Comparing data from mobile and static traffic sensors for travel time assessment

Nicolas Saunier and Catherine Morency

Department of civil, geological and mining engineering, École Polytechnique de Montréal, P.O. Box 6079, Station Centre-Ville, Montreal, PQ, CANADA, H3C 3A7; emails: nicolas.saunier@polymtl.ca, cmorency@polymtl.ca

ABSTRACT

Travel time and speed measures on road networks provide key information to identify critical spots of congestion and evaluate the scale of this phenomenon across an urban area. Many technologies are currently available to measure travel time and speed, but each has its limitations. As part of a wider project aiming to develop travel time reliability indicators, this paper presents a comparison and validation of data collected through four different non-intrusive techniques: floating cars, GPS traces, Bluetooth detections and video processing. First, some background information regarding the project and the modeling of travel times and speed on highway networks is provided. Then, a comparison between the various sources of data is presented. Finally, the characteristics of the various data sources are discussed based on the relevance and the availability of the sources.

INTRODUCTION

With the increasing problems of traffic congestion in urban areas, transportation planners need better tools to assess its evolution through the years. Travel time measures on road networks provide key information to identify critical spots of congestion and evaluate the scale of this phenomenon across the area. Many technologies are currently available to measure travel time at specific locations: license plate matching, loop detectors, Bluetooth (BT) device matching, and video processing. Traditionally, at a larger scale, travel time is estimated using mandated floating cars (FC) that run on specific routes to gather data on the road network. Nowadays, ad-hoc, or randomly collected, data can also be obtained from GPS devices aboard individual cars and commercial vehicles. In addition, interest for video data is growing since various computer vision techniques permit automated traffic monitoring and data collection. However, each of these data collection methods has limitations that must be addressed through the right aggregation method.

As part of a wider project aiming to develop travel time reliability indicators, this paper presents a comparison and validation of data collected through four different non-intrusive techniques: FC sample, GPS traces, BT detections and video-based traffic data. These sources are available for various locations of the Montreal highway network and time spans: the first two mobile sources, GPS and FC, provide traffic data on large spatial areas, but limited in time, while BT devices and video data may be collected continuously, but from static sensors at specific locations.

First, some background information on the research project and available sets of data is presented. Then, the analysis framework is presented along with a statistical description of the travel times and spot speed data extracted. The next section presents the possible comparisons between travel condition patterns identified in the various datasets. The conclusion and perspectives are finally provided.

BACKGROUND AND INFORMATION SYSTEM

Project: assessing the reliability of the highway network

In 2008, the Quebec Ministry of Transportation (MTQ) mandated Polytechnique to assess the potentialities of historical FC data to provide relevant estimates of the reliability of the Greater Montreal Area highway network. Six years of FC data were analyzed and used to model the evolution of travel times and variability of travel times over time. Also part of the project was the evaluation of the available technologies to monitor the evolution of travel times and/or speed on the main highway network. Experimentations are currently conducted to compare outputs of various data collection methods. The purpose of this comparison process is to formulate recommendations with respect to the value and relevance of available datasets, tools or technologies to provide critical information for the strategic planning of the network by decision makers.

Evaluating congestion on road networks

Better understanding the use of transportation network is a key factor for the enhancement of transportation planners and models. Hence, more and more studies are conducted to identify the various negative drawbacks of increasing congestion in urban areas. Issues related to sustainable development are actually providing new incentives to better understand and monitor congestion on transportation networks. For instance, studies on environmental impacts of congestion (Nesamani et al. 2005) as well as on the modification of activity rhythms of households and individuals, namely the reduction in shopping related trips (Schmöcker et al. 2006), are being conducted. Also, numerous research works are linked to the definition of sustainability and the various indicators to assess its level. Litman (2008) provides a very interesting view as well as an extensive list of indicators assisting the objective assessment of sustainability. Indicators such as commuting travel time or delays due to congestion are set as economical indicators in the evaluation framework.

Along with the increasing need to better understand congestion problems is the increasing availability of information technology that output multitudes of data on the movements of objects (people, vehicle, phones and other devices). This data can be processed to contribute to the measurement of critical indicators such as travel times or speeds. The emergence of services provided by companies (such as INRIX and Google) relying on distributed data collection and aggregation is an obvious witness of the spreading availability of multiple layers of transportation-related data and of the value of such information.

In the United States, the recurrent reports by Schrank and Lomax (2003, 2005, 2007) illustrate the relevance of estimating comparable congestion-related indicators for the main areas of this country. In the same perspective, FHWA (2004) is pursuing a continuous effort to develop and estimate indicators through its Mobility Monitoring Program. This program relies on the data outputted from detectors available in some thirty areas. Another interesting contribution is from the

Pennsylvania DOT (Szekeres and Heckman 2005) that implemented a Congestion Management System at the state level.

Information system

Study area

The study relies on various datasets gathered for routes and highway segments of the Greater Montreal Area (GMA). The GMA accounts for more than half of the Québec population and is the second largest area in Canada, after the Toronto Area. The most congested and intensely used highway segment of the Province of Québec is located on the Montreal Island (Highway 40, between the two connections with Highway 15). Many urban highways of the region act as main through corridors and face increasing congestion level.

Floating cars

This research project first started with the systematic processing of FC data that were collected at regular intervals by the MTQ to assess travel times over the main highway corridors. From 1998 to 2004 (no data collection in 2003), 29,229 FC routes were collected (over 51 months). The data collection process was articulated around 55 different routes for a total of around 800 kilometers of road segments.

GPS traces

The GPS traces are made available by the Québec carsharing company (Communauto inc.) who owns in the GMA a fleet of 1000 cars from which 400 are equipped with GPS devices. The shared cars can be used by the 20,000 members of the area to either travel short or long distances. For the current project, all the data from the 2009 years was made available. More than 13 millions of data points were made available with an average sampling period of 5 minutes.

Video data

The Québec highways, especially in urban areas, are well covered with traffic cameras, for use in traffic management centers and for traveler information through their display on the MTQ website. In addition, the MTQ regularly carries out traffic studies that require video data collection, typically for manual data extraction or verification. This work demonstrates that such data can be analyzed to extract traffic data.

BT devices

MTQ has also collected travel times data from three experimental BT devices in operation for a 2 weeks period in the Greater Montréal Area. The experience was not as straightforward as expected since one of the device stopped working early in the process but they still manage to collect 5318 positive matches between the two remaining devices (Fournier 2009). Travel times data is derived from the matches.

METHODOLOGY AND STATISTICAL DESCRIPTION

Modeling the distribution of travel times using FC data

The first stage of the project was to propose a model of travel time distribution based solely on FC historical data. Since previous modeling attempts had failed to draw significant results, it was decided to transform the initial data source into travel time data over road segments of constant one-kilometer length. Each FC route was automatically cut into one-kilometer segments and travel times estimated for each.

The second stage of the modeling process was to create clusters of segments based on the similarity of frequency distribution of travel times. Various classifications were tested using a k-means algorithm and segments were finally assigned to one of eight clusters (two distinct sets of 8 clusters for two periods of the day) (Loustau et al., 2010 a and b). For each cluster, a modeled frequency distribution of travel times is derived using three additive log-normal laws. These models are used for comparison purposes in the following sections.

GPS traces to spot speed data

With their limited temporal resolution (one point every 4-5 min), the current GPS traces are not suitable to provide accurate tracking of the cars over the network. Hence, it was decided to extract spot speed data from every point by matching to road segments using proximity and heading similarity. Every GPS point is linked to the nearest segment using GIS proximity function and then extracted for further analysis if the two following criteria are respected:

-
-

where *GPS(t)* and *RS* stand respectively for the GPS point at time *t* and the nearest point on the nearest road segment, and *A()* is the angle of the vehicle heading or road segment with respect to a common direction, e.g. the north.

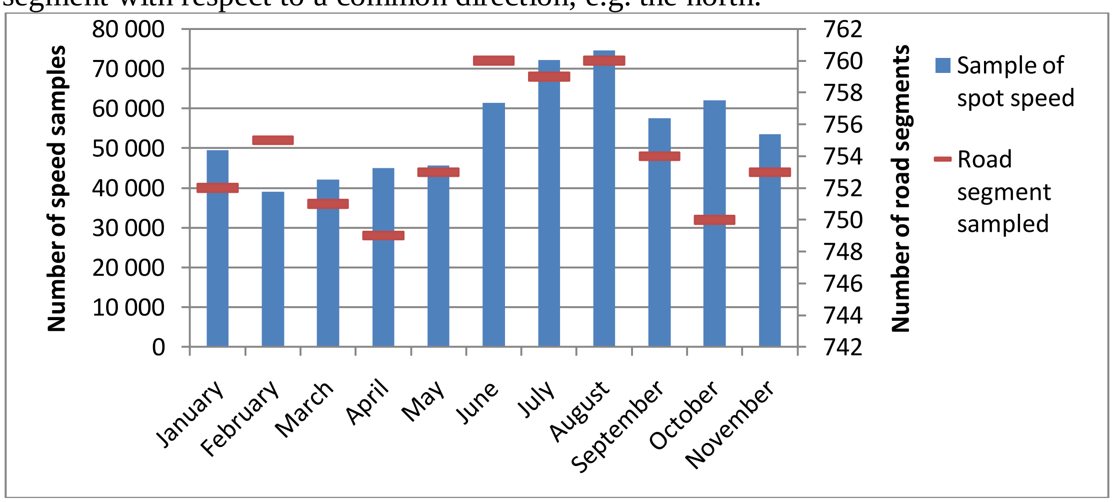

Figure 1. Description of the monthly sample available using GPS traces from shared cars.

Figure 1 summarizes the spot speed data extracted from the GPS traces over the 2009 year (11 months). The GPS traces provide between 39,000 and 74,000 spot speed data matched on the road segments that are typically surveyed using FC. The highest samples (both in terms of spot speed data and road segments with at least one observation) are observed during the summer months where the carsharing service has its highest level of usage.

BT detection to travel times

Because of the battery failing of one of the BT sensors, matches between the two stations can be obtained only for about 6 days, from September 17^{th} to 22^{nd} 2009. Data post-processing is necessary to derive the travel times (Haghani et al. 2009). The first detection at each station was used and very large travel times were discarded (more than 1000 s for the 1.555 km, or, given the mean travel time over 6 min and the corresponding standard error of the mean σ, travel times t such that). For aggregate results, the travel times were averaged per 6 min intervals. Travel times are then converted into speed data to allow comparison with spot speed data extracted from GPS traces, using the simplest hypothesis of constant speed between the two devices. Figure 2 presents the daily speed patterns obtained from the BT recognition process in one direction. We clearly see the differences in travel conditions between weekdays and week-end days as well as the anticipated peak period on Friday afternoons. Comparisons will hence be conducted separately for weekdays and weekends.

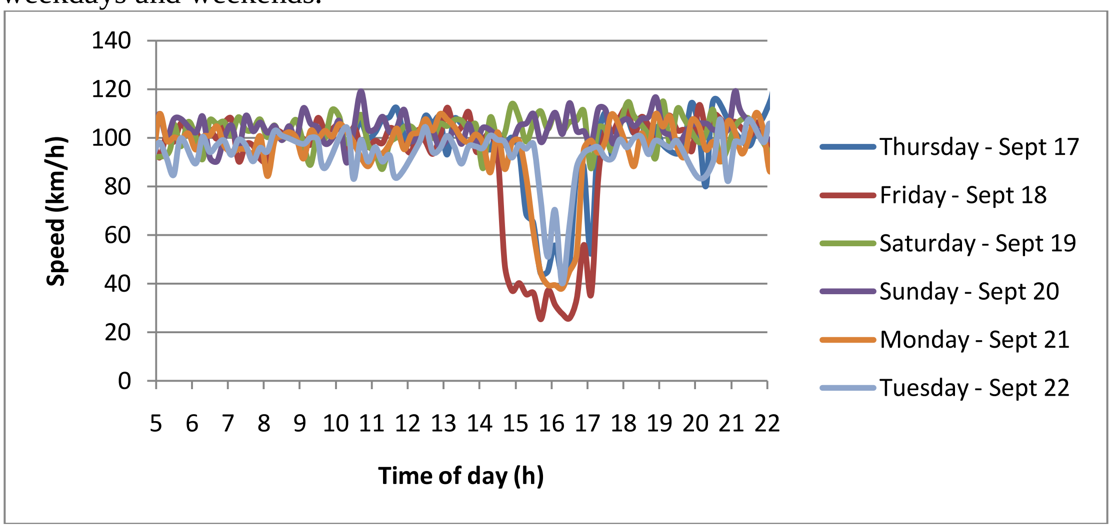

Figure 2. Pattern of mean speed extracted from 6 days of BT recognition data.

Video sequences to spot speed data

Although video cameras may cover large areas, a few tens to hundreds of meters, depending on their resolution and installation, these areas are limited with respect to the scale of the road network. Multiple cameras could cover long corridors, but are not available in this work. Spot speed, counts (flow) and density may be extracted for a specific location covered by a video camera.

Various computer vision techniques have been proposed to analyze movement in video data. Although extracting speed data does not require in itself to track all vehicles, this work makes use of a feature-based road user tracking system developed previously at the University of British Columbia for automated road safety analysis (Saunier et al. 2006, 2008 and 2010). Feature-based tracking is robust to partial occlusions and does not require any initialization step. One requirement is to be able to convert measurements done in the image space to measurements in real world coordinates, typically at the ground level. A robust camera calibration tool for urban traffic scenes developed previously was successfully used for this purpose (Ismail et al. 2010).

Once vehicle trajectories are extracted from video data, speeds are computed by a moving average of the vehicle displacements. At each location, two lines were identified between which vehicle speeds were averaged to yield a spot speed observation. Obvious outliers with very large mean speed, speed standard deviation and mean acceleration were discarded. The lane of each vehicle was roughly estimated based on their average lane positions; average speeds and 95 % confidence interval are computed for 5 min intervals. This process was applied to three locations and results for a busy tunnel entrance near the downtown area are shown in Figure 3. Spatial and temporal speed variations can be easily identified, especially the slowdowns on the exit of the highway and some spillover on other lanes. As can be expected, speeds are increasing from the rightmost to the leftmost lane.

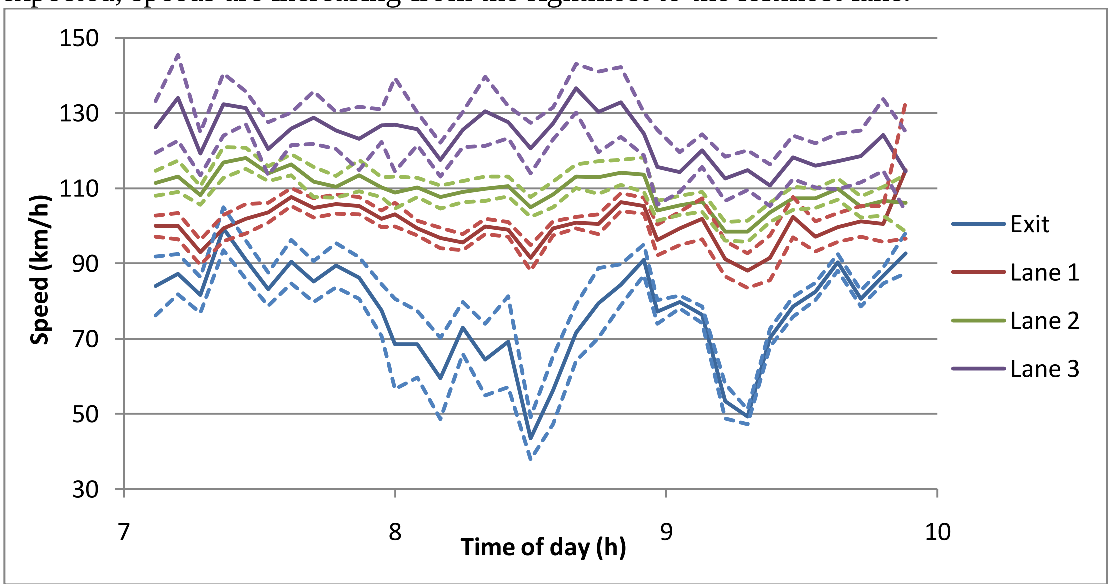

Figure 3. Average speeds and 95 % confidence interval (dotted lines) extracted from video collected at the entrance of Ville-Marie Tunnel in September 2009.

COMPARISONS: RESULTS AND ANALYSIS

Comparing GPS to FC travel times

Travel time estimations from FC data, based on one-kilometer segments, are compared with travel time estimation from GPS traces (a historical set of smaller size – 2006-2008 – as well as recent sets of 2009). Given the low temporal resolution of the GPS traces, spot speeds extracted from the traces are converted to travel time under the hypothesis of constant speed over 1 km. Figure 4 presents the comparison between travel times over highway 40 that namely contains the most congested segment of the Québec Province (between segment 22 and segment 37). All three sets of data manage to reveal the highly congested segments (all day periods pooled). Actually, the correlation between all series is very high:

- 0.833 between FC and historical GPS traces and from 0.800 to 0.852 between FC and monthly 2009 GPS data

- around 0.950 between historical and 2009 GPS estimates
- at least 0.970 between the 2009 patterns except for the July month for which correlation with fall or winter months is around 0.940

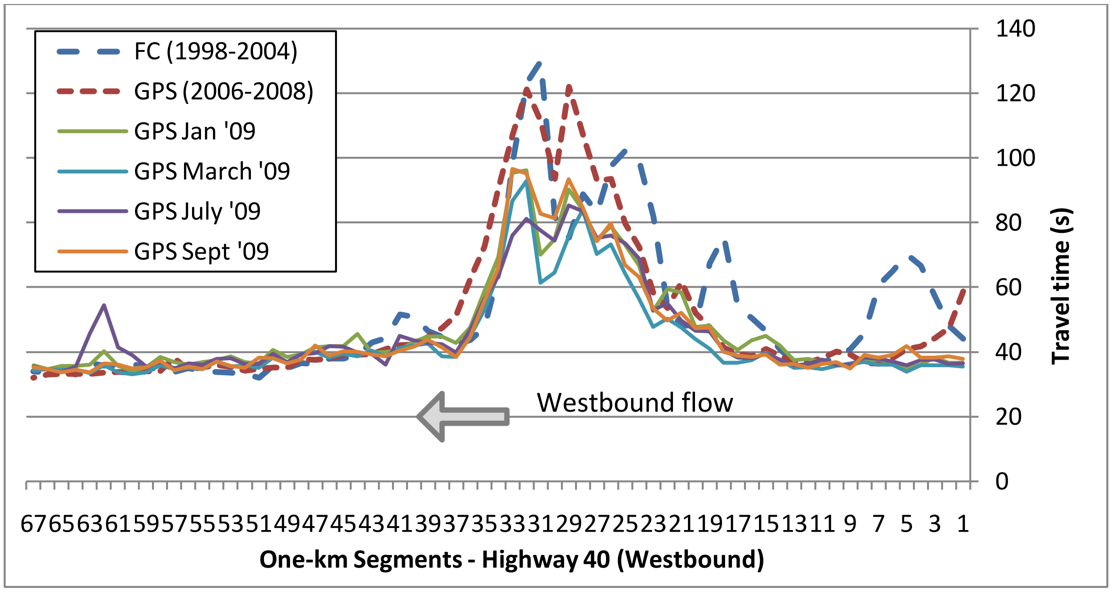

Figure 4. Average travel times estimation using the GPS and FC datasets for highway 40.

Comparing GPS to video-based spot speeds

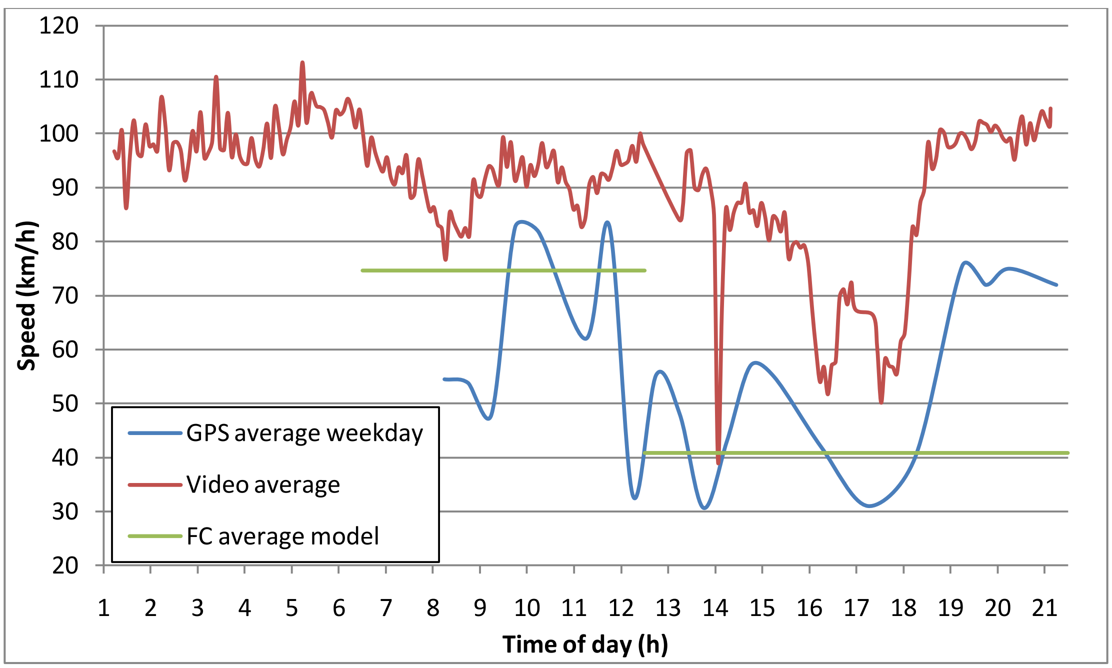

Figure 5. Video-based average speed, GPS speed and VF model average speed (respectively for morning and peak time periods) at the entrance of the Bridge-tunnel Louis-Hippolyte-Lafontaine.

Video data was collected at given locations for limited periods of time (all less than a day). The spot speeds extracted from the longest video recording (1h00 to 21h00 in November 2008) were compared to GPS traces (weekdays of October and November 2009) at the same location and a model estimated from FC data (1998 to 2004) (Figure 5). It is impossible to draw strong conclusions, but the trend is similar between the different sources of data: fairly fluid traffic conditions in the morning with some slowdowns around 8-9h00 and congestion during the afternoon peak hours, recovering after 18-19h00. The slowdowns detected in the video data were manually verified. The systematically higher speeds obtained from video data compared to the other sources may be caused by perspective.

Comparing GPS to BT devices

The experiment with BT data was unfortunately done in an area that was not covered by FC. Since only a limited number of GPS spot speed points were perfectly matched in space and time to BT recognition, GPS data from the fall period were processed to produce average patterns of speed over the studied segment (Figure 6).

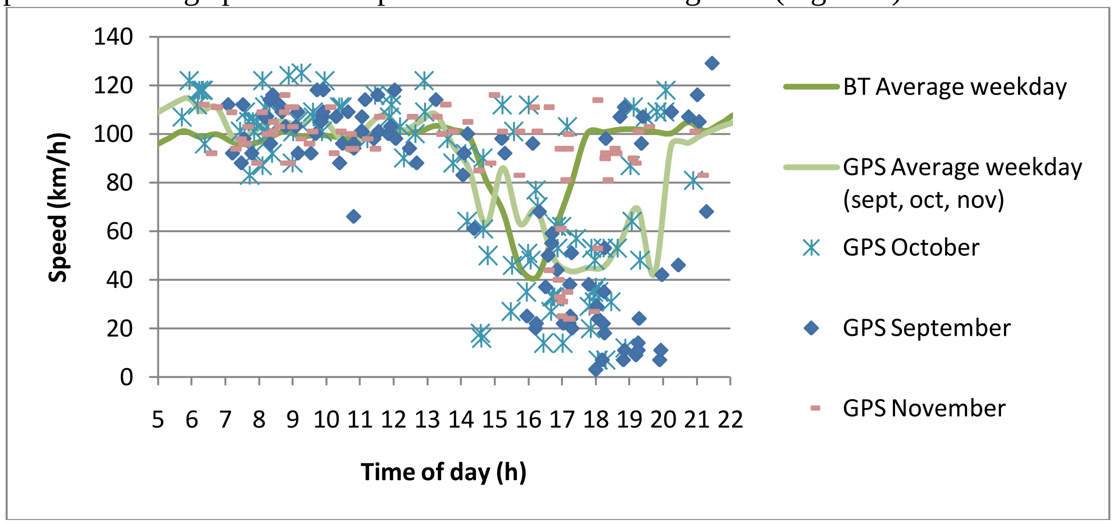

Figure 6. BT and GPS average weekday speeds near Sainte-Julie in September 2009.

Correlation between the two average speed patterns obtained with the two technologies is currently not convincing (Pearson r coefficient of 0.431): both technologies agree on non-congested conditions but the peak period is not similarly assessed. However, the three months of GPS data seem to provide similar weekday patterns of travel speed on the segment (r = 0.660 between September and October, r = 0.674 between September and November and r=0.623 between October and November). Uncertainties related to the recognition of the moving object by the BT device or measurement errors of the distance between the two devices could explain some observed differences.

CONCLUSION AND PERSPECTIVES

Assessing the travel conditions on the road network is a critical task for transportation planners and operators. On the one hand, they need to confirm whether conditions are getting better or worse and provide the relevant information to the users to ease their travels. On the other hand, they need to feed strategic planning with relevant travel time estimations to allow for comprehensive forecasts of future travel conditions, according to various interventions. There was a time when the lack of data limited the development of more precise and extended evaluation of travel conditions. Nowadays, data is becoming easily available via various sources and non-intrusive sensors but their coherent processing and integration is not simple. Every type of data can potentially contribute to increase the understanding of travel conditions but comes also with inherent limitations reflected in their mutual comparisons.

This paper acts as a first step into the integration of travel times and speed data for the GMA highway network. It has provided preliminary demonstration of the variability of the travel patterns observable on typical road segments. The characteristics of the data sources are summarized in Table 1. Large spatial coverage is useful to model typical travel conditions on various routes, but temporal coverage is required to inform on their variability and evolution. A recommendation is then to supplement costly or uncontrolled data collection through mobile sensors with continuous coverage at specific locations through static sensors: for example, FC data could be updated using GPS data and static sensors in key areas targeted for more complete monitoring. Future work will focus on the collection of larger datasets to allow better comparisons between the various data sources. The next step in this research will investigate the methods for the combination of the data sources.

Table 1. Comparison of the traffic data sources.

	Cost	Spatial coverage	Temporal coverage	Data types	Challenges
FC	high	moderate	small	speed, travel times	small sample size
GPS	low	large	large	speed, travel times	matching GPS positions to the network, varying sample size, no control
BT	low	small	continuous	travel times	ongoing development with some detection issues
Video	moderate	small	continuous	speed, counts, density, behavior	ongoing development, varying accuracy

ACKNOWLEDGMENTS

The authors wish to acknowledge the continuous contributions and support of the MTQ, namely the team of Pierre Tremblay, both in terms of financial support and data provision. They also wish to acknowledge the contributions of Communauto inc that provided access to datasets of GPS traces (Benoît Robert and Marco Viviani).

The authors wish to acknowledge the work of several students from École Polytechnique who assisted in the processing of various datasets: Ali El Husseini (video data), Vincent Grasset (GPS traces), Pierre Loustau (FC data) and Nakul Reddy Nalla (BT data).

REFERENCES

Fournier, P. (2009) Expérimentation des relevés de temps de parcours à l'aide de capteurs Bluetooth Ministère des transports du Québec.

Haghani, A., Hamedi, M. and Sadabadi, K. F. (2009) I-95 Corridor Coalition Vehicle Probe Project Validation of INRIX Data July-September 2008. I-95 Corridor Coalition. Final Report.
Ismail, K.; Sayed, T. and Saunier, N. (2010) Camera Calibration for Urban Traffic Scenes: Practical Issues and a Robust Approach, Presented at the 89th Annual Meeting of the Transportation Research Boarch (CD compendium of papers), Washington D.C.
Litman, T. (2008). Well Measured. Developing Indicators for Comprehensive and Sustainable Transport Planning, Report, Victoria Transport Policy Institute, 47 pages (http://www.vtpi.org/wellmeas.pdf, checked in May 2008).
Loustau, P., Morency, C., Trépanier, M. and L. Gourvil. (2010a) Travel Time Reliability on a Highway Network: Some Estimations using Floating Cars Data, Transportation Letters, Volume 2, Issue 1 / January 2010, J.Ross Publishing, pp.27-37.
Loustau, P., Morency, C. and Trépanier, M. (2010b) Measuring, Describing and Modeling Travel Time Reliability, Presented at the 89th Annual Meeting of the Transportation Research Boarch (CD compendium of papers), Washington D.C.
Nesamania, K.S., Chub, L., McNallyc, M.G., Jayakrishnan, R. (2005). Estimation of vehicular emissions by capturing traffic variations, Center for Activity Systems Analysis. Paper UCI-ITS-AS-WP-05-2. http://repositories.cdlib.org/itsirvine/casa/UCI-ITS-AS-WP-05-2
Saunier, N. and Sayed, T. (2006) A feature-based tracking algorithm for vehicles in intersections, Third Canadian Conference on Computer and Robot Vision, IEEE.
Saunier, N. and Sayed, T. (2008) A Probabilistic Framework for Automated Analysis of Exposure to Road Collisions, Transportation Research Record, 2083, pp. 96-104.
Saunier, N.; Sayed, T. and Ismail, K. (2010) Large Scale Automated Analysis of Vehicle Interactions and Collisions, Transportation Research Record, 2147:42-50, 2010.
Schmöcker, J.-D., Fonzone, A, Quddus, M., Bell, M.G.H. (2006). Changes in the frequency of shopping trips in response to a congestion charge, Transport Policy, Volume 13, pp. 217–228.
Schrank, D., Lomax, T. (2003). The 2003 Urban Mobility Report, Texas Transportation Institute, http://mobility.tamu.edu, May 2003.
Schrank, D., Lomax, T. (2005). The 2005 Urban Mobility Report, Texas Transportation Institute, http://mobility.tamu.edu, May 2005.
Schrank, D., Lomax, T. (2007). The 2007 Urban Mobility Report, Texas Transportation Institute, checked in May 2008), September 2007. http://tti.tamu.edu/documents/mobility_report_2007_wappx.pdf
Szekeres, D., Heckman, M. (2005). Statewide Congestion Measurement Study – Final Report, Pennsylvania Department of Transportation, Contract N° 030419, Juillet 2005.