Australasian Transport Research Forum 2013 Proceedings
2 - 4 October 2013, Brisbane, Australia
Publication website: http://www.patrec.org/atrf.aspx

Comparing Two Processing Routines for GPS Traces – Lessons Learnt

Peter R. Stopher[1], Stephen P. Greaves[1], Li (Andrew) Shen[1]

[1] Institute of Transport and Logistics Studies (ITLS),

The University of Sydney Business School, The University of Sydney NSW 2006, Australia

Corresponding author: Professor Peter Stopher
Phone: 02 9114 1808; email peter.stopher@sydney.edu.au

Abstract

This paper describes what may be one of the first side-by-side tests of two alternative software products for processing GPS traces into trips, and discusses some lessons learnt from the comparisons. For GPS to be useful as an alternative to self-report survey mechanisms, it is imperative that good processing software becomes available to reduce the data streams from the GPS devices into specific trips, with the various attributes of trips that are needed for modelling purposes. Currently, a number of agencies and researchers around the world have developed alternative software products, but none of these are generally open source, and comparisons between them are almost non-existent, although most make claims to certain levels of accuracy. In this paper, we describe an exercise in which two software products were used on the same GPS data set, following which a detailed comparison was made of the results. While it is interesting to see, overall, the accuracy differences between the two software products, what is of even more interest is the lessons that can be learnt about processing software in general. The paper draws some conclusions about the directions forward for processing software and processing routines in general.

1 Background and Objectives

1.1 Background

Over the past decade, the Institute of Transport and Logistics Studies of the University of Sydney Business School (ITLS) has developed a worldwide reputation in the use of Global Positioning System (GPS) technology to support the collection of travel data within a range of applications and contexts. Recent examples include: i) GPS surveys of personal travel in Victoria (Stopher and Greaves, 2009) and Ohio (Stopher and Wargelin, 2010), ii) GPS surveys to support the evaluation of travel behaviour change initiatives (Ampt et al., 2006; Stopher et al., 2009; Stopher et al., 2012; Stopher et al., 2013), iii) a GPS freight survey in Victoria (Greaves and Figliozzi, 2008), and iv) a GPS survey conducted as part of an investigation into motorist responses to variable-rate pricing schemes in Sydney (Greaves et al., 2010). Through these various applications, ITLS has been directly involved in both the design of GPS devices that are suitable for collecting personal travel data to the levels of accuracy required, and the processing of the data into useful/meaningful information.

In terms of the processing, the challenge is how to take the 'raw' data from the GPS devices, which are simply second-by-second records of the position, time, velocity, and quality of that record (e.g., number of satellites, horizontal dilution of precision), and turn this into useful travel information including trip-ends, mode (car, bus, train, bicycle, walk, etc.), trip purpose (work, school, shop, etc.), travel time, distance, route and (potentially) vehicle occupancy. This challenge is confounded by the fact that we are (often) dealing with millions of data points, inevitable data errors due to signal loss/degradation, and a desire to minimise the amount of time required from survey participants to both check and provide additional information that could be inferred from the GPS data. Taking this into account, ITLS has developed a sophisticated suite of software called G-TO-MAP to process GPS data based on heuristics/rules, which is described in more detail in Stopher et al. (2011). From tests that have been run to date, using prompted recall survey responses (in themselves not totally reliable) to verify inferred information, G-TO-MAP is about 96 percent accurate in identifying trips, about 90 percent accurate in identifying mode, and about 67 percent accurate in identifying purpose. While software enhancements have recently been made that are expected to improve these attributes, especially the purpose identification, few if any software packages that have been developed to date for processing GPS data achieve better accuracy than this (Wolf et al., 2001; Griffin and Huang, 2005; Troped et al., 2008; Bohte and Maat, 2009; Biljecki, 2010; Gonzalez et al., 2010).

Given the explosion in the use of GPS to support travel data collection, other groups around the world are facing similar issues and trying to deal with this problem. It seems that it would be useful if the results of using two or more processing packages could be compared to see whether there are material differences in the results obtained.

1.2 Objectives

ITLS was offered GPS data from another source that could be processed through G-TO-MAP and the results compared to the results of another agency that used the same data and their own processing software. This provided an unique opportunity to compare two processing routines side-by-side on identical raw GPS data, to determine whether or not there were differences in the results. If differences were found, then those differences could potentially identify possible improvements to either software package and would also provide the opportunity to determine factors affecting the output accuracy of processing GPS data streams. For a variety of reasons, the alternative software product was requested to be kept anonymous, so is referred to throughout this paper simply as the "other software". However, the software has been developed and is used by a highly reputable firm in a number of GPS surveys. The fact that the authors did not have direct access to the software is, however, a limitation on this study.

2 Data and Methodology

The data that were made available were recently collected GPS data from California, consisting of data from 3 to 7 days from 22 households. The data were assembled from both the other agency's processing and G-TO-MAP and set up for a trip-by-trip comparison. In comparing, the first task was to match each trip from the other agency's processing to a trip from G-TO-MAP, or to determine that there was no match. Where there was not a match, the next step was to investigate in more detail the raw data and determine the reasons for a match or lack of a match.

It was also discovered that, whereas G-TO-MAP is able to impute travel mode and purpose, the other processing software only splits the data stream into trips. Therefore, a comparison of mode and purpose from the two software packages could not be made.

2.1 Processing with G-TO-MAP

To understand the results of the processing comparison more fully, it is important to understand the processing steps involved with G-TO-MAP. In addition, a new capability was added to G-TO-MAP in this project. The new capability is a conversion program that allows G-TO-MAP to be run on data of any configuration from a GPS data collection. Prior to this project, G-TO-MAP required input data to be in a specific format, dictated by the GPS devices that have been used at ITLS over the time that the software has been developed. As part of this project, a new routine was added called DataConvert, which allows conversion of data from any input format to the subsequent requirements of G-TO-MAP. This project also represented the first real test of this new routine.

After the data were received, DataConvert was run and the reformatted data were then ready for full processing. At the outset of the processing, it is necessary to have a digital Geographic Information System (GIS) base map that includes layers for each means of travel – car, bus, train, etc. – so that the subsequent processing can use this information to partition the data streams into individual trips and also identify the means of travel used. Once the base map was completed, processing could begin.

The first step in the processing is to convert the GPS data into point layers for a GIS. Using a number of heuristic rules, the stream of data is then broken up into trips, where a trip is usually defined as a movement from one place to another, without any stops lasting more than 120 seconds. Additional processing steps are designed to remove data points that are not useful and to repair data where there may be gaps in the GPS points. Data points that are not useful are primarily those that are recorded when the GPS device is at rest (because the person is no longer travelling), and prior to the device switching to sleep mode. A few points may be ones that have too few satellites in view or too high a value of the HDOP to be useful.

When a GPS device is stationary, for example, it continues to record position information at the prescribed frequency (in this case, every second), but these estimates of position will vary and not record a fixed point. The reason for this is that the GPS device calculates its position by solving equations based on the positions of at least four satellites. Because of imprecision in the information about the actual distance to the satellites, the results of this solution will change by a number of metres each second. Thus, the GPS device generates what we term a 'cloud' of points around a stationary location. The next step in the software processing is to remove these clouds of data points and replace them by a single point at the apparent stopping place. A further step in the processing repairs 'cold starts' and signal losses in the records, a step that is not included in most other software available today. Cold starts occur when a device has been stationary for some time and the user of the device commences travel. The device has to solve for its initial position and this may take a number of seconds (even a minute or more in some cases typically depending on how long the device has been stationary and how fast the person is travelling). This produces a cold start, where there will be a gap between the end of the previous trip and the start of the next one. The software repairs this gap by changing the start location to match the previous end location. Similarly, if a gap appears in the middle of a trace, this is checked to determine if it is likely that the person travelled continuously through this gap, or if there might be time enough for a stop. In either case, an appropriate repair is made.

The final steps of the initial processing are to generate both a GIS map of each day's trips and also a Google Earth® map for each day. These maps are used for a manual review process, in which the mapped trips are reviewed to determine if any spurious trips (i.e., traces that are generated by the GPS device but which do not represent a real movement) remain. The software also detects situations where there is an apparent gap between the destination of one trip and the origin of the next trip, which is flagged to the reviewer. The reviewer also checks to see if there may be very brief stops along the route that were not detected by the software, but which may show up as a concentration of points in a particular location, usually off the relevant network.

Following this manual review process, which ITLS terms 'map editing', the resulting modifications are checked, certain data are automatically entered on changed trips, and then the trips for all days are combined into a single file for each person. This last file is an Excel® file that conforms to the normal format expected for a trip file in transport planning applications. This file is the one that is used in this research to compare the outcomes from the processing by G-TO-MAP with that from the other processing software.

3 Results of the Comparison

The initial results from the processing (including the manual map editing for G-TO-MAP) were that G-TO-MAP produced a file of 1,267 trips, while the other software produced 1,133 trips, or a difference of 134 trips. One could simply assume that G-TO-MAP had identified some additional trips, while there were 1,133 trips that presumably matched. However, this was not assumed and each trip from each of the two processing procedures was compared, to determine whether or not the two software packages had identified identical trips. The result of this trip-by-trip comparison is shown in Table 1.

Table 1: Results of the Comparison of G-TO-MAP and the Other Software Trips

Comparison Result	Number
Trip Matches in Both Results	1,031
Single Trip in the other software split to two or more trips by G-TO-MAP	67
Single Trip in G-TO-MAP split to two or more trips by the other software	64
Inserted missing trip in G-TO-MAP	8
Trip identified by G-TO-MAP only	131
Trip identified by the other software only	6
Trip retained by the other software with incorrect lat-long	2
TOTAL	1,309

As Table 1 shows, there was agreement on just over 81 percent of the trips identified by G-TO-MAP (or 78.8 percent of the total trips identified by either software). There is an almost equal incidence of cases where one software package split a trip into two or more separate trips, while the other software considered it to be one trip. However, these trips were reviewed and it was determined that the most probable situation is that all 67 trips that G-TO-MAP identified as multiple trips compared to the other software are, in fact, upon detailed inspection, trips that should have been split by the other software. Of the 64 trips that were split by the other software and not by G-TO-MAP, 25 probably should have been split by G-TO-MAP. Of the remainder, 30 trips show no obvious evidence of a stop within the trip, while 9 should not have been split (4 of these are at intersections and are simply long traffic waits, and 5 are all part of one continuous trip but have some signal missing issues). Of course, none of these figures are 100 percent certain, because there is no available ground truth (Shen and Stopher, 2013a). However, a careful and detailed review of the mapping of the recorded data indicates this is the most probable situation. One of the issues that causes problems in the other software becomes evident in this examination. This is that the other software seems to discard a number of valid data points, resulting in stop times of longer than 120 seconds, artificially, because data points have been edited out.

In eight cases, the ITLS procedure ended up inserting a missing trip, where there was a gap between the latitude-longitude of one trip end and the start of the next trip, which did not appear to be a case of a cold start or a premature termination of a trip. These were usually trips at the beginning or end of the day. After re-checking, all eight appear to be correct. There are also 131 trips that G-TO-MAP identified as trips, but which the other software deleted as spurious. After re-checking these, it appears that there could be nine of these trips that are actually spurious, but all of the remaining trips appear to be correct in G-TO-MAP. There are two trips in the remainder that a second examination suggests may be both part of a single trip, leaving 121 trips that were identified by G-TO-MAP, but not by the other software. Because ITLS does not have the details of the rules used by the other software, we surmise that it may have discarded these as spurious trips, when, in fact, they are not. This may indicate that rules for detecting spurious trips are either too simplistic, or too stringent in the other software, resulting in deletion of about 10 percent of actual real trips. In contrast, there were six trips that the other software retained as trips, but which G-TO-MAP did not identify as real trips. A re-examination of these trips confirms that five are actually spurious, and only one trip is a genuine trip, which G-TO-MAP did not identify, because there was a signal loss problem in the trip also. Figure 1 shows an example of the type of situation encountered in comparing results of the two softwares.

From the results of the other software package, the route of the tour shown in Figure 1 is O1→D1, O2→D2. Apparently, there is a missing part from D1 to O2. This missing part could be one trip or several trips. From the results of G-TO-MAP, it clearly shows that the route of this tour should be O1→D1→MD1→O2→D2. In this case, the other software package failed to identify two trips, i.e., from D1 to MD1, and from MD1 to O2.[1]

[1] O1 stands for origin 1, D1 stands for destination 1, and MD1 stands for missed destination 1.

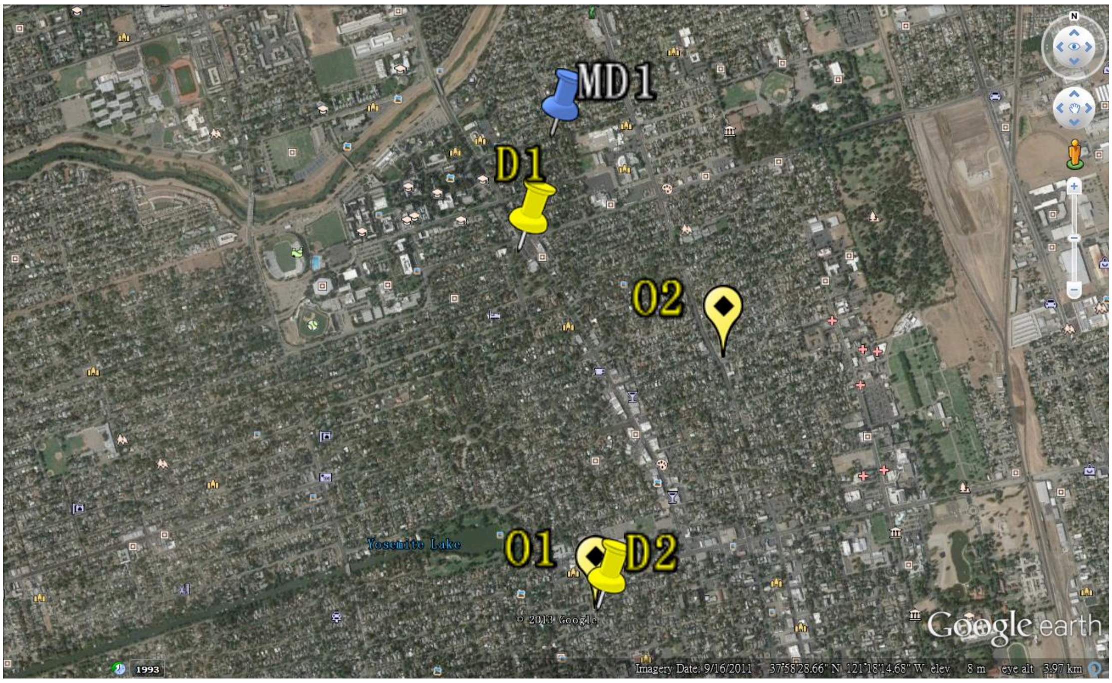

Figure 1: Example of the Difference between G-TO-MAP Result and the Other Package’s Result

Finally, there were two trips that had latitude and longitude from Memphis Tennessee, although there was no trip to there on a prior day or returning to California on a later day. These trips were excluded by G-TO-MAP, but retained by the other software.

While a majority of trips were identified by both software products, there were differences in the start and end times of virtually every trip. The average travel time of the matched trips from the other software was 15 minutes and 17 seconds, whereas for G-TO-MAP, it was 14 minutes and 15 seconds. Further, the minimum travel time from the other software trips was 2 minutes, whereas it was 30 seconds for G-TO-MAP. On the other hand, the maximum travel time in this data set from the other software was 3 hours, 14 minutes and 0 seconds, whereas it was 3 hours 14 minutes and 35 seconds from G-TO-MAP. On a trip-by-trip basis, there was almost no exact match on the times. Again, this appears to be a result of the more stringent validity rules in the other software, which results in data points being disregarded that G-TO-MAP considers as valid. This means that some data points at the start and end of a trip may be discarded by the other software and retained by G-TO-MAP, resulting in a different time start and end for the trip.

3.1 Summary of Differences

Considering first the trips identified by the other software: a total of 1,133 trips were identified by the software, of which 2 were in Memphis, TN and should be excluded, 5 were spurious traces that the other software counted as trips, and 39 trips were split by the other software and should not have been. These splits that should not have occurred represent what are, in fact 18 actual trips, meaning the other software over-counted trips by 21 in this area. Hence the final net total of trips that the other software identified that are genuine trips, as far as it is possible to ascertain from visual examination, is 1,105 trips.

It appears, however, that the other software failed to split 31 trips into two or more trips, which would have added 36 additional trips. There were 8 trips that needed to be inserted to connect a destination to the next origin, and 121 trips that were not identified by the other software as trips. This would bring that actual total to 1,270 trips. Based on this assessment, the accuracy of the other software is 87.0 percent on trip identification.

From G-TO-MAP, there were initially 1,267 trips identified. Of these, 9 trips were possibly spurious, and 2 trips should have been combined into a single trip. This leaves a net total of 1,256 trips identified by G-TO-MAP. In addition, the results from the other software suggest that there are a further 12 trips identified by G-TO-MAP that should have been split to create 25 trips in total, and there was one trip that the other software identified that G-TO-MAP did not, because of signal loss issues. Thus, it appears that there should have been 14 additional trips that G-TO-MAP should have identified, but did not. This suggests that the correct number of trips that should have been identified by either software product is 1,270. Given that G-TO-MAP identified 1,256 correctly, G-TO-MAP appears to have an accuracy of 98.9 percent.

4 Additional Results from G-TO-MAP

In addition to identifying trips, G-TO-MAP also has procedures to impute mode and purpose. Part of this data set was collected using in-vehicle devices, for which mode is given. However, for the hand-held devices, mode was determined using G-TO-MAP. The results of the mode identification are shown in Table 2, which includes the results from the in-vehicle devices as well as the hand-held devices.

Table 2: Mode Results from G-TO-MAP

Imputed Mode	Number
Car	1,227
Bicycle	18
Bus	1
Walk	20
TOTAL	1,266

There was one trip for which mode could not be identified. For this reason, the total in Table 2 is 1,266 instead of 1,267. The number of bicycle trips is almost certainly too high, because this has been the experience previously with mode identification. Some of the trips identified as bicycle are probably car trips, especially where the car is travelling in moderate congestion.

Trip purpose was also imputed by G-TO-MAP. Normally, ITLS collects data on the address of each workplace, each educational establishment, and the two most frequently used grocery stores. With this information, we would expect more accurate purpose identification. Table 3 shows the results of purpose imputation from G-TO-MAP.

Table 3: Results of Purpose Imputation by G-TO-MAP

Trip Purpose	Number

Home-based work	114
Home-based education	20
Home-based shopping	34
Home-based other	428
Home-based mode change	4
Non-home-based work	156
Non-home-based education	17
Non-home-based shopping	30
Non-home-based other	445
Non-home-based mode change	16
Non-home-based serve passenger	3
TOTAL	1,267

The rather large number of trips with the purpose ‘other’ is a result of the lack of geocodable address information for shops and schools. The inclusion of mode change and serve passenger purposes indicates that the trips are unlinked trips, since trip linking would normally remove these purposes.

Unfortunately, for mode and purpose there is no ground truth available, so that it is not possible to state claims of accuracy for these statistics. In addition, the other software did not have this capability, so a comparison of software is also not possible.

5 Conclusions

Both programs – G-TO-MAP and the other software – were run on a single small data set, consisting of a mixture of hand-held and in-vehicle GPS devices used by a sample of persons in California. From the results of this test, the previously-claimed accuracy of G-TO-MAP of about 99 percent in identifying trips seems to have been confirmed. In contrast, the other software appears to be about 87 percent accurate, with the largest problem apparently arising from a too stringent definition of what comprises a valid observation from the GPS device, and an overly simplistic definition of spurious trips, which results in deleting a number of actual trips, especially walk trips. It must be kept in mind, however, that there was no ground truth against which to assess these results, so there remains some element of subjectivity in the final assessment.

In terms of improvement to G-TO-MAP, two conclusions can be drawn. First, the program that was developed to convert data of different formats from different GPS devices worked well. Thus, G-TO-MAP should be able to function effectively on data from any GPS device. However, it is also necessary to keep in mind that both the number of satellites in view and the HDOP are necessary for correct assessment of validity of data points. If either of these is missing, then the identification of valid data points is compromised. Second, the only improvement that appears to be warranted, which could reduce the errors from G-TO-MAP relates to identifying short stops in the midst of a trip. Currently, both the other software and G-TO-MAP use a 120-second rule, i.e., that there is a trip end if the time elapsed between two periods of movement is 120 seconds or greater, then this is a stop. G-TO-MAP then relies on a manual review of the output maps to identify shorter stops, which are clearly subject to the level of attention of the reviewer, and his or her ability to pick out, from a scan of a trip, a potential short stop (Shen and Stopher, 2013b).

It must be further noted that it is the manual review step that represents the biggest difference between the other software and G-TO-MAP and that it is this process that probably accounts for most of the 12 percent difference in accuracy between the two products. Ways to automate this process are needed to reduce the overall processing costs, however. One possible way is using probabilistic methods that involve learning algorithms to see whether they can help to reduce the manual work and achieve better results. This is future research that is needed.

6 References

Ampt, E., P. Stopher, and J. Wundke (2006). Households on the Move – Experiences of a New Approach to Voluntary Travel Behaviour Change, *Transportation Research Record No. 1985*, pp. 98-105.

Biljecki, F. (2010). Automatic segmentation and classification of movement trajectories for transportation modes, *Dissertation, MSc in Geomatrics*, Delft University of Technology: TU Delft

Bohte, W. and K. Maat (2009). Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands, *Transportation Research Part C*, 17 (2009) 285–297

Gonzalez, P., Weinstein, J., Barbeau, S, Labrador, M., Winters, P, Georggi, N. and Perez, R. (2010). Automating mode detection for travel behaviour analysis by using global positioning systems-enabled mobile phones and neural networks, *IET Intelligent Transport Systems*, Vol. 4, Issue. 1, 03/2010, pp. 37 - 49

Griffin, T. and Huang, Y., (2005). A Decision Tree Classification Model to Automate Trip Purpose Derivation. *Proceedings of the ISCA 18th International Conference on Computer Applications in Industry and Engineering*, November 9-11, 2005, Sheraton Moana Surfrider, Honolulu, Hawaii, USA. 44-49.

Greaves, S.P. and Figliozzi, M. (2008) 'Collecting Commercial Vehicle Tour Data with Passive Global Positioning System Technology: Issues and Potential Applications', *Transportation Research Record*, vol. 2049, pp. 158-66.

Greaves, S.P., Fifer, S., Ellison, R. and Germanos, G. (2010) 'Development of a Global Positioning System Web-Based Prompted Recall Solution for Longitudinal Travel Surveys', *Transportation Research Record*, vol. 2183, pp. 69-77.

Shen, L. and Stopher, P.R. (2013a) 'Using SenseCam to Pursue ‘Ground Truth’ for GPS Travel Surveys’, 14th TRB National Transportation Planning Applications Conference, Colombus, United States, 5-9 May 2013

Shen, L. and Stopher, P.R. (2013b) ‘Should we change the rules for trip identification for GPS travel records’, paper submitted to the 36th ATRF Meeting, Brisbane, October.

Stopher, P.R. and Greaves, S.P. (2009) 'Missing and Inaccurate Information from Travel Surveys – Pilot Results', *32nd Australasian Transport Research Forum ATRF 2009*, Auckland, New Zealand, 29 September - 1 October.

Stopher, P.R., Y. Zhang, J. Zhang, and B. Halling (2009) 'Results of an Evaluation of TravelSmart in South Australia', *32nd Australasian Transport Research Forum ATRF 2009*, Auckland, New Zealand, 29 September - 1 October.

Stopher, P. and L. Wargelin (2010) 'Conducting a Household Travel Survey with GPS: Reports on a Pilot Study', *12th World Conference on Transport Research WCTR Conference 2010*, Lisbon, Portugal, 11 - 15 July.

Stopher, P.R. (2009). Collecting and Processing Data from Mobile Technologies, In *Travel Survey Methods: Keeping Up With a Changing World,* Patrick Bonnel, Martin Lee-Gosselin, Johanna Zmud, Jean-Loup Madre (editors), Emerald Group Publishing Limited, pp. 361-391.

Stopher, P.R., J. Zhang, and C. Prasad (2011). Evaluating and Improving Software for Identifying Trips, Occupancy, Mode and Purpose from GPS Traces, paper presented to the International Conference on Travel Survey Methods, Chile, November 2011.

Stopher. P.R., E. Clifford and B. Halling (2012). 'Evaluating a Voluntary Travel Behaviour Change by Means of a 3-Year GPS Panel', in G. Sammer, M.Bell, W. Saleh (eds) *Visions, Concepts, and Experiences of Travel Demand Management*, Ashgate Publishing Ltd., Farnham, Surrey, UK. (in press).

Stopher, P.R., C. Moutou, and W. Liu (2013). 'Sustainability of Voluntary Travel Behaviour Change – a 5-Year Study', paper submitted to the 36th ATRF Meeting, Brisbane, October.

Troped, P. J., Oliveira, M. S., Matthews, C. E., Cromley, E. K., Melly, S. J. and Craig, B.A. (2008). Prediction of activity mode with global positioning system and accelerometer data, *Med Sci Sports Exerc,* 2008 May, 40(5):972-8

Wolf, J. R. Guensler, and W. Bachman (2001). Elimination of the Travel Diary: An Experiment to Derive Trip Purpose from GPS Travel Data, *Transportation Research Record,* No 1768, Paper No. 01-3255. 125-134