I T L S

WORKING PAPER

ITLS-WP-08-01

Determining trip information using GPS data

By

Eoin Clifford, Jun Zhang and Peter Stopher

January 2008

ISSN 1832-570X

INSTITUTE of TRANSPORT and LOGISTICS STUDIES

The Australian Key Centre in
Transport and Logistics Management

The University of Sydney

Established under the Australian Research Council's Key Centre Program.

NUMBER: Working Paper ITLS-WP-08-01

TITLE: **Determining trip information using GPS data**

ABSTRACT: With the development of lightweight, high sensitivity Global Positioning System (GPS) devices, there has been increasing interest in their use as a means to measure people's travel for travel surveys. Data-logging enables a person's position, speed, and heading to be recorded on a second by second basis. GPS devices cannot, however, collect data on the mode or the purpose of travel, both of which are frequently required for transport planning purposes. This paper describes a set of heuristic rules developed by the authors for determining both the mode of travel and the purpose of trips recorded on GPS devices. This processing utilises comprehensive GIS databases for the areas where the GPS surveys have been conducted. These include information on all public transport routes in the region of interest. The rules developed consider the average, maximum and minimum speeds when determining mode of transport. They also use information about the transport network and the availability of bicycles and cars to the survey participants. Data-cleaning procedures are incorporated to eliminate erroneous information obtained from the devices themselves. To determine the purpose of a trip, it is necessary to have full land use records for all locations in the survey area. The addresses of all workplaces, schools, and most frequently used grocery stores visited by household members are also collected. Our procedure checks trip origins and destinations against these locations, and also the amount of time spent at the destination. We have applied our procedure to data obtained from a survey in Adelaide. The procedure gives results which are comparable to those obtained using a more burdensome travel diary.

KEY WORDS: ***Travel surveys, transport, GPS applications, trip analysis***

AUTHORS: Eoin Clifford, Jun Zhang and Peter Stopher

CONTACT: Institute of Transport and Logistics Studies **(C37)**
The Australian Key Centre in Transport Management
The University of Sydney NSW 2006 Australia
Telephone: +61 9351 0071
Facsimile: +61 9351 0088
E-mail: itlsinfo@itls.usyd.edu.au
Internet: http://www.itls.usyd.edu.au

DATE: January 2008

1. Introduction

Global Positioning System (GPS) technology has been used for examining various transport-related issues since the mid-1990s. GPS devices have been used for numerous transport applications, including objective measurement of infrastructure (Baffour, 2002), various measurements of traffic flow, system performance, and related phenomena (Quiroga et al, 2002), (Ranjitkar et al, 2002), (Wee et al, 2002), (Bullock and Jiang, 2003), validating the measurement of personal travel behaviour (Wolf et al, 2003), (Forrest and Pearson, 2005), (Li et al, 2004), (Stopher et al, 2002), and evaluating behaviour-change policies (Stopher et al, 2005), (Stopher et al, 2006). It has been suggested that GPS devices could take the place of more conventional household travel surveys (HTSs), which require participants to record there own movements on paper for a full day (Wolf et al, 2001). However this notion has not yet taken off, although GPS devices are sometimes used as a complement to HTSs.

Originally, GPS devices required an external power source, which limited them to being used in vehicles, and sometimes only in specially-equipped vehicles (Wagner, 1997), (Quiroga, 2000), (Wolf et al, 2000), (Wolf et al, 2003). Advents in GPS technology allowed the development of devices with an attached or built-in power source (Draijer et al, 2000), (Stopher et al, 2005), (Wolf, 2006). Further developments have produced small and lightweight personal devices that are very sensitive and able to pick up GPS signals in locations previously inaccessible, such as public transport vehicles, high-rise urban zones (urban canyons), areas of dense tree cover, and even in some buildings and shopping centres. Considering these improvements, it may now be possible for GPS devices to completely replace the traditional HTS.

Traditional HTSs depend on respondents to report the details of their travel. Significant variables for transport modelling include origin and destination, travel time, time of day and travel distance. GPS units have the capacity to record all of this information in fine detail and provide information about routes that was previously unobtainable. The main advantages of using GPS over HTSs are, therefore, a reduced burden on participants, and a comprehensive, unbiased record of trip data. It is well documented in HTSs that respondents misestimate trip lengths and forget about shorter journeys – this should not be a problem with GPS devices. However, there are limits to what kind of information a GPS device can record. GPS is able to collect accurate data on:

- Location
- Time of each measurement
- Speed of movement
- Direction of movement (heading)
- Data quality measures

Data can be collected as frequently as every second, or as far apart in time as the user desires. For travel survey purposes, data-logging devices are ideal; i.e. the data collected are stored in the device for later retrieval and analysis. Transmission of the data in real time is not generally advantageous in this case and may be expensive to achieve. It can also raise major issues about privacy and intrusion for the user. A device with 8 Mb of memory is sufficient to store second-by-second travel for several months for the average urban dweller.

However, to fully replace a conventional HST, additional data are needed that GPS devices do not collect. The most important of these are method of travel used for each trip (mode identification), and the reason each trip was carried out (purpose identification). This paper will show that, while current GPS devices cannot record this information directly, it is possible to infer this information with a high degree of accuracy, provided that certain supplemental data are also available.

2. Mode and purpose analysis

GPS devices do not directly provide any information on purpose or mode. However, both of these can be deduced, with high accuracy, from the trip information recorded by the GPS device, provided there are adequate Geographic Information System (GIS) databases for the urban area and adequate demographic information about the GPS user.

For the purposes of this paper, it is assumed that the GPS data have already been divided into trips (Fitzgerald et al, 2006). Simplistically, a trip end is defined as occurring when there is no movement for a period of 120 seconds or more. A trip start is then defined as the next instance of movement following a trip end. This processing of the GPS data into trips is done using software developed by the authors.

After the data have been divided into trips, we generate maps of all trips and do a visual check to look for possible missed trip ends and spurious trip ends – the trip repair process. This process also checks for unconnected consecutive trips i.e. where the start of one trip is not the same as where the previous one ended. In this case “artificial data” is introduced to link the two, based on the street network. Generally, the program misses about 5 percent of actual trip ends, because the stop lasts less than 120 seconds, and also identifies stops longer than 120 seconds as false stops (e.g. traffic stops) about 5 percent of the time. Most of these are corrected through the visual checking. This is required for mode identification, as discussed in the next section of this paper.

2.1 Mode identification

In this study, 4 different modes of transport are considered. These are by foot (walk), bicycle (bike), private vehicle (car) and public transport. Public transport can be further divided into rail, bus, tram and ferry, depending on the region of interest. For mode identification, the following GIS information is required

- street network
- all public transport routes (including rail and subway lines)
- all bus stops and station locations

This information needs to be stored using a geographic referencing system compatible with the data collected by the device (typically latitude and longitude).

The identification of travel mode is a hierarchical process, using heuristics based on speed and route of travel, as well as some demographic information. At the outset, the easiest mode to identify is walk, because of the consistently low speeds for the entire trip segment. Off-network public transport modes, such as rail and ferry, are identified next. This is a relatively simple process, as the trip route will coincide with rail lines or ferry routes which are not on the street network. For underground rail trips, gaps will appear in the trip data, as the device may not be able to pick up signal. This is fixed in the data repair process.

In all cases, when speeds and acceleration or deceleration are used to assist in identifying a mode, we use the 85th percentile value. This helps by removing excessive values that are erroneously created by the GPS device, and do not reflect actual movement. For example, in second-by-second GPS data, the device may give an occasional point that is located off the route, as a result of momentary problems in determining its position. Such off-route points can also produce high figures for speed and acceleration. Using the 85th percentile value means that such spurious values are ignored.

After identifying the walk and off-network public transport modes, the next to be identified is bus. This is based again on maximum (85th percentile) and average speed, acceleration, and on the trip segment beginning and ending close to a bus stop. The trip should also decelerate near at least two bus stops along the trip.

Bicycle trips are identified next. The demographic information is examined to see if that person has a bicycle in their household. If not, then no trip segments are assumed to be by bicycle. If at least one is owned, then the bicycle trips are identified by examining the maximum speed, average speed and acceleration, and that the trip origin is either home or is a location to which this person has already travelled by bicycle.

At this point, all remaining trips should be trips by car. However, a further check is made of maximum speed and acceleration, and also that the trip segment remains on the road network. If these are correct, then the trip segment is identified as being by car. At this point it is not determined whether the trip is as a driver or a car passenger. However, if the person does not hold a current driving license, then it could be assumed that the person is a passenger.

2.2 Purpose identification

To identify the purpose for each trip, certain demographic data must be obtained for all survey participants. The following information is collected

- Address of main workplace for each working household member
- Address of main educational establishment for household members that are studying
- Addresses of the two most frequently used supermarkets/grocery stores

Where a survey participant does not have a fixed work address (eg home care nurse, tradesman etc) the usual hours of work are collected. The home address is already known, as it is used for delivery and collection of the GPS devices and other survey materials. All of these locations are then geocoded i.e. their positions are located within our GIS databases and registered. Since participants rarely give complete address information, it is important to have a GIS database which lists the use of each location on the map – a “land use” layer.

The purpose identification procedure depends on the level of information required. For this study, it was sufficient to break trip purpose down into the following categories:

- Home-based work – a trip from home to work
- Home-based education – a trip from home to school
- Home-based shopping – a trip from home to a store
- Home-based other – any other trip from home
- Non-home-based work – a trip to work not starting at home
- Non-home-based school – a trip to school not starting at home
- Non-home-based shopping – a trip to a store not starting at home
- Non-home-based other – any other trip

The purpose identification process is also hierarchical. First, a search radius is defined – typically 200m. Then the endpoint of the trip is identified. The work location (if any) is checked to see if it is within the search radius of the end point. If so, we have a match, and the trip is a work trip. If not, the school location is checked, then finally the shopping locations. The start-point details are taken from the end point of the previous trip.

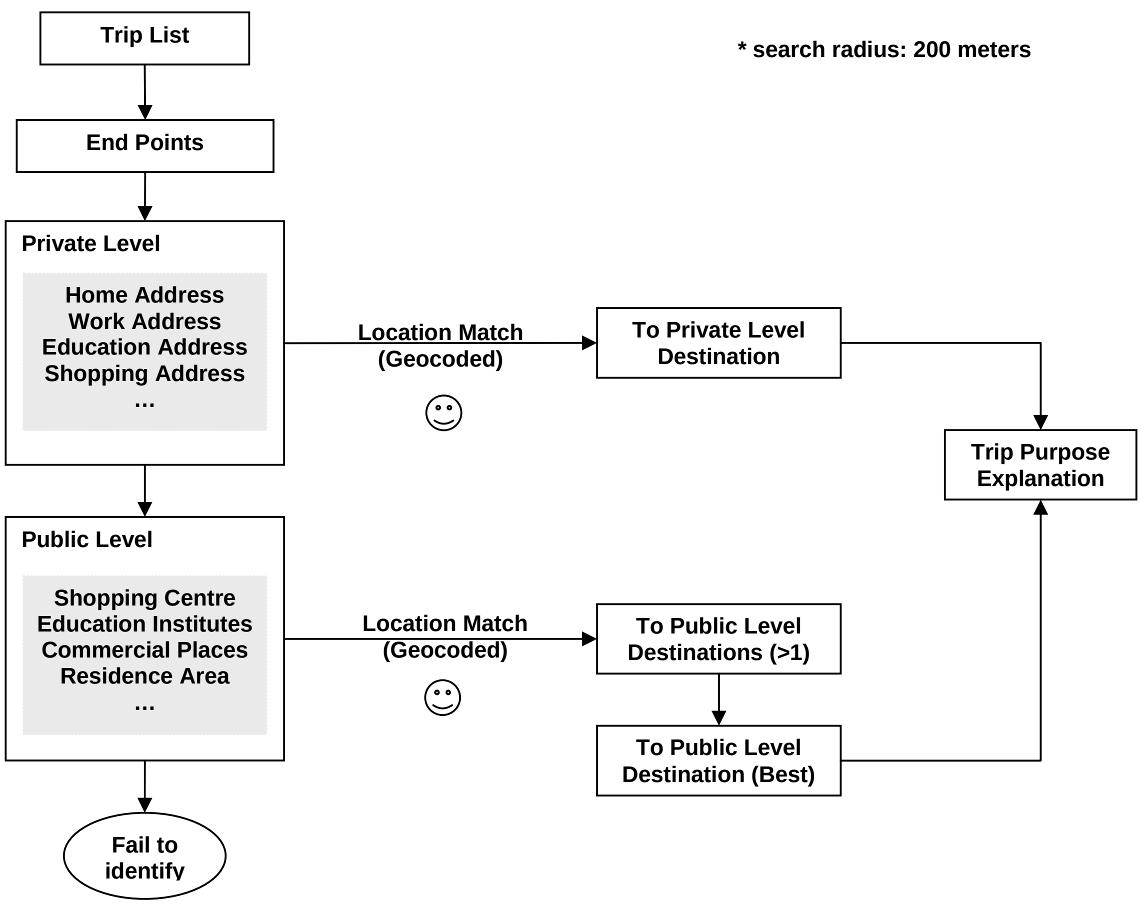


Figure 1: Trip Purpose identification process

If more detail is required, then the procedure may follow the steps shown in Figure 1. This takes into account 'public level' addresses, such as cinemas, libraries etc, which may be visited by anyone in the locality.

3. Example of trip determination

The following example shows results obtained from a travel survey carried out in the greater Adelaide area of South Australia. In this survey, GPS devices were carried by respondents for a period of time. Once the devices are returned, the stored data are downloaded and processed. The initial data output, converted to a csv format, will appear generally like that shown in Figure 2.

```
V,07/03/2006,12:58:49,138.509622,-34.843843,500,78,3,3,8.3
V,07/03/2006,12:58:51,138.509650,-34.843825,500,77,3,3,8.3
V,07/03/2006,12:58:53,138.509686,-34.843809,500,65,2,3,8.3
V,07/03/2006,12:58:55,138.509714,-34.843786,500,66,3,3,8.3
V,07/03/2006,12:58:57,138.509732,-34.843779,500,68,2,3,8.3
A,07/03/2006,12:59:00,138.509741,-34.843580,500,26,8,4,11.6
A,07/03/2006,12:59:02,138.509778,-34.843454,500,34,3,4,11.6
A,07/03/2006,12:59:04,138.509805,-34.843373,500,53,3,4,11.6
```

Figure 2: Data as downloaded from a GPS device

Deployment information is stored in the file name itself. The data shown here includes a data quality indicator, position, time, heading, speed, etc. for each data point, where data are collected every two seconds. The entire processing procedure is shown in Figure 3.

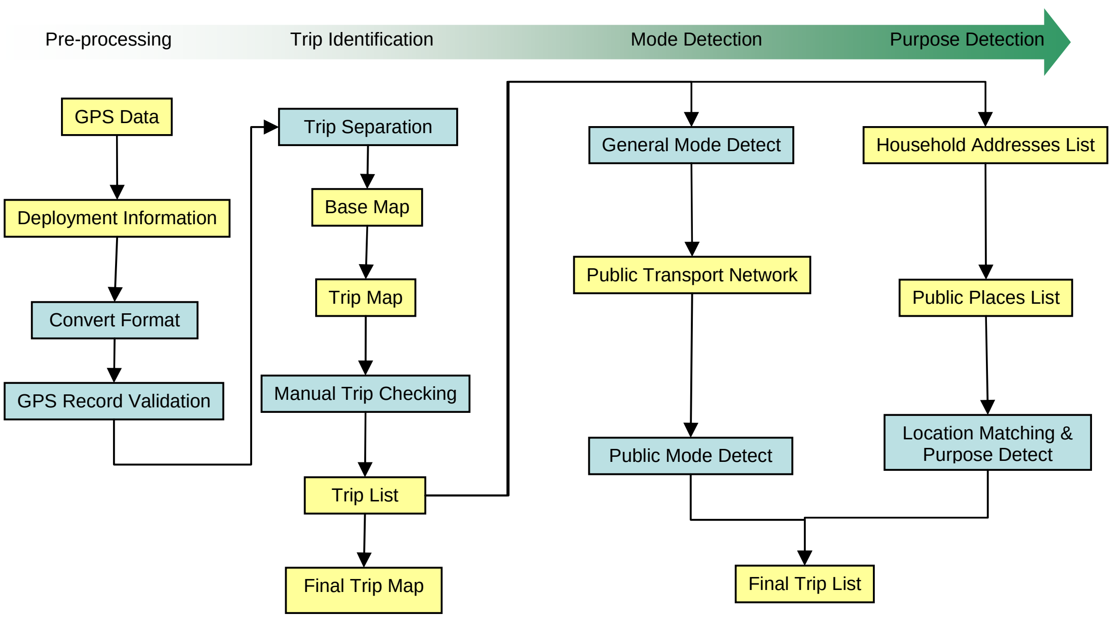


Figure 3: GPS data processing procedure

Figure 4 shows an example of what the data might typically look like before editing, and on a background GIS map that shows layers of the streets, bus stops, bus routes, etc.

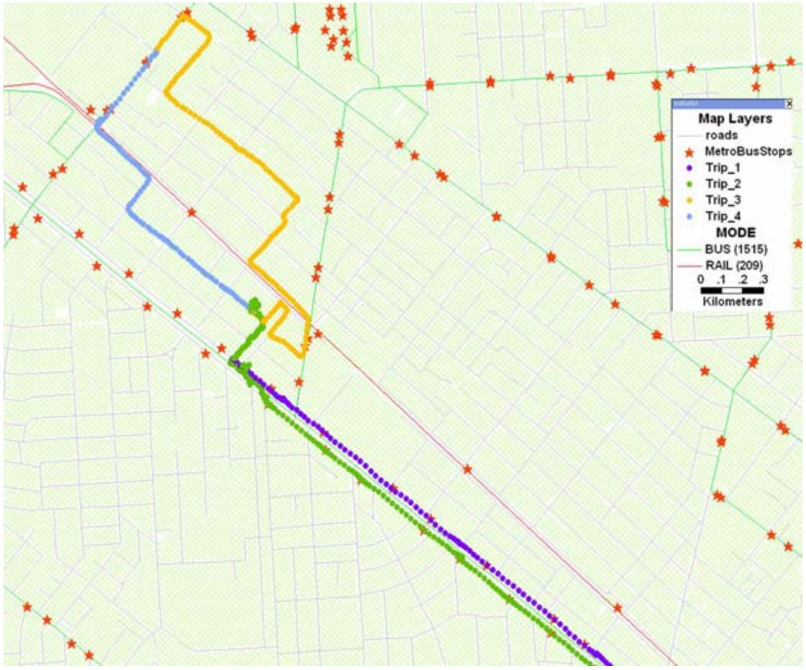


Figure 4: Unprocessed GPS data on a base map

The actual traces of the person's travel are shown by coloured points for each of the four trips that were identified by the initial processing of the data. After editing, the trips are converted to lines with directional arrows included. This is shown in Figure 5. Trips are also separated into segments if a different mode of transport is suspected (eg walking from bus stop – trip 2a). Any gaps in travel are filled in.

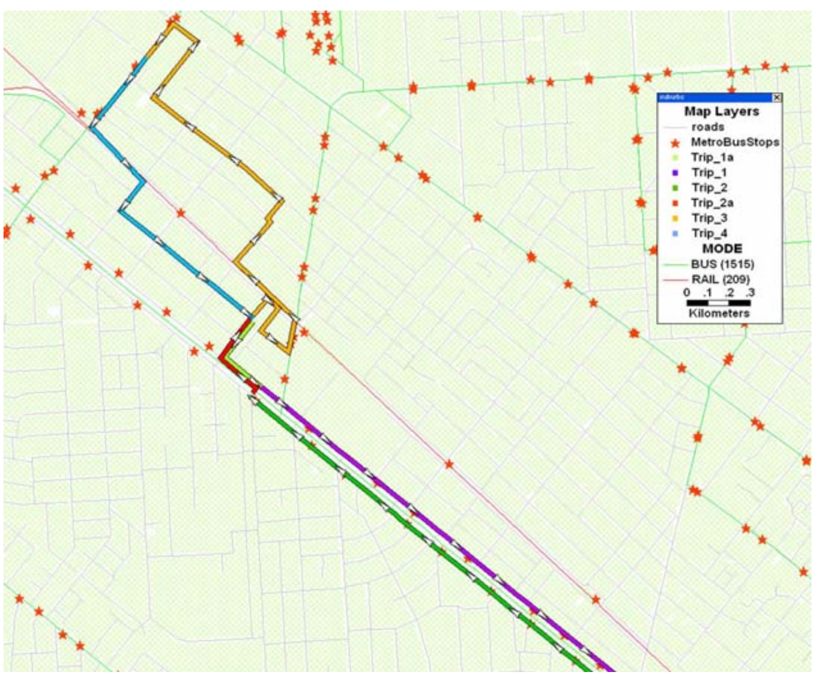
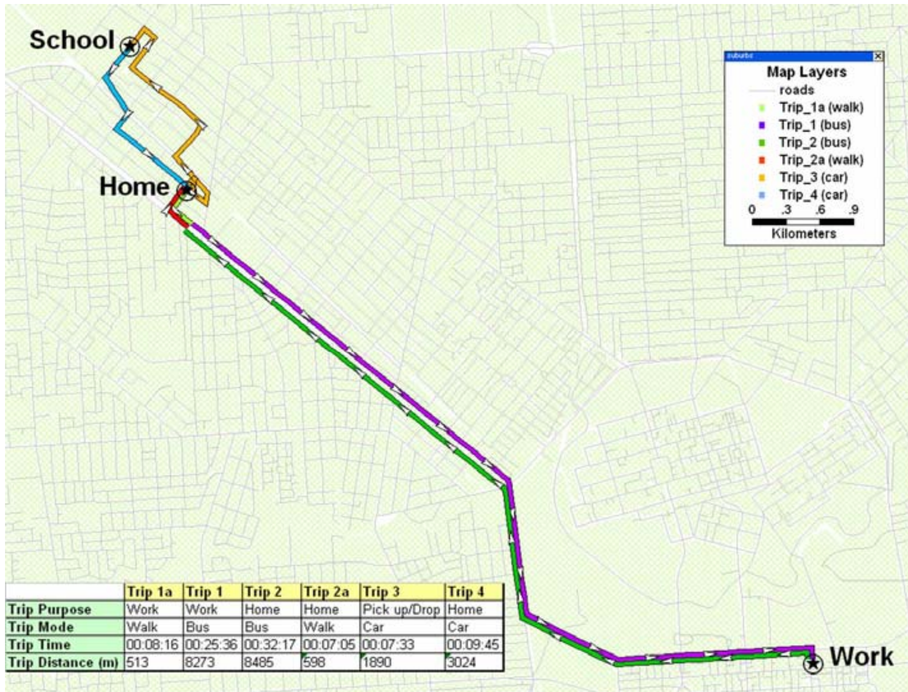


Figure 5: Post editing of the GPS traces

At the next processing step, the mode of travel is identified for each of the trips, as shown in Figure 6. This is indicated in the key to the map, in this case, where trips 1a and 2a are identified as walk, trips 1 and 2 as bus, and trips 3 and 4 as car. Finally, the purpose of the trips is deduced, and a final map and summary table are produced.



	Trip 1a	Trip 1	Trip 2	Trip 2a	Trip 3	Trip 4
Trip Purpose	Work	Work	Home	Home	Pick up/Drop	Home
Trip Mode	Walk	Bus	Bus	Walk	Car	Car
Trip Time	00:08:16	00:25:36	00:32:17	00:07:05	00:07:33	00:09:45
Trip Distance (m)	513	8273	8485	598	1890	3024

Figure 6: Final result of processing with explanatory table

4. Conclusions

Mode and purpose can be deduced with a quite high degree of accuracy from GPS traces, provided that certain supplementary data are available. Appropriate GIS layers are essential for this process, and supplementary data are required from participants, such as bicycle ownership, driver's licence status and addresses of certain frequently-used locations. All of the remaining information required for mode and purpose identification is readily available from the GPS record.

It would be useful to assess the proportion of trips reported in conventional surveys where the mode and purpose are well defined. It is possible that the deduced mode and purpose, as described in this paper, achieves a similar level of accuracy to that achieved in self-report diaries. A further study is planned by the authors to investigate this. Further refinement of the procedures outlined here is also currently underway, and improvements may be possible by introducing some form of Artificial Intelligence procedures to further refine the results obtained.

References

Baffour, R.A. (2002), "Collecting Roadway Cross Slope Data Using Multi-Antenna-Single Receiver Global Positioning System (GPS)", in *Applications of Advanced Technologies in Transportation: Proceedings of the 7th International Conference*, K.C.P. Wang, S. Madanat, S. Nambisan, and G. Spring (editors), American Society of Civil Engineers, Reston, VA.

Bullock, P. and Jiang, Q. (2003) 'Measuring Bus Performance Using GPS Technology', Paper presented at the 26th Australasian Transport Research Forum, 1-3 October 2003, Wellington, New Zealand.

http://www.itls.usyd.edu.au/publications/working_papers/wp2003/its-wp-03-15.pdf

Draijer, G., N. Kalfs, and J. Perdok (2000), "GPS as a Data Collection Method for Travel Research", *Transportation Research Record No. 1719*, National Academy of Sciences, Washington, DC, pp.147-153.

FitzGerald, C., Zhang J. and Stopher P. (2006). Processing GPS Data for Travel Surveys, paper presented at the IGNSS Annual Meeting, Brisbane, July.

orrest, T.L. and Pearson, D. (2005), 'Comparison of Trip Determination Methods in Global Positioning System-Enhanced Household Travel Surveys', Paper presented at the 84th Annual Meeting of the Transportation Research Board, January 9-13, 2005, Washington DC.

Li, H., Guensler, R., Ogle, J. and Wang, J. (2004), 'Using GPS Data to Understand the Day-to-Day Dynamics of the Morning Commute Behavior', Paper presented at the 83rd Annual Meeting of the Transportation Research Board, January 11-15, 2004, Washington DC.

http://commuteatlanta.ce.gatech.edu/Resources/TRB2004-003029.pdf

Quiroga, C. A. (2000), "Performance measures and data requirements for congestion management systems", *Transportation Research Part C*, 8, pp. 287-306.

Quiroga, C. A., Perez M., and Venglar S. (2002), "Tool for Measuring Travel Time and Delay on Arterial Corridors", in *Applications of Advanced Technologies in Transportation: Proceedings of the 7th International Conference*, K.C.P. Wang, S. Madanat, S. Nambisan, and G. Spring (editors), American Society of Civil Engineers, Reston, VA.

Ranjitkar, P., Nakatsuji T., Gurusinghe G.S., and Azuta Y. (2002), "Car-Following Experiments Using Real Time Kinematic (RTK) Global Positioning System (GPS) and Stability Characteristics of Followers in Platoon", in *Applications of Advanced Technologies in Transportation: Proceedings of the 7th International Conference*, K.C.P. Wang, S. Madanat, S. Nambisan, and G. Spring (editors), American Society of Civil Engineers, Reston, VA.

Stopher, P.R., Bullock P. and F. Horst (2002). "Exploring the Use of Passive GPS Devices to Measure Travel," in *Proceedings of the 7th International Conference on Applications of Advanced Technologies to Transportation*, edited by K.C.P. Wang, S. Medanat, S. Nambisan, and G. Spring, ASCE, Reston VA, pp. 959-967.

Stopher, P.R., Greaves, S.P., Xu, M., FitzGerald, C., and Lauer, N. (2005) "A Panel Approach to Evaluating TravelSmart Initiatives in the Short Term – South Australia

Pilot Survey", paper presented to the 28th Australasian Transport Research Forum, Sydney, 28-30 September.

http://www.patrec.org/atrf/papers/2005/Stopher,%20Greaves,%20Xu,%20Fitzgerald,%20Lauer%20&%20Perkins%20(2005).pdf

Stopher, P.R., FitzGerald C., and Biddle T. (2006). "Pilot Testing a GPS Panel for Evaluating TravelSmart", paper submitted for presentation to the 22nd ARRB Research Conference, Canberra, October-November.

http://www.its.usyd.edu.au/publications/working_papers/wp2006/itls-wp-06-16.pdf

Stopher, P.R., S.P. Greaves, and FitzGerald C. (2005). "Developing and Deploying a New Wearable GPS Device for Transport Applications", paper presented to the Second International Colloquium on the Behavioural Foundations of Integrated Land Use and Transportation Models: Frameworks, Models, and Applications, Toronto, June 13-15.

http://www.patrec.org/atrf/papers/2005/Stopher,%20Greaves%20&%20Fitzgerald%20(2005).pdf

Wee, S.H., Cheu R.L., Lee D-H, and Chan W.T. (2002), "Using Global Positioning System (GPS) Data to Calibrate Arterial Speed Profile in PARAMICS", in *Applications of Advanced Technologies in Transportation: Proceedings of the 7th International Conference*, K.C.P. Wang, S. Madanat, S. Nambisan, and G. Spring (editors), American Society of Civil Engineers, Reston, VA.

Wolf, J., Loechl, M., Thompson, M. and Arce. C. (2003), "Trip Rate Analysis in GPS Enhance Personal Travel Surveys" in *'Transport Survey Quality and Innovation'*, P. Stopher and P. Jones, (editors) Elsevier, Oxford, pp. 483-498.

Wolf, J., Guensler R., and Bachman W. (2001). "Elimination of the Travel Diary. An experiment to derive trip purpose from GPS travel data". *Transportation Research Board*, 17688, pp. 125-134

Wagner, D.P. (1997), "Global Positioning Systems for Personal Travel Surveys: Lexington Area Travel Data Collection Test", *Report to the Federal Highway Administration, U.S.DOT by Battelle Transportation Division*, Washington, DC, September.

Wolf, J., Guensler R., Frank L., and Ogle J. (2000), "The Use of Electronic Travel Diaries and Vehicle Instrumentation Packages in the Year 2000 Atlanta Regional Travel Survey: Test Results, Package Configurations, and Deployment Plans", paper presented at the 9th International Association for Travel Behaviour Research Conference, Gold Coast, Australia.

http://onlinepubs.trb.org/onlinepubs/circulars/ec026/19_wolf.pdf

Wolf, J., Loechl, M., Thompson, M. and Arce. C. (2003), "Trip Rate Analysis in GPS Enhance Personal Travel Surveys" in *'Transport Survey Quality and Innovation'*, P. Stopher and P. Jones, (editors) Elsevier, Oxford, pp. 483-498.

Wolf, J. (2006). 'Application of New Technologies in Travel Surveys', in *Travel Survey Methods -- Quality and Future Directions*, P. Stopher and C. Stecher (editors), Elsevier, Oxford.