PERGAMON

Transportation Research Part C 8 (2000) 287–306

TRANSPORTATION
RESEARCH
PART C

www.elsevier.com/locate/trc

Performance measures and data requirements for congestion management systems

Cesar A. Quiroga *

Texas Transportation Institute, Texas A&M University System, 3500 N.W. Loop 410, Suite 315, San Antonio, TX 78229, USA

Abstract

Many metropolitan areas have started programs to monitor the performance of their transportation network and to develop systems to measure and manage congestion. This paper presents a review of issues, procedures, and examples of application of geographic information system (GIS) technology to the development of congestion management systems (CMSs). The paper examines transportation network performance measures and discusses the benefit of using travel time as a robust, easy to understand performance measure. The paper addresses data needs and examines the use of global positioning system (GPS) technology for the collection of travel time and speed data. The paper also describes GIS platforms and sample user interfaces to process the data collected in the field, data attribute requirements and database schemas, and examples of application of GIS technology for the production of maps and tabular reports. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Congestion management systems; Travel time; GPS; GIS; Dynamic segmentation; Performance measures

1. Introduction

Traffic congestion is a critical problem in urban areas. Several indicators confirm this trend. For example, between 1976 and 1996, the number of vehicle-miles traveled (VMT) in the United States increased by 77%, while the mileage of roads and streets increased only by 2% (FHWA, 1998). Over the years, the percentage of the peak-hour VMT that occurs under congested conditions has steadily increased, although at a slower pace in recent years. In 1996, that percentage was 54% for the urban interstate system and 45% for the urban national highway system. Congestion usually results in time delays, increased fuel consumption, pollution, stress, health hazards, and added

* Tel.: +1-210-731-9938; fax: +1-210-731-8904.
E-mail address: c-quiroga@tamu.edu (C.A. Quiroga).

PII: S0968-090X(00)00008-5

vehicle wear. The associated cost is huge. For example, in 1997 congestion cost travelers in 68 urban areas in the United States 4.3 billion hours of delay, 6.6 billion gallons of wasted fuel consumed, and $72 billion of time and fuel cost (Schrank and Lomax, 1999).

Travel delay is perceived by many as the most noticeable impact of congestion. Not surprisingly, numerous efforts have been made to eliminate it or, at least, to alleviate its effects. In the past, adding capacity was considered the main solution to eliminate or reduce travel delays. However, this approach has frequently proved to be insufficient. Faced with this reality, many urban areas have opted for implementing alternative management measures. Examples of these alternative management measures include improved traffic surveillance and control systems, dedicated high-occupancy vehicle (HOV) lanes, improved transit service, and congestion and parking pricing. The objective of these measures is to manage and reduce congestion by improving traffic flow, enhancing mobility and safety, and reducing demand for car use. Many of these management measures are the result of federal, state or local legislation. Such is the case, for example, of the measures that had to be implemented in urban areas designated as non-attainment for ozone or carbon monoxide (CAAA, 1990; ISTEA, 1991) and that continue to be eligible for federal funding under TEA 21 (1998).

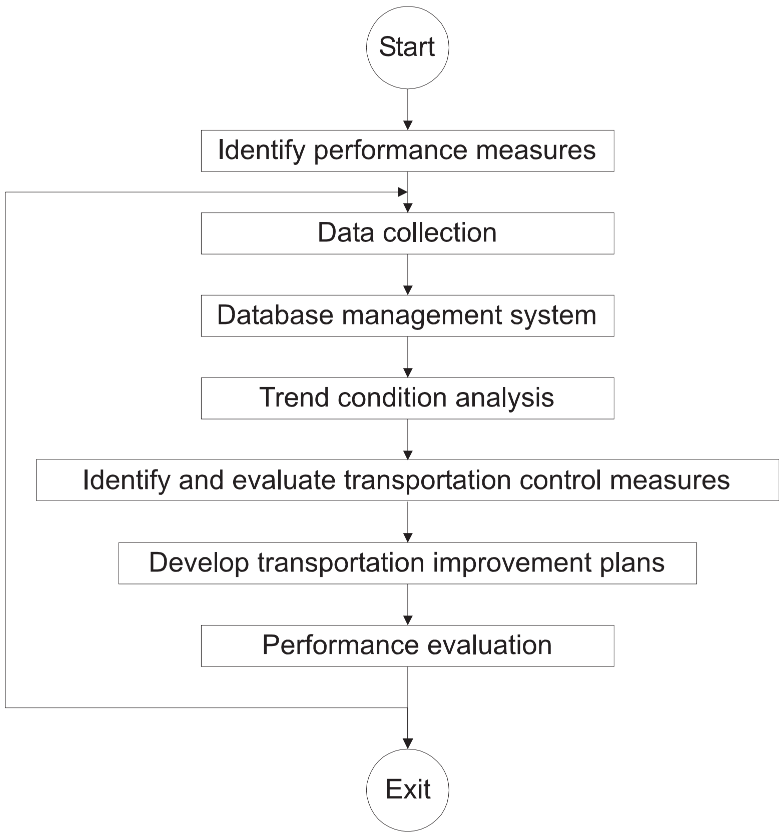

Fig. 1. Possible CMS flowchart.

A congestion management system (CMS) is intended to provide information on transportation system performance and identify alternative actions to alleviate congested roadway conditions (Lindquist, 1999). Although a CMS can use data obtained from a traffic operation surveillance and control system, in general a CMS is considered to be more a planning tool than an operations tool. This characteristic determines to a large degree the type of data collected and the types of data collection procedures implemented. In general, a CMS follows a decision making process that involves the identification of performance measures, a system-wide program of data collection and system monitoring, identification and evaluation of possible transportation control measures (TCMs), development of transportation improvement plans, and evaluation of the effectiveness of implemented strategies. Fig. 1 summarizes this process.

This paper focuses on those steps in Fig. 1 that have a strong geographic information system (GIS) component: identification of performance measures, data collection, data management, and identification and evaluation of transportation control measures. However, rather than focusing on GIS per se, this paper attempts to provide a review of technical and non-technical issues that influence the development of CMSs from a transportation engineering perspective and describes the role that GIS techniques can play in that development. It summarizes recent work in the area of travel time-based performance measures and describes the results of a recent project that involved the use of global positioning system (GPS) technology for the collection of travel time and speed data.

2. Performance measures

Numerous performance measures have been suggested or used for quantifying congestion (Francois and Willis, 1995; Schwartz et al., 1995; Lomax et al., 1997). Following Lomax et al., (1997), congestion measures can be grouped into highway capacity manual (HCM) measures, queuing-related measures, and travel time-based measures. Three commonly used HCM measures are volume to capacity (V/C) ratio, average intersection delay, and level of service (LOS). V/C ratios are frequently used because of the relative ease of traffic volume data collection and because surrogate measures such as LOS can be derived from V/C values. Average intersection delays are normally used on arterial streets and also provide a foundation for surrogate measures such as LOS. LOS measures normally range from "A" – best service to "F" – worst service. While conceptually simple and easy to understand by the professional transportation community, HCM performance measures tend to be somewhat abstract for the traveling public. They usually require detailed, location-specific input data, which makes them more appropriate for localized analyses and design than for area-wide planning. Finally, HCM measures are difficult to use for long-range comparisons because concepts such as capacity and speed-flow rate relationships tend to change over time (HRB, 1950; HRB, 1965; TRB, 1985; TRB, 1994; TRB, 1997).

Two commonly used queuing-related measures are queue length and lane occupancy. Queue length and duration can be determined by direct observation. Lane occupancy (or percentage of time a traffic lane is occupied by traffic) can be measured from vehicle detectors that are part of roadway surveillance and control systems. Queuing-related measures are increasingly being used to quantify roadway congestion because of the increasing availability of vehicle detectors and other sensors. However, although queues best reflect the public's perception of congestion, measuring queues remain laborious, site-specific, and time-specific. Because it is usually impractical to measure queues on a broader spatial scale, queuing-related measures tend to be inappropriate for planning and policy-related analyses.

Three commonly used travel time-based measures are travel time, travel speed, and delay. Travel time data collection is an integral component of traffic engineering studies. Travel speed and delay data can be derived from travel time data by using a reference desired/acceptable travel time or speed. Travel time-based measures are easy to understand by both the professional transportation community and the traveling public. They are flexible enough to describe traffic conditions at various levels of resolution in both space and time. This makes travel time-based measures appropriate for handling specific locations as well as entire corridors. It also allows analysts to perform comparisons over long periods of time, e.g., years or decades. Travel time-based measures translate easily into other measures like user costs, and can be used directly to validate planning models such as travel demand forecasting models (Laird, 1996). Travel time-based measures are applicable across modes. So important is travel time in this regard that the year 2000 edition of the HCM is being structured around travel time as a common measure of effectiveness for all modes (JHK, 1996). All these reasons make travel time-based measures extremely powerful, versatile, and desirable. Not surprisingly, an increasing number of transportation agencies are switching to travel time measures to monitor and manage congestion.

Unfortunately, budgetary limitations usually impose severe restrictions on the number and coverage of travel time studies. Recent technological advances, however, are assisting transportation officials in providing the necessary tools to make travel time data collection more affordable and reliable. Examples of those technological advances include GPS, GIS and a variety of subsystems and components normally associated with intelligent transportation system (ITS) deployments.

A number of travel time-based measures could be used (Lomax et al., 1997; Quiroga and Bullock, 1999). A sample of measures follows.

2.1. Travel time

Travel time (t_L) is the total time required to traverse a roadway segment of length L. It can be measured directly using field studies, although it could also be derived using simulation models or empirical relationships between volume and roadway characteristics.

2.2. Acceptable travel time

Acceptable travel time (t_{L_0}) is the travel time associated with a performance goal established for the transportation facility. The acceptable travel time should be influenced by community input and should, explicitly or implicitly, provide a balance between transportation quality, economic activity, land use patterns, environmental issues, and political concerns. In the absence of a more detailed analysis, a number of transportation agencies define the acceptable travel time as that associated with free flow conditions (or sometimes posted speed limits). A more detailed analysis should provide a differentiation by time period (i.e., peak vs off-peak), by functional class (e.g., freeway vs major arterial), and by geographic location (e.g., central business district vs suburban area) (Lomax et al, 1997). For example, based on a consensus involving technical and non-technical groups, an acceptable travel time per km (or travel rate, as defined below) during peak periods for a major arterial located in the central business could be, say, 3 min. For the same arterial, an acceptable travel time per km during off-peak periods could be, say, 2 min.

2.3. Segment speed

Segment speed (u) is the result of dividing the length (L) of the segment along which travel time data are collected by the corresponding travel time t_L:

$$u = L/t_L. \tag{1}$$

Length is an established item in a roadway inventory database and is normally measured in the field with a distance measuring instrument (DMI). GIS packages can also be used to provide estimates of distance, but it is clear that the accuracy of these estimates depends on the accuracy of the underlying digital base map. Digital base map accuracy has dramatically increased in recent years (submeter accuracy is now commonplace) and, as a result, measuring distances with GIS packages has become a feasible alternative. Readers should realize, however, that GIS packages usually measure distances on the ellipsoid, i.e., they "simulate" ground distances by using a mathematical model of the surface of the Earth. Some GIS packages take into consideration the eccentricity of the ellipsoid (i.e., explicitly account for the fact the Earth is not a perfect sphere), while other GIS packages simply assume a spherical model of the surface of the Earth. GIS packages that do consider the eccentricity effect provide more accurate estimates of distance than GIS packages that assume a simple spherical model of the surface of the Earth. Differences between the two models can be quite significant. For example, at 30° of latitude, differences between the two models could be up to 3 m/km on the N–S direction and up to 2 m/km on the E–W direction (Quiroga, 1999).

In general, travel time studies involve several runs and more than one segment. In this case, it may also be of interest to compute representative speeds and travel times for each segment and/or for all segments combined. In general, as shown in Fig. 2, if there are interchanges or intersections along the route, the number of runs per segment may be different.

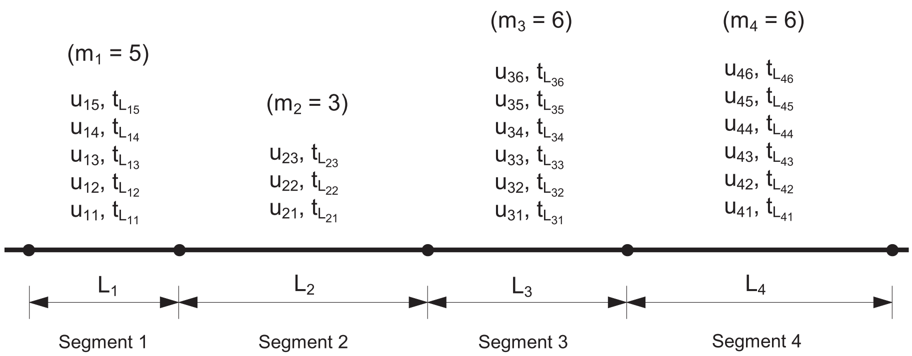

Fig. 2. Sample segment speeds and travel times.

Assume a representative segment travel time is given by the arithmetic average of all travel time values associated with a segment. Following Quiroga and Bullock (1999), the total representative travel time and speed over all segments can be expressed as

$$t_{T_L} = \sum_{i=1}^{n} \bar{t}_i = \sum_{i=1}^{n} \left[L_i \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{1}{u_{ij}} \right], \tag{2}$$

$$\bar{u}_L = \frac{L_T}{t_{T_L}} = \frac{\sum_{i=1}^{n} L_i}{\sum_{i=1}^{n} \bar{t}_i} = \frac{1}{\sum_{i=1}^{n} \left[\frac{L_i}{L_T} \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{1}{u_{ij}} \right]}, \tag{3}$$

where t_{T_L} is the total representative travel time over all segments, n the number of contiguous segments, $\bar{t}_i$ the average (arithmetic mean) of all travel time values associated with segment i, L_i the length of segment i, m_i the number of travel time and speed records per segment, u_{ij} the jth speed record associated with segment i, L_T the total length considered, and $\bar{u}_L$ is the overall average speed.

Eq. (3) represents a "weighted" harmonic mean of segment speeds, where the weight is the ratio of the length of each segment to the total length considered. One disadvantage of this equation is its sensitivity to outlying low speeds (which tend to occur on atypically adverse traffic conditions), resulting sometimes in very small average speeds. A more solid, robust estimator of central tendency can be obtained by using median segment travel times instead of arithmetic mean segment travel times. The median speed formulation is

$$\bar{u}_L = \frac{L_T}{\sum_{i=1}^{n} t_{m_i}} = \frac{1}{\sum_{i=1}^{n} \left[\frac{L_i}{L_T} \frac{1}{u_{m_i}} \right]}, \tag{4}$$

where t_{m_i} is the median travel time associated with segment i and, u_{m_i} is the median speed associated with segment i.

2.4. Travel rate

Travel rate (t_r) is the inverse of average speed and is usually expressed in min/km (or min/mile). While not readily understood by all audiences, travel rate provides a useful measure that can be averaged for a facility, geographic area, or mode. It can also be used to compare performance among transportation facilities more effectively than speed. Travel rate can be expressed as

$$t_r = t_L / L. \tag{5}$$

2.5. Delay

Delay (d_L) is the difference between travel time and the acceptable travel time on a road segment. Delay can be expressed as

$$d_L = t_L - t_{L_0}. \tag{6}$$

2.6. Total delay

Total delay ($D_{L_{\mathrm{tp}}}$) is the sum of delays for all vehicles traversing the segment during the time period for which travel time data are available and is normally expressed in vehicle-minutes or vehicle-hours. Total delay can be expressed as

$$D_{L_{\mathrm{tp}}} = V_{\mathrm{tp}} d_{L_{\mathrm{tp}}}, \tag{7}$$

where tp is the subscript associated with the time period for which data are available (e.g., 15 min), V_{tp} is the number of vehicles traversing the segment during the time period for which travel time data are available. Traffic volumes are established items in a roadway inventory database.

2.7. Delay rate

Delay rate ($d_{r_{\mathrm{L}}}$) is the rate of time loss for a specified roadway segment. It is calculated as the difference between the travel rate and the acceptable travel rate. Delay rate can be expressed as

$$d_{r_{\mathrm{L}}} = \frac{t_L - t_{L_0}}{L} = \frac{d_L}{L}. \tag{8}$$

2.8. Relative delay rate

Relative delay rate ($d_{r_{\mathrm{R}}}$) is a dimensionless index that can be used to compare congestion on facilities, modes, or systems in relation to different mobility standards. It is calculated as the delay rate divided by the acceptable travel rate. Relative delay rate can be expressed as

$$d_{r_{\mathrm{R}}} = \frac{t_L - t_{L_0}}{t_{L_0}} = \frac{d_L}{t_{L_0}}. \tag{9}$$

3. Data collection

There are essentially two groups of travel time data collection techniques: roadside techniques and vehicle techniques. Roadside techniques are based on the use of detecting devices physically located along the study routes at pre-specified intervals. They obtain travel time data from vehicles traversing the route by recording passing times at predefined checkpoints. Examples of these techniques include license plate matching and automatic vehicle identification (AVI). License plate matching is based on recording of the license plate number of individual vehicles and the corresponding time stamps as they pass checkpoints. Travel times are determined as differences in time stamps between checkpoints. An assumption of this technique is that each individual vehicle does not make intermediate stops. This may be limiting, particularly if there are intersections, on-ramps, off-ramps, or interchanges between checkpoints.

AVI is an example of a data collection technique included as part of traffic surveillance and control system deployments at traffic management centers. AVI systems are based on the used of in-vehicle transponders (or tags), roadside reading units, a communication network, and a central computer system. The roadside reading units detect individual vehicles equipped with transponders as they pass nearby and transmit the corresponding transponder data to the central computer system. Travel times between consecutive checkpoints are computed in a similar manner as with the license plate technique, except that transponder identification numbers are used to compare time stamps instead of vehicle license plate numbers. One advantage of AVI technology is that area-wide real-time travel time data collection and dissemination are possible. With GIS-based Internet tools, for example, cities like Houston, Chicago, and Seattle are using AVI technology to disseminate up-to-date geo-referenced travel time and speed data to the traveling public.

Vehicle techniques are based on the use of detection devices carried inside the vehicle. Examples of these techniques include the traditional stopwatch and clipboard technique and automatic vehicle location (AVL). In the stopwatch and clipboard technique, travel time and passage of specific landmarks are manually recorded along the route. Two technicians are required in the vehicle: one of them to drive and the other one to manually record items such as the location and time of individual checkpoints and the length and time spent in queues. Unfortunately, this process tends to be labor intensive during the data collection and data reduction phases, and spatial resolution and coverage are limited. In addition, problems such as missing checkpoints or inaccurately marked checkpoints are common. To avoid some of these problems, many transportation agencies use distance measuring instruments (DMIs) in their probe vehicles. With a DMI, only one technician is needed in the vehicle. In some cases, it is even possible to log route and checkpoint locations. However, DMIs require frequent calibrations to avoid inaccurate speed and distance readings (Benz and Ogden, 1996). In addition, DMI data (which by definition are linearly referenced) are not always compatible with geographic databases because of the difficulty to ensure those critical checkpoints on the survey, mainly the beginning and ending points, have been properly geo-referenced in the field.

AVL is a generic term that groups several techniques that use receivers or transmitters on board to determine vehicle location (in latitude and longitude) and speed. Examples of these techniques are ground-based radio navigational system techniques and GPS techniques. GPS techniques are particularly advantageous because they do not need receiving towers on the ground as traditional radio navigational systems do. Several GPS-based techniques have been developed in recent years (Guo and Poling, 1995; Laird, 1996; Quiroga and Bullock, 1996; Quiroga and Bullock, 1999; Zito et al., 1995) and the number of implementations in urban areas is constantly increasing.

One of the significant advantages of AVL techniques compared to other techniques is that traffic monitoring is roadway network and driver independent. This makes AVL suitable for many applications, including tracking the motion of special-purpose probe vehicles and entire fleets. When used with single probe vehicles, AVL systems are usually configured so that data are collected and stored on board, and then post-processed in the office. When used with entire fleets, AVL systems are usually configured so that data are collected and transmitted via radio or cellular phone to a central location where they can be automatically processed.

Table 1 is a summary of characteristics and applicability of the travel time data collection techniques described previously. Roadside techniques are obviously infrastructure dependent, as opposed to vehicle techniques. Roadside techniques have lower levels of resolution and accuracy than vehicle techniques. However, vehicle techniques are generally based on a limited number of probe vehicles, which means that area wide coverage is limited. This makes roadside technique (specifically AVI) better suited for daily or real-time monitoring. In contrast, vehicle techniques are best for determining initial conditions and for annual monitoring.

Table 1
Comparison of travel time data collection techniques (adapted from Liu and Haines, 1996; Turner, 1996)

Criteria	Roadside techniques		Vehicle techniques	
	License plate matching	AVI	DMI	AVL
Characteristics				
Infrastructure dependent	Yes	Yes	No	No
Travel time/speed resolution	Low	Low	High	High
Travel time/speed accuracy	Good	Good	Good	Very good
Area wide coverage	Low	Very good	Low	Low
Technology status	Proven	Proven	Proven	Proven[a]
Capital costs	Low	High	Low	Low to moderate
Operating costs per unit	Moderate	Low	High	Low to moderate
Applicability				
Annual monitoring	Yes	Yes	Yes	Yes
Daily monitoring	Limited	Yes	Limited	Limited
Real-time travel information	Limited	Yes	No	Yes
Incident detection	Limited	Yes	Limited	Limited

[a] GPS is a proven technology. However, its applicability to travel time studies has been limited until recently.

4. Data management

Regardless of the data collection technique used to collect travel time data, the data management component of a CMS is critical. In general, that component should be built using a geographic relational database model and provide all the necessary interfaces and procedures to provide the capability to measure congestion accurately, reliably, and efficiently. Obviously, the structure of the data management component depends on the data collection techniques used and the performance measures chosen.

As an illustration, this section summarizes the data management component of a travel time data collection application developed recently in support of a congestion management system for Baton Rouge, Louisiana (Quiroga and Bullock, 1998; Quiroga and Bullock, 1999). This application, called travel time with GPS (TTG), is based on the use of GPS receivers to collect travel time data and GIS-based procedures to manage the data collected in the field. TTG includes a spatial model, a geographic relational database, a procedure for linearly referencing GPS data using dynamic segmentation tools, and data reporting procedures. TTG was used to process 2.4 million GPS records on 40,000 km (25,000 miles) of travel time runs on a 240-km (151 miles) highway network.

4.1. Spatial model

The spatial model is based on highway links, where highway links are defined as directional centerlines delimited by physical discontinuities like signalized intersections, ramps, and interchanges (Fig. 3). Fig. 3 also shows sample GPS data being mapped to highway links. Linearly referencing these GPS points involves computing cumulative linear distances for GPS points located along the route of interest.

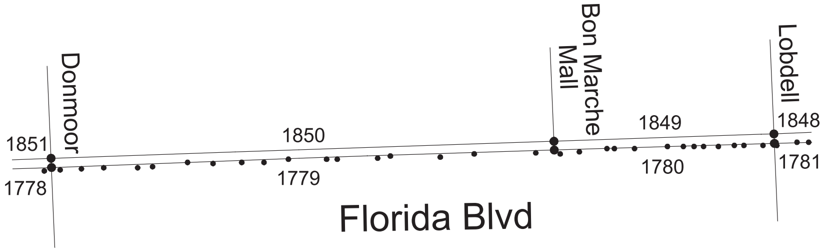

Fig. 3. Highway links on Florida Boulevard in Baton Rouge, Louisiana, with overlaid GPS data collected during a travel time run.

Fig. 4 shows a generic time–distance diagram for a probe vehicle that traverses a segment of length L (Note: a segment does not necessarily have to be same spatial entity as a link). The segment travel time, t_L, could be estimated by interpolating 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. If instantaneous speed values are recorded along with the GPS positional data, a representative speed value for the segment could

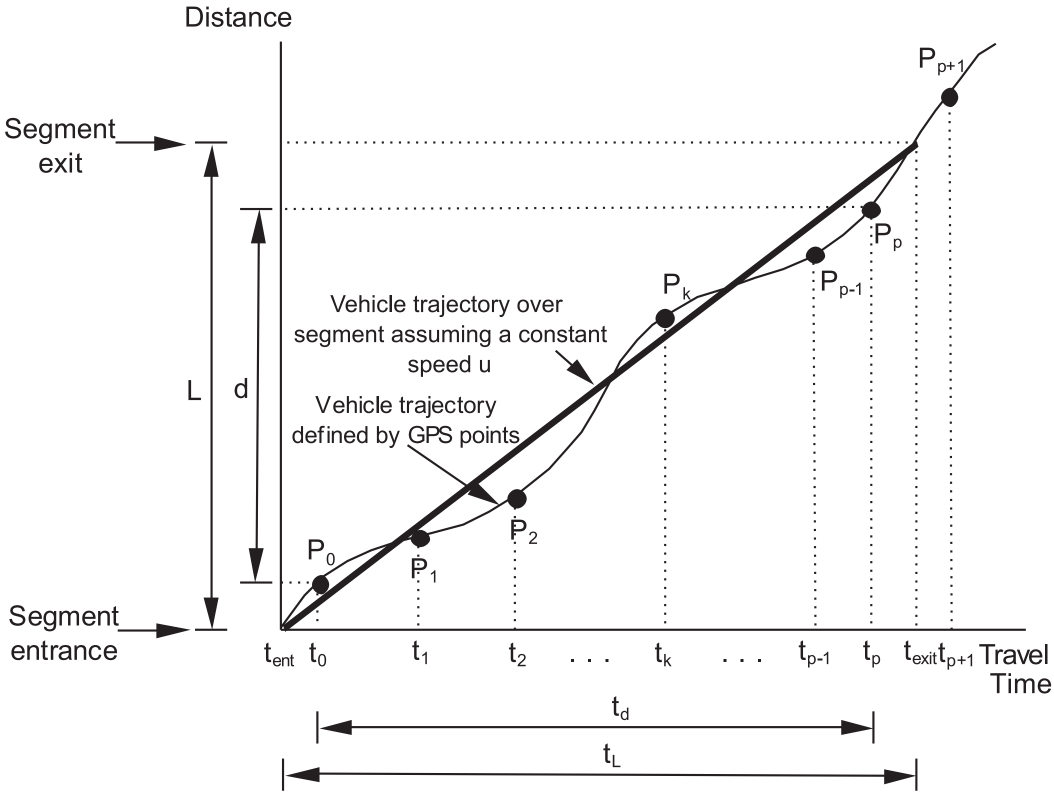

Fig. 4. Time–distance diagram for GPS points on a segment.

be computed. In this case, the segment travel time, t_L, could be estimated by dividing the segment length by the representative segment speed value.

4.2. Geographic database

Each GPS data file contains data such as time stamps, speed, latitude, longitude, and satellite navigational data at regular time intervals, say every 1 s. These data need to be linearly referenced so that GPS data can be associated with routes on the highway network. To manage all this information efficiently, a geographic relational database was developed. For illustration purposes, Fig. 5 shows the database schema (or database structure) and includes tables, field names and relationships (both one-to-one and one-to-many) among tables. To assist readers in the process of understanding the database structure, Fig. 6 shows a few sample records of each table included in the database.

The database structure 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 (e.g., Link code = 1779 in Fig. 3 and table GPS_DATA of Figs. 5 and 6). This way, users can easily build queries based on the same segmentation scheme as that used for generating the highway network map. Strictly speaking, however, all that is required from the linear referencing process is a route

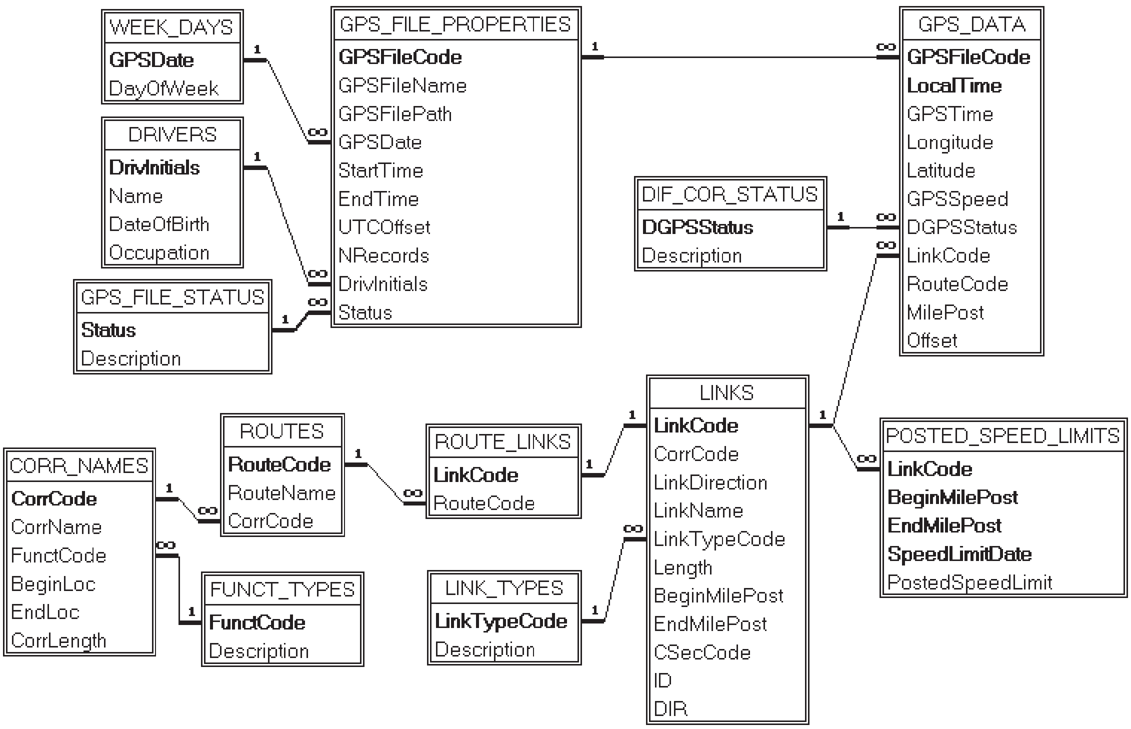

Fig. 5. Geographic database schema showing tables and relationships among tables (primary keys are indicated in bold).

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

Table LINK_TYPES

LinkTypeCode	Description
1	main
2	interchange

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

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
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

Fig. 6. Sample of records from the database (primary keys are indicated in bold).

code and a cumulative distance value for each GPS point (attributes RouteCode and MilePost in table GPS_DATA). With this information and any table containing cumulative distances associated with links or segments along routes, generic GPS data tables for any highway segmentation scheme can be produced to generate segment aggregated travel time and speed data.

TransCAD was used to linearly reference GPS data and to display results in a map format. Like other desktop GIS packages, TransCAD's architecture is based on a fairly large number of associated files. For example, a geographic file can easily involve 10–20 associated files including graphical elements, tables, indexes, and data dictionaries. Most tables are in dBase IV format (.dbf extension), which means that each table is stored in a separate dBase file. This kind of architecture is intended to provide flexibility to typical TransCAD users. However, it can also complicate data management problems in an environment where tens or hundreds of GPS data files are being generated and processed. With GPS data scattered in several dBase files, the process of enforcing data integrity constraints, building queries, and producing reports that involve aggregating or summarizing travel time and speed data by, say, time period or corridor, could be quite challenging. To address this issue, all database tables were stored 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

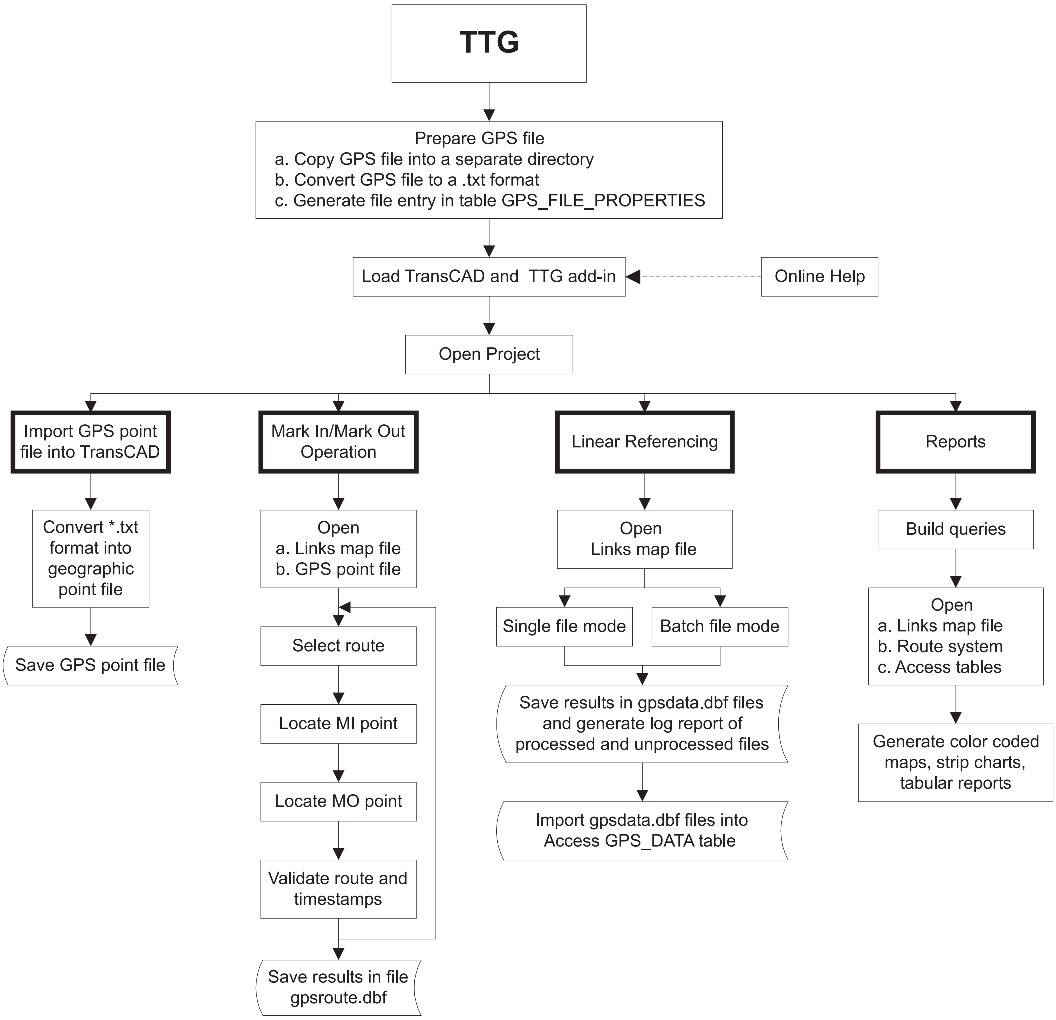

Fig. 7. Data reduction work flow.

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 that results from the linear referencing process in TransCAD.

4.3. Data reduction procedure

Fig. 7 shows a generic view of a typical work flow using TTG. For completeness, Fig. 7 shows both data reduction steps and data reporting steps. In this section, only the data reduction steps are discussed. In summary, the data reduction procedure allows users to

- Import GPS data file into the GIS to generate a point geographic file that can be overlaid on the highway network vector map;
- Specify when and where the vehicle enters and exits a study route using an animated GPS playback utility (Fig. 8). Of particular interest in Fig. 8 are the mark-in (MI) and mark-out (MO) buttons. The MI button is used to specify when the vehicle enters a route by marking the first GPS point associated with that route. Similarly, the MO button is used to specify when the vehicle exits a route by marking the last GPS point associated with that route. For added flexibility, the GPS player utility allows users to define MI–MO pairs anywhere along a route and define more than one MI–MO pair per route. This technique is useful for filtering out spurious GPS points and for partial route analysis;

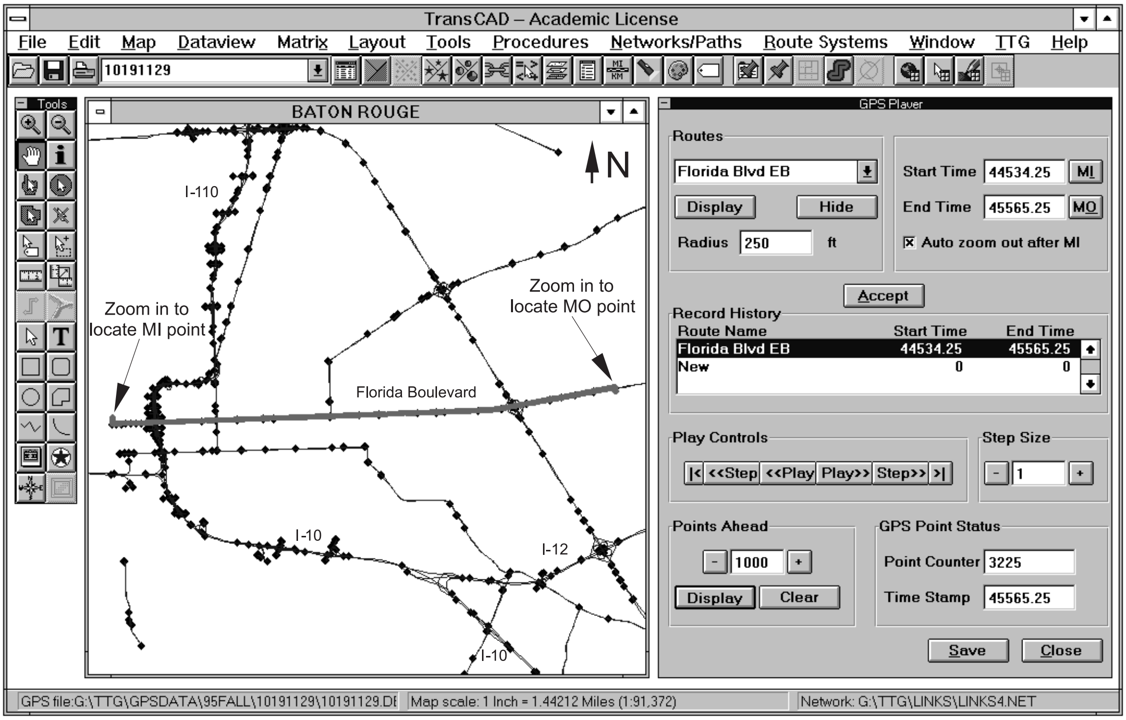

Fig. 8. Links map, GPS point file, and TTG GPS player utility.

- Linearly reference GPS points to the highway network. TTG also measures the transversal distance from each GPS point to the mapped link on the network (Offset field in table GPS_DATA of Fig. 6). This offset provides a verification of GPS positional accuracy and allows users to flag GPS points that may have been referenced to incorrect routes;
- Import the linearly referenced GPS data into a repository database (Microsoft Access).

4.4. Data reporting procedure

After storing the linearly referenced GPS data in a repository database, the next step involves constructing queries and reports. For the sake of brevity, only a sample of databases querying and data reporting options are included here. For additional data reporting examples, readers are referred to other sources (Quiroga, 1997; Quiroga and Bullock, 1998; Turner et al., 1998). Suppose it is of interest to produce reports showing average link speeds on a corridor. The procedure to do this can be summarized as follows:

- Build a query to retrieve the GPS records associated with the time period of interest (e.g., 7:00–7:15 am).
- Calculate link travel time and speed for each travel time run conducted during the time period of interest. Because relational database and GIS packages do not have tools to readily perform numerical interpolations, a special purpose utility was developed. This utility automatically calculates link speeds based on a table such as GPS_DATA and outputs the results to a table called SEGMENT_SPEEDS_x (where x represents the number associated with the procedure to calculate link or segment speed chosen). For example, by selecting a time interpolation procedure (Fig. 4), the output table would be called SEGMENT_SPEEDS_1. The utility provides users with the capability to compute segment speeds and travel times for any highway segmentation scheme.

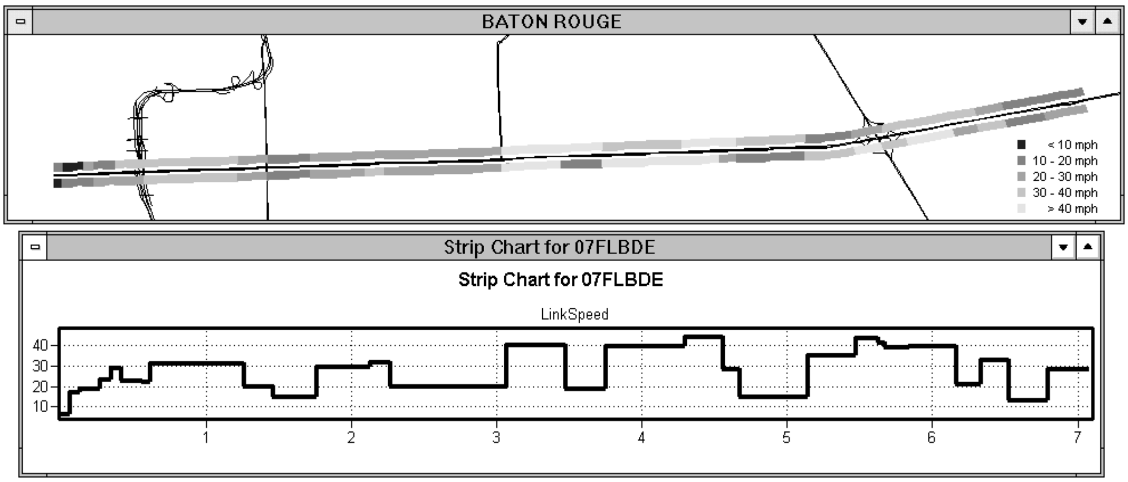

Fig. 9. Link speeds on Florida Boulevard in Baton Rouge, Louisiana.

- Calculate representative segment speeds (Eqs. (2)–(4)). Off-the-shelf database functions give users the capability to compute minimum, average, and maximum speeds, but not median speeds. As a result, a utility to compute median speeds was developed.
- Produce reports documenting average speeds on the corridor of interest. Examples of reports include maps and strip charts (Fig. 9) and tabular reports.

4.5. Discussion of results

TTG was tested using GPS data from 428 files collected in Baton Rouge, Louisiana. The 428 files included 2.4 million GPS records on 40,000 km (25,000 miles) of travel time runs on a 240-km (151 miles) 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 the other parts of the network including service roads, on-ramps, off-ramps, and intersecting streets. The 1.8 million GPS records located on the main routes were stored in table GPS_DATA in the access database file. The resulting size of this database file, including the other tables shown in Figs. 5 and 6, 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 provided the capability to improve and optimize the data reduction application, develop generalized data quality control checks, and detect limitations and bugs of the developing platform. In general, the centralized database 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, it appears the extra effort is worth the benefits the system provides, particularly with respect to the ability to build comprehensive queries to generate travel time reports, the ability to conduct generalized data quality control checks, and the ability to process data much faster than with manual data collection methods. An analysis of data reduction speeds indicates that the automated data reduction process can result in 15–17 min of data reduction time for two hours worth of GPS data. By comparison, traditional data reduction procedures based on manual data collection procedures require about one hour of data reduction per hour of data collection (Turner et al., 1998). In other words, the GPS/GIS data reduction procedure described here is about 7.5 times faster than traditional manual data collection procedures.

During the processing of the GPS data files, a number of areas were detected where errors tend to occur frequently. To assist readers in this process, a set of procedures or checks for data quality control were developed. Some of these checks are listed below. Additional checks can be found in Quiroga and Bullock (1998).

- *File system*. Make sure each GPS data file is stored in a separate subdirectory under the GPS-DATA subdirectory. Likewise, make sure the file entry in table GPS_FILE_PROPERTIES, particularly the file name and path is correct. The need to check for the location of the GPS data file could be eliminated by developing a script to automatically store GPS data files in the appropriate subdirectories as they are being uploaded from the data collection equipment. The script would also generate an entry in the database automatically.
- *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 the output file from the

data reduction process. 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 correct, i.e., that routes contain only valid links and that the beginning and ending mileposts of individual links are correct. TransCAD automatically calculate mileposts, however, if links must be edited, the GIS does not always recalculate distances correctly.
- *Linear referencing*. Verify that link codes, mileposts, and offsets are correct and meaningful. Offsets are transverse distances from the GPS points to the highway links and provide an indirect measurement of either GPS data positional errors (if the underlying highway network base map is more accurate than the GPS data) or GPS data mapping errors (unusually high offsets could be an indication that the route assignment was incorrect). Table GPS_DATA can be used to check for large offset values.
- *Average link/segment speeds*. Verify that representative link/segment speeds are meaningful. TTG includes two formulations for the computation of representative link/segment speeds: using a harmonic mean formulation (Eq. (3)), and using a median formulation (Eq. (4)). 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 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.

5. Conclusions

This paper presented a summary of procedures, collectively called travel time with GPS (TTG), and examples of application of GPS and GIS technologies for the collection of travel time data needed for monitoring and managing congestion. The paper examined transportation network performance measures and discussed the benefit of using travel time as a robust, easy to understand performance measure. The paper addressed data needs and examined the use of new technologies such as global positioning system (GPS) technology for the collection of travel time and speed data.

TTG is built using 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 directional vector representation of the 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 data. The procedure for linearly reference GPS data uses GIS dynamic segmentation tools. To automate this process, 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 was developed.

TTG was implemented using a PC-based TransCAD-Access environment. This environment is relatively inexpensive and allows users to process vast amounts of GPS data and produce reports quickly and cost-effectively. This environment works well in most cases although it was found that the software had some deficiencies that could produce erroneous results if care is not taken in processing the data.

Following an analysis of the travel time and speed data and an evaluation of the congestion situation, the next step would to identify and evaluate potential TCMs. If appropriate and properly implemented, TCMs could result in reductions of congestion levels. A sample of TCMs is included in Table 2. Notice that some of the TCMs can be evaluated using GIS techniques, particularly those that involve spatial analyses such as accessibility analysis, routing analysis, and demand/market program analysis. As an illustration, consider the case of defining appropriate locations for park-and-ride lots. Each park-and-ride location has an associated cost to locate, build, and operate. In addition, each location is expected to serve a number of drivers. For each driver using the park-and-ride lot, there is an associated cost of service, e.g., the travel time between the driver's home and the park-and-ride location. A goal for the park-and-ride location could be to minimize the cost of service for all drivers. Cost of service values is stored in cost matrices. In a cost matrix, each row represents the location of each alternative park-and-ride lot and each column represents a driver. Each cell, therefore, represents the cost of service for a single location-driver combination. For the GIS analysis, the following layers of data are needed: a network map layer with travel time and/or speed data, a park-and-ride location layer, and a driver location layer. Using GIS routing and dynamic segmentation functions, it is possible to construct the corresponding cost matrix. The minimum total cost of service could be obtained by adding the cost of service for all drivers who are expected to use each park-and-ride location and by comparing the total cost of service among locations. The locations with the lowest total cost of service are then retained for further analysis.

Another example could be that of development of para-transit operation improvements. The objective would be to service all origins and destinations based on a specified fleet of vehicles, while minimizing the total travel time for their customers. Each vehicle has a fixed capacity and each destination has a demand given by the number of passengers that must be transported there. This is a typical routing problem that requires the construction of a vehicle routing matrix. The vehicle routing matrix contains the distance and travel time between each origin (i.e., each customer's home) and each destination and between every pair of origins. It is

Table 2
Transportation control measures likely to have a positive impact on congestion

Transportation control measure group	Examples
High occupancy vehicle (HOV) lanes	Entrance ramp priority, dedicated HOV lanes
Traffic flow improvements	Traffic signal optimization, incident management systems
Parking management	Preferential parking for HOVs, parking zoning regulations, park-and-ride facilities, shuttle services
Vehicle use restrictions	Route diversion, downtown vehicle restrictions, no-drive days, truck movement control
Special event and activity center programs	Remote parking with shuttle service, parking management
Improved public transit	Service expansion, operational improvements, rail expansion
Employer-based transportation management programs	Carpooling, transit, financial incentives to employees, telecommuting, flextime, compressed work weeks
Trip reduction ordinances	Special use permits, mandated ridersharing
Rideshare incentives	Commute management organizations, tax incentives
Bicycle and pedestrian programs	Bicycle routes and storage facilities, sidewalks and walkways

possible to build the vehicle routing matrix by using GIS routing functions applied to a network map layer, a customer home layer, and a destination layer. Once the vehicle routing matrix is developed, the system attempts to find efficient routes that service as many customers and destinations as possible while trying to minimize the total travel time. The output from the procedure is an itinerary for each vehicle summarizing the route and all stops and destinations associated with that route.

References

Benz, R.J., Ogden, M.A., 1996. Development and benefits of computer-aided travel time data collection. Transportation Research Record 1551, TRB, National Research Council, Washington, DC, pp. 1–7.

CAAA, 1990. Clean Air Act Amendment. US Code, Title 1, Section 103.

FHWA, 1998. Our Nation's Highways – Selected Facts and Figures. Publication No. FHWA-PL-98-015, US Department of Transportation, Washington, DC, p. 28.

Francois, M.I., Willis, A., 1995. Developing effective congestion management systems. Federal Highway Administration, Technical Report No. 8, p. 22.

Guo, P., Poling, A.D., 1995. Geographic information Systems/Global positioning systems design for network travel time study. Transportation Research Record 1497, TRB, National Research Council, Washington, DC, pp. 135–139.

HRB, 1950. Highway Capacity Manual; Practical Applications of Research. National Research Council, Washington, DC, p. 147.

HRB, 1965. Highway Capacity Manual, second ed. Special Report 87, National Research Council, Washington, DC, p. 397.

ISTEA, 1991. Intermodal Surface Transportation Efficiency Act. US Code, Title 23, Chapter 3, Section 303.

JHK, 1996. Performance Measures and Levels of Service in the Year 2000 Highway Capacity Manual. NCHRP Project 3-55(4), TRB, National Research Council, Washington, DC, p. 25.

Laird, D., 1996. Emerging Issues in the Use of GPS for Travel Time Data Collection. National Traffic Data Acquisition Conference, Albuquerque, NM, 6–9 May, 1, pp. 117–123.

Lindquist, E., 1999. Assessing effectiveness measures in the ISTEA management systems. Southwest Region University Transportation Center, Texas Transportation Institute, College Station, TX, p. 103.

Liu, T.K., Haines, M., 1996. Travel Time Data Collection Field Tests – Lessons Learned. Report FHWA A-PL-96-010, US Department of Transportation, p. 116.

Lomax, T., Turner, S., Shunk, G., Levinson, H.S., Pratt, R.H., Bay, P.N., Douglas, G.B., 1997. Quantifying congestion. Final Report, National Cooperative Highway Research Program, Transportation Research Board, p. 184.

Quiroga, C.A., 1997. An integrated GPS-GIS methodology for performing travel time studies. Ph.D. Dissertation, Louisiana State University, Baton Rouge, LA, p. 171.

Quiroga, C.A., 1999. Accuracy of linearly referenced data using GIS. Transportation Research Record 1660, TRB, National Research Council, Washington, DC, pp. 100–107.

Quiroga, C.A., Bullock, D., 1996. Architecture of a congestion management system for controlled-access facilities. Transportation Research Record 1551, TRB, National Research Council, Washington, DC, pp. 105–113.

Quiroga, C.A., Bullock, D., 1998. Development of CMS Monitoring Procedures. Report No. FHWA/LA-314, Louisiana Transportation Research Center, p. 87.

Quiroga, C.A., Bullock, D., 1999. Travel time information using GPS and dynamic segmentation techniques. Transportation Research Record 1660, TRB, National Research Council, Washington, DC, pp. 48–57.

Schrank, D.L., Lomax, T.J., 1999. The 1999 Annual Mobility Report. Texas Transportation Institute, p. 123.

Schwartz, W.L., Suhrbier, J.H., Gardner, B.J., 1995. Data collection and analysis methods to support congestion management systems. ASCE Transportation Congress, Proceedings V. 2, San Diego, CA, pp. 2012–2023.

TEA 21, 1998. Transportation Equity Act for the 21st Century. US Code, Title 23, Section 149.

TRB, 1985. Highway Capacity Manual, third ed. Special Report 209, National Research Council, Washington, DC, p. 504.

TRB, 1994. Highway Capacity Manual, third ed. Update, Special Report 209, National Research Council, Washington, DC.

TRB, 1997. Highway Capacity Manual, third ed. Update, Special Report 209, National Research Council, Washington, DC.

Turner, S.M., 1996. Advanced techniques for travel time data collection. Transportation Research Record 1551, TRB, National Research Council, Washington, DC, pp. 51–58.

Turner, S.M., Eisele, W.L., Benz, R.J., Holdener, D.J., 1998. Travel Time Data Collection Handbook. Report No. FHWA-PL-98-035, Texas Transportation Institute, College Station, TX, p. 346.

Zito, R., D'Este, G., Taylor, M.A.P., 1995. Global positioning systems in the time domain: how useful a tool for intelligent vehicle-highway systems? Transportation Research Part C 3 (4), 193–209.