



## GPS, GIS and personal travel surveys: an exercise in visualisation

Peter Stopher, Philip Bullock and Qingjian Jiang

*The Institute of Transport Studies*

---

### Abstract

In the past three years, a number of attempts have been made to use GPS devices to measure elements of person travel that have never been successfully measured by conventional interview and self-administered surveys. A key component of this application of GPS devices is to process the track points recorded, so as to produce maps and other visual representations of the travel conducted with the device. These maps and other visualisations of the travel are subsequently used in a prompted recall survey, to obtain additional data about the travel that cannot be measured by the GPS devices, such as travel purposes, number in the travelling party, and costs associated with the travel. Determining what constitutes a trip, and processing the data to produce a recognizable map of the travel is essential to the success of a prompted recall survey. In turn, the use of this type of survey avoids the need for survey subjects to enter data in a diary or electronic device during the travel – a task that is both burdensome and likely to be forgotten or omitted sufficiently often to negate most of the benefits of a GPS survey.

This paper describes the use of the GPS devices in this type of survey, the paradigms used to convert the track points to coherent trips, examination and correction of the visualised travel, and presentation of the resulting maps and other visual tools to subjects in a prompted recall survey. Methods to display the trips and other information that can be gained from alternative ways of presenting the data are also outlined in the paper. These include the ability to determine when a person travels in congested conditions, examination of delays at traffic lights and other controlled intersections, and identification of the locations of acceleration and deceleration episodes.

---

### Contact author

Peter Stopher, Professor of Transport Planning,  
The Institute of Transport Studies, The University of Sydney, NSW 2006.  
Phone: (02) 9351 0010 Fax: (02) 9351 0088  
e-mail: [peters@its.usyd.edu.au](mailto:peters@its.usyd.edu.au)

Philip Bullock, Research Analyst,  
The Institute of Transport Studies, The University of Sydney, NSW 2006.  
Phone: (02) 9351 0075 Fax: (02) 9351 0088  
e-mail: [philipb@its.usyd.edu.au](mailto:philipb@its.usyd.edu.au)

## Background

Over the past 3 years, there has been growing interest in measuring human travel patterns using GPS or other similar semi-automatic devices (Wolf and Arce, 2002; Stopher and Bullock, 2001). In most cases, individuals are recruited and asked to plug a GPS device into the car that the person usually drives. The GPS device then logs track points at a pre-specified interval, usually 1, 2, or 5 seconds. The device is capable of recording considerable quantities of data, which must subsequently be analysed so as to provide information useful to transport planners and others.

As discussed elsewhere (Stopher and Bullock, 2001), there are generally two types of GPS device. One type is a GPS device that is connected to a Personal Data Assistant (PDA) and which requires the respondent to enter data about each trip as the trip is started. This type of device was pioneered originally by the U.S. Federal Highway Administration, through work undertaken initially by Battelle (Wagner *et al.*, 1998; Wagner, 1997). It has been developed further in recent work by Guensler and Wolf (1999). The second type of device is a passive device that requires little or no intervention by the respondent, but which may require the use of a prompted recall survey, after the data have been collected and downloaded (Stopher and Wilmot, 2000; Wolf *et al.*, 2001; Stopher, 2001). It is this type of device that is the major focus of this paper.

## Objectives

The objectives of the work reported in this paper are to determine how to convert the data obtained from the GPS device into discrete trips, display these for a variety of potential uses, and also display other information of relevance that is collected by the device, such as speed or acceleration data.

In some applications, particularly those that ITS, Sydney has been working on, there is a need to be able to append other data from a separate prompted recall survey that is conducted with the respondent after the GPS data have been collected and analysed. This requires that the data collected by the GPS are processed to provide a visual record that is meaningful to a respondent. Such a visual record can be either or both maps and tabular presentations of the data. If the only issue were to come up with a reasonable way to show a respondent a set of track points on a map, this would be relatively trivial, involving simply transferring the data collected to a GIS and printing the resulting map. However, this is not the issue. Rather, the data need to be transformed from a set of track points collected over some period of time – ranging from a day to several weeks – to a set of line records, wherein each line represents a trip from an origin to a destination, and where the origins and destinations each represent locations where the respondent undertook some non-travel activity. In addition, a tabular presentation is needed that indicates the location and time of the strip start and

end for each trip, the length of the trip in distance and time, and the day on which the trip was performed.

### **Passive GPS Devices**

Two types of devices, shown in Figures 1 and 2, are currently being used: an in-vehicle device that is on whenever the ignition is on, and a wearable device that is on continuously. The wearable device is still under development, but the in-vehicle devices have been used extensively. The primary issue to be addressed is how to present the information contained in the records of these devices so that it is understandable both to the survey respondents and to decision makers. There is also an issue of how to present to analysts and decision makers additional information that is present in the records, but that may not be needed in interacting with respondents.

**Figure 1 In-Vehicle GPS Device**



**Figure2 Wearable GPS**



An example of the type of data available is shown in Figure 3. The data available from these devices is the standard output data from a GPS antenna, consisting of:

- Latitude and longitude in degrees and decimal degrees, with hemispheric (E,W,N,S) designation
- Altitude in metres above sea level
- Heading in degrees from North
- UTC Time
- UTC Date
- Speed in km/h
- HDOP (Horizontal dispersion of precision)
- Satellites in view

### **Data Manipulation**

Certain items of the data require manipulation to make them useful and usable for GIS representation. First, the record must be re-ordered so that it can be recognised by the GIS software as a geographic file. The software being used

for this is Caliper Corporation's TransCAD® software. This requires that each line of the file is uniquely numbered with an ID number, and that the first column after the ID is the longitude, and the second is the latitude. Further, longitude and latitude are to be expressed in millionths of a degree, with negative values indicating south or west, and positive values indicating north or east. The compass quadrant is therefore removed, and replaced by the appropriate positive or negative sign for the latitude and longitude. Next, the UTC date and time are converted to local date and time. At the same time, these records are converted so that they are recognised by the database procedures to be times and dates, so that calculations involving times and dates can be performed.

**Figure 3 Example Output from the GPS Devices**

VALI D	LATITUDE	HEMI- SPHERE	LONGITUDE	HEMI- SPHERE	UTC TIME	UTC DATE	SPEED (KPH)	HEAD ING	ALTI- TUDE	HDOP	SATEL- LITES
V	33.89243	S	151.19162	E	50655	190202	0	0	43	2.4	3
A	33.89237	S	151.19165	E	50658	190202	20	23	43	2.4	3
V	33.89205	S	151.19153	E	50703	190202	0	0	43	2.4	0
A	33.8916	S	151.19128	E	50708	190202	40.2	340	43	2.4	3
V	33.89112	S	151.19105	E	50713	190202	0	0	43	2.4	0
A	33.89077	S	151.19088	E	50718	190202	28	324	43	2.4	3
A	33.8904	S	151.1906	E	50723	190202	35.4	327	43	2.4	3
A	33.88993	S	151.19023	E	50728	190202	45.9	326	43	2.4	3
A	33.88963	S	151.19005	E	50733	190202	19.3	351	43	2.4	3
A	33.8893	S	151.19028	E	50738	190202	39.9	41	44	2.4	4
A	33.88895	S	151.1908	E	50743	190202	45.9	53	42	2.4	4
A	33.88863	S	151.1914	E	50748	190202	45.2	58	42	2.4	4
A	33.88832	S	151.19198	E	50753	190202	45.5	57	43	2.4	4
A	33.88797	S	151.19262	E	50758	190202	55.3	51	43	2.4	4
A	33.88747	S	151.19325	E	50803	190202	57.9	43	42	2.2	4

Some additional computations are needed to enable better visualisation of the results. First, the elapsed time and distance between successive track points are desired. These are used to determine trip lengths in time and distance in subsequent steps. For the purposes of better communication, it is also desirable to add the day of the week to the date and time. Additionally, acceleration or deceleration between successive points may be desired, and are calculated from the speed data.

The most challenging part of the data manipulation is breaking the record into individual trips. To do this, it is first necessary to understand how the data are recorded. For the in-vehicle GeoLogger, we assume that the car's ignition turns the accessory slot on and off. In this case, the GeoLogger is on only when the engine is running. (There are some cars in which the accessory slot is always live. We have excluded vehicles of this type from our surveys, so that we are able to determine times when the engine is off.) For this situation, long elapsed times between two successive points would clearly indicate the end of one trip and the beginning of the next. A situation of this type is shown in Figure 4, which also shows some of the results of data manipulation.

**Figure 4 Partially Manipulated Data Showing a Trip End**

ID	LONGITUDE	LATITUDE		Actual Time	Actual Date	Day	Speed kph	Time in Seconds	Elapsed Time	Distance (m)
485	151.276200	-33.741083	154448	15:44:48	26-Jun-01	3 Tuesday	4.17	56688.0	5.00	5.8
486	151.276100	-33.741133	154453	15:44:53	26-Jun-01	3 Tuesday	9.00	56693.0	5.00	12.5
487	151.275967	-33.741183	154458	15:44:58	26-Jun-01	3 Tuesday	0.00	56698.0	5.00	0.0
488	151.275950	-33.741200	154503	15:45:03	26-Jun-01	3 Tuesday	4.17	56703.0	5.00	5.8
489	151.276000	-33.741200	170332	17:03:32	26-Jun-01	3 Tuesday	0.00	61412.0	4709.00	0.0
490	151.276000	-33.741183	170337	17:03:37	26-Jun-01	3 Tuesday	0.00	61417.0	5.00	0.0
491	151.275983	-33.741167	170342	17:03:42	26-Jun-01	3 Tuesday	0.00	61422.0	5.00	0.0
492	151.276000	-33.741150	170347	17:03:47	26-Jun-01	3 Tuesday	0.00	61427.0	5.00	0.0
493	151.276000	-33.741150	170352	17:03:52	26-Jun-01	3 Tuesday	0.00	61432.0	5.00	0.0
494	151.276000	-33.741150	170357	17:03:57	26-Jun-01	3 Tuesday	0.00	61437.0	5.00	0.0
495	151.276000	-33.741167	170402	17:04:02	26-Jun-01	3 Tuesday	0.00	61442.0	5.00	0.0
496	151.276000	-33.741167	170407	17:04:07	26-Jun-01	3 Tuesday	0.00	61447.0	5.00	0.0
497	151.276000	-33.741183	170412	17:04:12	26-Jun-01	3 Tuesday	0.00	61452.0	5.00	0.0
498	151.276000	-33.741183	170417	17:04:17	26-Jun-01	3 Tuesday	0.00	61457.0	5.00	0.0
499	151.276000	-33.741183	170422	17:04:22	26-Jun-01	3 Tuesday	0.00	61462.0	5.00	0.0
500	151.276000	-33.741183	170428	17:04:28	26-Jun-01	3 Tuesday	0.00	61468.0	6.00	0.0
501	151.276017	-33.741183	170433	17:04:33	26-Jun-01	3 Tuesday	0.00	61473.0	5.00	0.0
502	151.276100	-33.741133	170438	17:04:38	26-Jun-01	3 Tuesday	1.67	61478.0	5.00	2.3
503	151.276200	-33.741117	170443	17:04:43	26-Jun-01	3 Tuesday	3.17	61483.0	5.00	4.4
504	151.276250	-33.741067	170448	17:04:48	26-Jun-01	3 Tuesday	3.17	61488.0	5.00	4.4
505	151.276500	-33.741067	170453	17:04:53	26-Jun-01	3 Tuesday	24.83	61493.0	5.00	34.5
506	151.276800	-33.741100	170458	17:04:58	26-Jun-01	3 Tuesday	15.50	61498.0	5.00	21.5
507	151.276917	-33.740933	170503	17:05:03	26-Jun-01	3 Tuesday	28.83	61503.0	5.00	40.0
508	151.276983	-33.740467	170508	17:05:08	26-Jun-01	3 Tuesday	37.33	61508.0	5.00	51.9

The highlighted line (dark grey) in Figure 4 shows an increment of 4,709 seconds from the previous track point, or an elapsed time of 1 hour 18 minutes and 29 seconds. Clearly, this must be a trip end, occurring between 3:45 p.m. and 5:03 p.m. A further give away that this is a stop is provided by the following rows of data (highlighted in a lighter grey), for which the speed is zero for approximately 1 minute, and no distance is covered. This is most probably due to the vehicle remaining stationary while the driver gets ready to leave, or waits for a passenger to embark. Such situations make for obvious trip ends. The more difficult ones are very short stops, such as may occur when a person fills up the car with petrol, stops to post a letter, or does some similar activity. Even more difficult to detect are stops where the engine is left running, such as at a drive-through fast food outlet, or to drop off or pick up a person.

To determine how to detect such short stops and stops where the engine is not turned off, a number of experimental runs were done by staff, for which the actual activities along the route were known. From an analysis of these situations, it was determined that the optimal rule would be to select any location where there was a break of more than two minutes, during which speed remained zero, and distance travelled was also zero. This selection is not foolproof. There are instances where traffic stops may exceed two minutes, particularly when a vehicle is attempting to make a right turn at an unsignalised

intersection, where the oncoming traffic volume is heavy and traffic is not platooned markedly. Another situation that can be detected by software is one in which there is either no detectable stop, or a very short stop, followed by a 180° change in heading. To detect the change in heading, it is necessary to look further afield than the immediate two or three track points, because it may take some large number of seconds to reverse a vehicle's direction.

### Potential Problems

There are also some potential problems that may arise in the track records. The first of these is signal loss during a trip. This is most often caused by urban canyons, heavy tree canopies, or other problems such as tunnels. The second is the time taken for the device to acquire a signal. From our experiments, we found that the rated time of acquisition of position applies only to a stationary device. The devices in use are rated to acquire signal in 15-45 seconds, and generally succeeded, when stationary, to acquire position within no more than 15 to 20 seconds. However, if the device is immediately in motion when it is turned on, such as if an in-vehicle device is used and a person gets in the car and drives off immediately, a much longer time may be required to acquire position. This results from the vehicle motion, which requires the device to take longer to fix its position, and may also be exacerbated if there are interruptions to the signal, resulting from tall buildings or heavy tree canopies, while the device is attempting to acquire position. In our experiments, we found that in-motion acquisition time depended on how long it had been since the device was last turned on. For elapsed stop times of less than an hour or two, signal acquisition was still relatively fast and generally around 30-60 seconds. However, if a longer time had elapsed since the last use, the position acquisition could become lengthy, and exceeded, in a few cases, one kilometre of travel distance and about 2 minutes or more of time.

Both of these situations require resolution if the track record is to be converted sensibly to a set of trips that can be shown to participants in the survey, and if reliable information is to be gleaned from the records for analysis and presentation to decision makers. The former situation is also more troublesome if a stop occurs within the portion of the record where the signal has been lost. This would occur, for example, with someone who drives from a suburban area into the CBD, stops to transact some business within the CBD, and then resumes travel and emerges in a different part of the CBD boundary then entry. In this case, it is necessary to determine if a stop might have been made, which would constitute the end of one trip and the beginning of the next trip.

### **Requirements for Analysis**

Because the collection of data by GPS devices has the potential to provide very large quantities of data, it is necessary to develop automated procedures to analyse the data, and convert it to readily visualised information for

respondents, analysts, and decision makers. The need is for some rule-based algorithms that can do the following:

- Detect a trip end when the ignition has not been turned off;
- Detect a trip end that occurs when direction is reversed, but there is a very short or non-detectable stop;
- Determine whether a trip ended during a period of signal loss in the middle of a trip; and
- Determine whether there is a loss of signal at the beginning of a trip and repair the data record if there is.

As noted previously, we experimented and found that a rule in which a trip end was assumed to occur if there was 120 seconds or more of elapsed time between two successive in-motion data points worked well to detect trip ends where the ignition was not turned off. Such locations are found by removing from the data any data points where the movement between successive data points is less than the accuracy rating of the GPS device. In our case, the GPS devices in use are rated to have an accuracy of within  $\pm 20$  metres. However, when stationary, the position rarely changes by more than a few metres, with a speed of 0.0 kph, or nearly so. Therefore, by removing those points where the speed is shown to be zero, and there is little change in position, we can detect when there is a stop lasting two minutes or longer, and define that as a probable trip end. Again, by inspection of known short stops, we found that setting a limit on the change of latitude or longitude to  $0.000051^\circ$ , which represents a distance of movement of less than 7.4 metres, we could be fairly certain of detecting a stop. In summary, we flag a trip end if:

- The difference in successive latitude and longitude values is less than  $0.000051$  degrees AND
- The heading is unchanged or is zero AND
- Speed is zero AND
- Elapsed time during which these conditions hold is equal to or greater than 120 seconds.

If there is a break in the record, meaning that the engine was turned off, for between 30 and 120 seconds, this is also defined as a potential trip end.

The third situation that is automatically detected is when the vehicle reverses direction. Reversal is defined as having occurred if the heading changes within 6 track points (30 seconds) by between 178 and 182 degrees. This algorithm, however, may need further experimentation, since it is not clear whether 30 seconds is enough to capture most reversals. Such occurrences also may arise when a driver realises that he or she has taken a wrong turning, or missed a turning, and reverses to find the correct location. It is unlikely that any rule-based algorithm will be able to detect and differentiate this from a situation in which a driver serves a passenger, or performs some other activity and immediately reverses direction.

The signal loss problem is not very serious if the vehicle drives straight through the area of signal loss, because the vehicle is picked up again as soon as the

conditions causing signal loss end. Because there is a substantial change in location, this situation can be distinguished from a stop by estimating the implied speed over the distance between the last point before signal drop out and the first one after the signal is re-acquired. The rule created for this situation is to check the speeds over the ten or twelve track points prior to signal loss, the speed over the ten or twelve track points after signal loss, and estimate the average of these two speeds. The implied speed over the signal loss is determined from the straight line distance between the point of loss and the point of signal re-acquisition, and estimating the average speed for this segment. If this speed is comparable to the average of the speeds before and after signal loss, it is assumed that no stop occurred in the signal loss segment.

If the average speed over the lost segment is much lower than the average speed of the points immediately before and after signal loss, then a potential stop is assumed. However, before completing this step, the length of the stop is estimated by applying the before and after average speed to the distance of the lost signal. The time that would have been required to negotiate that distance at that speed is subtracted from the elapsed time of the signal loss. If the difference is greater than 120 seconds, a trip end is assumed to have occurred. If the difference is less than 120 seconds, then a trip end is assumed not to have occurred.

Work is currently in progress to deal with the antenna warm up problem, and to interpolate time and distance for the initial part of a trip where no position is acquired. The reason for doing this is to be able to provide a better estimate of when the travel began, and to correct the trip length in both distance and time from the GPS device. This is important if the GPS device is being used to estimate emissions, or when presenting the data in a prompted recall survey, so that respondents are presented with more accurate data. If signal acquisition occurs within 100 metres of the real origin of the trip, this is generally not a problem. However, when distances exceed 100 metres, both the time and distance information from the trip may appear incorrect when compared to another trip between the same two geographic locations.

### **Mapping of the Resulting Data**

Once the data have been processed through our programs to manipulate the data and to define trip ends, two steps remain. The first is to present the data in map form, so that various pieces of information can be communicated. The second is to provide a tabular summary of the data. In GIS terms, the track point data, even after processing, is a point layer for a GIS. Thus, a trip is represented by a sequence of dots in a layer. This is useful for analyses that are focused on what is happening moment-by-moment through the trip. For example, colour coding the dots by speed, or acceleration/deceleration allows one to see how the trip is progressing along the route. Figure 5 shows an example of colour coding by speed.



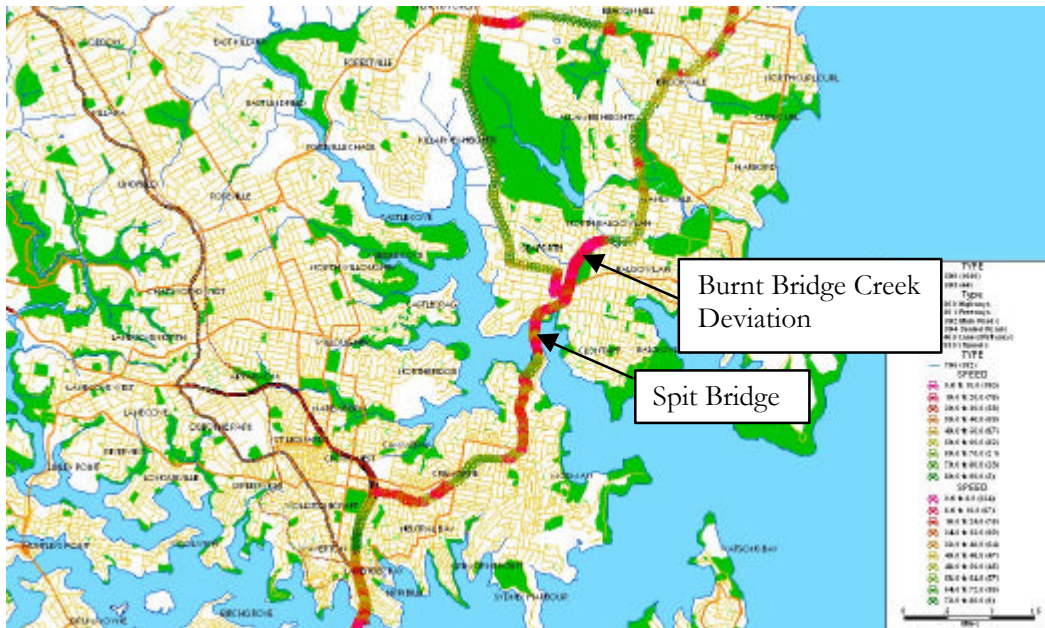
**Figure 5 Point Data Coded to Show Speed**



In Figure 5, the direction of travel is from the top of the picture to the bottom. The faster the speed, the more green (or light tone) is the symbol, while the slower the speed, the more red is the symbol (or dark tone). This clearly shows problems of severe congestion along the roadway north of the bridge. Sporadic and short-lived congestion occurs further south, followed by a fairly long distance at relatively steady speeds. Similarly, a second map (Figure 6), with this and another approach to the bridge at a slightly different time in the morning peak period, shows that there is congestion on all approaches, while south of the bridge, there is indication of increasing congestion later in the peak period.

Using data coded in this way provides a ready mechanism to demonstrate problems to policy and decision makers. The graphical presentation is much more effective than figures relating to amount of congestion. One can, of course, also derive direct graphical data from the GPS track points, as shown in Figure 7. Here, the speed is shown against distance. The spikes in the speed indicate the short sections where a fairly high speed can be obtained, while the extensive distance of low speeds indicates severe congestion for a period.

**Figure 6 Representation of Congestion on Two Routes Approaching the Bridge**



**Figure 7 Representation of Speed Data over Trip Length**

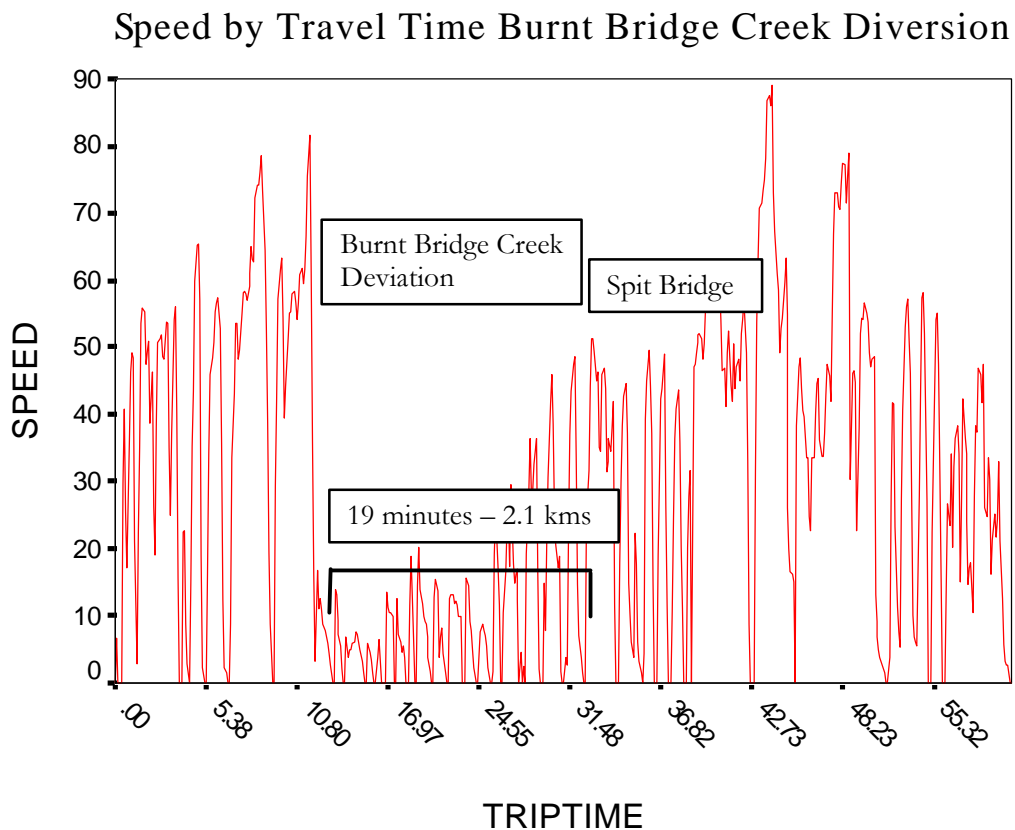


Figure 8 shows acceleration and deceleration on a point-by-point basis, for part of the same trip. Some fairly hard, but brief acceleration and deceleration

incidents can be seen on this map. Most of the points show close to zero acceleration and deceleration, as would be expected. However, there is one hard deceleration followed by an equally hard acceleration in the middle of the map.

**Figure 8 Point Layer Colour Coded for Acceleration and Deceleration**



The next step in the process is to convert the point layers to line layers, and to do this in such a way that the individual trips are offset from each other, to allow one to distinguish individual trips that occur over the same stretch of highway. Figure 9 shows the initial trip layers, colour coded to distinguish one trip from another, but still represented by point data.

Figure 10 shows the result of converting these to line layers and offsetting them from the road centre line, to show multiple trips along the same road. Figure 11 shows the tabular presentation of the same data, that can be used for those who prefer tables to maps.

## Conclusions

The visual displays shown in this paper and that have been produced using the rule-based algorithms discussed in this paper provide effective communication of the extensive data available from the GPS devices. However, it is clear that a comprehensive set of rules are required to permit automated reduction of the huge amounts of data available to parsimonious displays of the data.

**Figure 9 Point Layer, Colour Coded to Trips**



**Figure 10 Point Layers Converted to Line Layers and Offset**



**Figure 11 Tabular Presentation of Travel Recorded by GPS**

## Summary of Travel Activity for John Smith

Trip No.	Start	Origin	End Time	Destination	Travel Time (min)	Distance (Km)	Av. Speed (KPH)	See Map No.	Notes
<i>Saturday 07/07/01</i>									
1	9:11:01	Home	9:17:26	Vumbaca's	6.42	3.18	35	1	
2	9:34:03	Vumbaca's	9:35:53	Coles	1.83	0.76	15	1	
3	10:15:50	Coles	10:21:42	Home	5.87	2.46	18	1	

From an analysis of known trips recorded by GPS and the application of the rules discussed in this paper, we have determined that there is about a five percent error rate in both detecting false trip ends, and in failing to detect real trip ends. This five percent is only partially correctable by manual procedures. Some false trip ends are likely to remain in the final data presented to respondents, and some trip ends will not be detected. However, the prompted recall survey provides a mechanism to correct most of this error.

## **References**

- Guensler, Randall and Jean Wolf, (1999) "Development of a Handheld Electronic Travel Diary for Monitoring Individual Trip Making Behavior", *Transportation Research Board Annual Meeting*.
- Stopher, Peter R. (2001), "Using Passive GPS as a Means to Improve Spatial Travel Data", paper presented to the 8<sup>th</sup> World Congress on Intelligent Transportation Systems, Sydney, Australia, Sept.-Oct., 2001.
- Stopher, Peter and Chester Wilmot, (2000) "Some New Approaches to Designing Household Travel Surveys – Time Use Diaries and GPS", *Transportation Research Board Annual Meeting, Paper # 00-1163*.
- Stopher, Peter R. and Philip Bullock (2001), "Using Passive GPS as a Means to Improve Spatial Travel Data: Further Findings," Paper presented to the Conference of Australian Institutes of Transport Research (CAITR), Monash University, December.
- Wagner, David P., Murakami Elaine and David M. Neumeister, (1998) "Global Positioning System for Personal Travel Surveys", *Federal Highway Administration*.
- Wagner, David 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 1997.
- Wolf, Jean L., Randall Guensler and William Bachman, (2001) "Elimination of the Travel Diary: An Experiment to Derive Trip Purpose from GPS Travel Data", *Transportation Research Board Annual Meeting, Paper #01-3255*.