPS file that were effectively associated with routes but that were duplicate records. These records were not imported into the Access database.
8. This file contained data which was not collected on the highway network.

DATA COLLECTION SUMMARY IN BATON ROUGE

Raw GPS Data Files						Reduced Data	
TT	DI	Code	Name	# Pts	% DC	# Pts	Dupl
a.m.	bm	241	12011218	6,428	99.8	5,548	10
p.m.	pt	242	12012143	7,033	99.9	5,278	4
p.m.	ms	243	12012208	976	100.0	399	0
p.m.	ms	244	12012229	6,057	99.9	2,699	1
a.m.	bm	245	12041229	6,897	100.0	5,493	8
p.m.	pt	246	12042041	10,475	99.9	8,038	13
p.m.	pt	247	12052144	5,418	99.9	4,097	2
p.m.	pt	248	12062131	4,134	99.9	2,455	6
a.m.	bm	249	12071228	3,118	100.0	1,256	0
a.m.	ms	250	12071229	7,907	99.9	5,423	10
p.m.	bp	251	12072134	7,280	99.9	4,466	8
a.m.	bm	252	12081234	4,089	100.0	3,577	7
p.m.	bp	253	12081942	12,060	100.0	7,813	5
p.m.	ms	254	12082151	8,032	99.9	7,353	6
a.m.	bm	255	12111216	3,215	100.0	2,216	7
p.m.	pt	256	12112146	6,665	100.0	5,676	2
a.m.	bm	257	12121228	4,958	100.0	4,326	3
p.m.	pt	258	12122047	7,936	100.0	5,558	4
a.m.	bm	259	12131220	6,874	100.0	5,954	7
p.m.	bp	260	12132107	6,700	100.0	5,223	1
p.m.	pt	261	12132111	7,198	99.9	5,540	1
a.m.	bm	262	12141224	3,281	100.0	1,280	2
p.m.	bp	263	12142203	6,038	99.2	916	10
o.p.	ma	264	12181440	4,878	99.8	3,732	2
o.p.	ma	265	12181844	6,961	99.9	5,980	10
p.m.	ma	266	12182104	2,954	100.0	2,163	0
a.m.	ma	267	12191223	5,988	99.9	5,289	5
a.m.	bm	268	12191232	11,324	100.0	10,619	7
o.p.	bp	269	12191840	4,058	99.9	1,143	3
o.p.	ma	270	12191438	5,740	99.9	5,145	2
o.p.	ma	271	12191901	3,857	100.0	2,299	1
p.m.	ma	272	12192122	4,260	99.8	2,350	3
p.m.	bp	273	12192123	5,348	99.9	3,877	2
o.p.	ma	274	12201539	5,677	99.7	4,279	6
o.p.	ms	275	12201608	3,715	100.0	2,630	1
o.p.	ms	276	12201837	6,491	99.4	5,283	139
o.p.	ma	277	12201838	5,832	100.0	4,563	11
o.p.	ma	278	12211603	3,004	100.0	2,807	3
o.p.	ms	279	12211632	4,408	100.0	1,424	65
o.p.	ms	280	12211824	4,876	100.0	3,923	5
o.p.	ma	281	12211854	5,692	99.9	5,131	7
o.p.	ma	282	12221453	2,876	99.9	1,970	1
o.p.	ma	283	12271523	4,139	99.1	3,540	6
o.p.	ma	284	12271834	5,561	100.0	5,294	6
o.p.	ma	285	12281510	5,035	100.0	3,651	2
o.p.	ma	286	12281838	5,993	99.9	5,406	4
a.m.	bm	287	01041219	11,115	100.0	8,625	7
a.m.	bm	288	01051214	6,387	99.9	5,616	3
o.p.	bm	289	01051423	3,146	100.0	2,948	0
a.m.	bm	290	01081229	7,486	100.0	6,501	3
a.m.	bm	291	01091217	6,534	100.0	5,229	3
a.m.	bm	292	01111207	6,983	99.8	5,983	5
p.m.	bp	293	01112151	8,030	99.9	6,310	7
a.m.	bm	294	01121215	8,776	99.9	7,705	6
a.m.	bm	295	01161223	6,391	100.0	5,257	2
a.m.	ms	296	01161240	817	100.0	—	Note 8
a.m.	bm	297	01171211	6,373	100.0	5,661	6
a.m.	bm	298	01181239	3,861	100.0	3,255	0
a.m.	ms	299	01181301	3,429	99.9	3,162	5
p.m.	bp	300	01182207	6,036	99.8	3,512	2

Raw GPS Data Files						Reduced Data	
TT	DI	Code	Name	# Pts	% DC	# Pts	Dupl
a.m.	bm	301	01191216	8,318	99.9	6,947	2
p.m.	ms	302	01192324	2,188	100.0	1,665	1
a.m.	bm	303	01221219	6,956	100.0	5,610	4
a.m.	bm	304	01231234	5,876	99.8	5,275	36
a.m.	ms	305	01231245	4,821	100.0	4,230	16
p.m.	bp	306	01232135	5,707	99.7	3,661	1
a.m.	bm	307	01241210	4,248	99.9	3,424	4
a.m.	bm	308	01251211	3,385	99.9	2,420	3
p.m.	bp	309	01252157	3,664	99.9	2,251	1
a.m.	bm	310	01291212	8,211	100.0	6,784	4
a.m.	bm	311	01301220	7,020	100.0	5,817	3
p.m.	bp	312	01302229	5,229	99.9	4,016	4
a.m.	bm	313	01311211	6,974	100.0	6,366	3
a.m.	bm	314	02011222	5,482	100.0	4,929	3
p.m.	bp	315	02012242	8,439	100.0	6,101	12
a.m.	bm	316	02051223	5,479	100.0	4,398	4
a.m.	bm	317	02061228	5,496	99.7	4,735	7
p.m.	bp	318	02062200	7,508	100.0	5,341	7
a.m.	bm	319	02071212	8,720	99.9	7,259	5
a.m.	bm	320	02081240	4,771	99.9	3,895	124
p.m.	bp	321	02082258	4,943	100.0	3,196	5
a.m.	bm	322	02091222	7,850	95.1	6,650	40
p.m.	ms	323	02092047	1,529	100.0	641	0
p.m.	ms	324	02092122	8,198	99.9	6,685	9
a.m.	bm	325	02121222	7,176	100.0	6,523	16
a.m.	bm	326	02131208	5,426	100.0	4,378	22
p.m.	bp	327	02132229	6,609	100.0	4,987	10
a.m.	bm	328	02151210	5,669	99.6	4,659	3
p.m.	bp	329	02152159	1,707	100.0	1,171	1
p.m.	bp	330	02152258	5,453	100.0	4,160	2
a.m.	bm	331	02161230	6,381	100.0	3,869	6
a.m.	bm	332	02221215	6,340	97.5	5,126	6
a.m.	ms	333	02221245	5,541	94.3	4,755	14
p.m.	bp	334	02222208	7,813	98.5	5,321	4
a.m.	bm	335	02231207	7,346	98.5	5,993	5
p.m.	ms	336	02232217	5,721	99.7	4,753	3
a.m.	bm	337	02261210	3,023	100.0	2,092	3
a.m.	bm	338	02271208	4,139	99.5	2,837	0
p.m.	bp	339	02272102	6,378	99.8	4,733	1
a.m.	bm	340	02281214	3,426	100.0	2,516	2
p.m.	bp	341	02292144	9,064	98.6	7,642	3
p.m.	ms	342	03012230	4,317	97.3	3,530	5
a.m.	ms	343	03051227	7,579	99.4	5,548	7
p.m.	bm	344	03052139	6,333	98.8	5,266	6
p.m.	bp	345	03052231	7,008	99.4	6,637	5
p.m.	bm	346	03062132	5,440	99.4	4,225	3
p.m.	bm	347	03072140	6,203	99.9	4,342	4
p.m.	bp	348	03072149	8,220	100.0	5,764	2
p.m.	ms	349	03082155	1,954	88.2	1,470	4
p.m.	ms	350	03082247	3,520	98.7	3,088	12
a.m.	ms	351	03121252	6,634	98.2	4,939	17
a.m.	ms	352	03141235	6,725	100.0	6,016	17
p.m.	ms	353	03152242	4,709	98.5	4,070	7
a.m.	ms	354	03191308	4,369	99.7	3,033	64
p.m.	bp	355	03192213	4,043	0.0	3,045	2
a.m.	ms	356	03211241	6,578	100.0	4,993	11
p.m.	bm	357	03212143	4,809	99.8	3,793	5
p.m.	bm	358	03222141	4,463	99.3	3,672	4
p.m.	ms	359	03222146	3,705	99.8	3,014	2
p.m.	ms	360	03222303	1,435	100.0	999	1

DATA COLLECTION SUMMARY IN BATON ROUGE

Raw GPS Data Files						Reduced Data	
TT	DI	Code	Name	# Pts	% DC	# Pts	Dupl
p.m.	bm	361	03252157	3,498	100.0	3,017	2
a.m.	ms	362	03261319	6,283	99.9	5,008	1
p.m.	bm	363	03262139	5,525	99.8	4,783	6
p.m.	bp	364	03262145	7,317	98.0	6,396	4
p.m.	bm	365	03282148	3,490	97.6	2,196	43
p.m.	bp	366	03282225	6,046	99.9	4,229	2
p.m.	bm	367	03292151	4,377	99.2	2,671	3
a.m.	ms	368	04091213	6,458	99.8	4,923	12
p.m.	bp	369	04092118	8,915	100.0	6,836	43
a.m.	ms	370	04111132	6,933	100.0	5,581	7
p.m.	bp	371	04112202	6,894	99.8	6,160	10
p.m.	ms	372	04122113	2,601	100.0	2,344	1
p.m.	ms	373	04122226	3,288	100.0	2,397	7
a.m.	ms	374	04161046	6,687	100.0	4,864	13
p.m.	bm	375	04162056	4,173	99.8	3,211	4
p.m.	bp	376	04162102	8,704	100.0	7,001	10
p.m.	bm	377	04172102	3,525	99.5	2,677	6
p.m.	bm	378	04182107	6,322	99.9	4,546	13
p.m.	bm	379	04192042	4,436	99.9	3,550	3
p.m.	bm	380	04232033	6,753	99.9	5,957	5
p.m.	bp	381	04232215	5,489	99.3	4,831	9
a.m.	ms	382	04231205	1,823	90.6	1,113	3
p.m.	bm	383	04252116	5,601	99.9	4,398	16
p.m.	bp	384	04252038	8,569	99.9	7,151	3
a.m.	ms	385	04251150	4,659	96.7	3,679	8
p.m.	ms	386	04262018	5,902	100.0	5,400	2
a.m.	ms	387	04301114	5,223	98.8	310	9
p.m.	bm	388	04302057	6,384	99.9	4,741	12
p.m.	bp	389	04302144	5,713	99.4	4,569	7
p.m.	bp	390	05022158	6,923	99.8	3,734	5
o.p.	mn	391	05141354	6,696	99.4	2,374	1
o.p.	mn	392	05141832	6,014	99.8	4,392	9
o.p.	mn	393	05151506	3,770	99.3	2,668	3
o.p.	mn	394	05151856	2,674	98.0	1,732	1
o.p.	mn	395	05161316	7,174	99.4	3,943	62
o.p.	mn	396	05161916	5,892	0.0	4,780	8
o.p.	mn	397	05171456	6,393	99.8	4,827	5
o.p.	mn	398	05171932	2,699	99.9	1,866	1
o.p.	mn	399	05201359	1,997	97.7	1,341	1
o.p.	mn	400	05201833	8,669	99.2	2,897	7
o.p.	mn	401	05211335	4,014	96.1	2,654	4
a.m.	mn	402	05221201	4,232	98.2	3,169	1
p.m.	mn	403	06141519	5,182	99.6	4,712	6
a.m.	mn	404	06170639	4,942	99.8	4,388	0
p.m.	mn	405	06171530	4,771	99.7	4,363	8
a.m.	mn	406	06200650	6,709	99.8	5,992	2
p.m.	mn	407	06201625	2,774	99.8	2,336	2
p.m.	mn	408	0624m001	4,265	0.0	4,143	0
a.m.	mn	409	0625m001	5,696	0.0	5,006	0
p.m.	mn	410	0625m002	5,732	0.0	5,323	0
p.m.	mn	411	07011539	4,776	98.7	4,422	4
a.m.	mn	412	07100703	4,687	48.9	4,361	4
p.m.	mn	413	07121557	3,709	7.9	3,581	1
a.m.	mn	414	07160624	1,285	98.6	1,215	22
a.m.	mn	415	07160756	1,894	94.1	1,825	34
p.m.	mn	416	07161518	5,634	99.9	5,463	10
a.m.	mn	417	07170630	4,327	99.5	4,077	58
p.m.	mn	418	07171524	4,438	6.4	4,161	11
a.m.	mn	419	07180618	6,741	13.8	6,229	4
a.m.	mn	420	07220647	5,665	99.8	4,665	70

Raw GPS Data Files						Reduced Data	
TT	DI	Code	Name	# Pts	% DC	# Pts	Dupl
p.m.	mn	421	07221517	4,015	99.6	3,737	5
a.m.	mn	422	07230624	1,849	33.3	1,787	7
p.m.	mn	423	07231554	5,222	99.8	4,972	9
a.m.	mn	424	07240619	4,067	100.0	3,713	2
p.m.	mn	425	07241537	5,882	100.0	5,550	5
a.m.	mn	426	07250638	4,141	100.0	3,787	6
a.m.	mn	427	07260648	7,124	100.0	6,609	12
p.m.	mn	428	07261532	2,590	100.0	2,504	5

This public document is published at a total cost of $1238.36. Two hundred and twenty copies of this public document were published at a cost of $750.86. The total cost of all printings of this document including reprints is $1238.36. This document was published by Louisiana State University, Graphic Services, 3555 River Road, Baton Rouge, Louisiana 70802, to report and publish research findings of the Louisiana Transportation Research Center as required in R.S.48:105. This material was printed in accordance with standards for printing by state agencies established pursuant to R.S.43:31. Printing of this material was purchased in accordance with the provisions of Title 43 of the Louisiana Revised Statutes.