WENDY BOHTE, KEES MAAT, WILKO QUAK, TU DELFT

A METHOD FOR DERIVING TRIP DESTINATIONS AND MODES FOR GPS-BASED TRAVEL SURVEYS

INTRODUCTION

Travel patterns of individuals are becoming more and more varied in time and space. To understand complex travel patterns, current travel behaviour research increasingly focuses on trip chaining, complete daily and weekly activity patterns and relationships with spatial structure on a detailed level (e.g. Krizek, 2003; Golob, 2000; Boarnet & Sarmiento, 1998; Maat & Timmermans, 2006).

Due to this shift to a more integral approach, it has become essential to use data collection methods that are able to obtain detailed travel behaviour characteristics of all trips an individual makes throughout a day and preferably for several consecutive days. Current travel behaviour research combines in its analyses data on the location of origins and destinations, destination type, trip length, trip duration, departure and arrival times and mode choices. In addition, even more specific data such as exact routes, activities and the people who accompany respondents during their trips are being used. Moreover, the availability of detailed spatial data enables researchers to perform analyses on the effect of spatial structure on travel behaviour at a very detailed level. Paper travel diaries, as frequently used in the past decade, provide much better data than traditional recall surveys (e.g. Stopher & Wilmot, 2000; Arentze et al., 2001). However, the burden of taking these detailed notes all day increases the risks of non-response and reduced accuracy. Moreover, a major drawback of the paper diary method is the difficulties respondents have in determining the exact locations of their visits.

A recent solution for these drawbacks is a collection method on the basis of the Global Positioning System (GPS), which is potentially more accurate and less of a burden on respondents than paper diary methods. Not only can exact locations of trip destinations and travel times be recorded, but also additional characteristics such as exact routes (Wolf, et al., 2003; Forest and Pearson, 2005; Steer Davies Gleave & Geostats, 2003; Ohmori et al., 2005). GPS is a satellite-based navigation system which makes it possible to determine locations with an accuracy of approximately 10 metres. Especially when GPS data are placed in a GIS application for further interpretation, the possibilities of the use of GPS are promising (cf. chapter 4). Moreover, the lower burden on the respondent allows a longer collection period than just a few days. Furthermore, data collection by means of GPS yields an advantage in terms of data processing. The data is immediately available in digital format, thereby avoiding the need for time-consuming data entry and data entry errors.

To date, data collection by means of handheld GPS receivers has only been applied in a few, largely experimental studies. The majority of GPS-based studies have been conducted in the USA, with GPS receivers being placed in cars. However, as the focus in many travel behaviour studies is not only on the car, and as in many countries modes other than the car have a considerable modal share, travel behaviour data should be collected for all different modes. With the introduction of light-weight handheld GPS receivers with an increasingly better reception, a sharp increase has recently been observed in the use of GPS receivers for measuring trips by other travel modes (Steer Davies Gleave & Geostats, 2003; Kochan et al., 2006; cf. chapters 5, 6, 7 and 9). Nevertheless, few studies include the registration of modal choice and destination types due to the fact that in contrast to travel times and distances, these travel characteristics cannot be derived directly from GPS logs and require a more complex derivation method.

This chapter contributes to the improvement of GPS-based travel surveying by introducing a combined method of GPS, GIS and web-based user interaction, which has been applied in large-scale fieldwork in the Netherlands. With over 1000 participants, as far as we know, this is the first time that a GPS-based method that measures travel mode choice as well as the location and type of destinations that are visited has been used on such a large scale. The chapter focuses in particular on the identification of travel modes and destinations, which is still an under-researched issue (see also chapter 8). For the greater context of this study, we refer to Bohte et al (2007).

Our approach concentrates on the issue of deriving and validating the purpose of trip destinations and travel modes, while also allowing reliable multi-day data collection. The method consists of an interpretation process and a validation process. The interpretation process uses spatial data (e.g. railways, shops) and characteristics of the respondents (e.g.

home address, possession of cars) to interpret data from the logs. The travel behaviour data that result from this interpretation round can be adjusted and added to by the respondents in a validation application. The link between both processes is interactive; when new individual characteristics (e.g. the address of a friend's house) are entered by the respondents, these characteristics will be used for further interpretation of the data.

The remainder of this chapter is structured as follows. The following section gives an overview of the advantages and drawbacks of current GPS-based data collection methods that are suitable for measuring choice of travel mode and/or trip destinations. The subsequent section describes the GPS-based method that we developed and in section four the value of our method is evaluated by presenting the results of the fieldwork we recently undertook. The results are compared with results from an internet survey that was carried out at an earlier date and also with the Dutch Travel Survey (DTS) that uses paper diaries. The chapter ends with conclusions on the use of GPS-based methods for the collection of travel behaviour data and a discussion of future possibilities.

LITERATURE REVIEW

Travel researchers are currently experimenting with different GPS-based data collection methods. This review focuses on methods that can be used to derive modal choice and destination types visited. Travel behaviour characteristics such as travel times and distances can be derived almost directly from GPS logs, as a GPS receiver records exact positions and exact times.

A simple, straightforward method for collecting the data that cannot be derived from the GPS logs alone is to ask the respondents directly. Some studies asked respondents to use a paper diary in combination with a GPS receiver, while others asked respondents to enter the data in a GPS-enabled mobile phone (Ohmori et al., 2005) or PDA (Kochan et al., 2006). However, these methods do not solve all the accuracy and burden drawbacks of paper diary methods. Therefore, for deriving modal choice and destination types, GPS data have to be combined with spatial data and respondent characteristics using smart algorithms.

Tsui and Shalaby (2006) deduced the travel mode from the logs by taking the average and maximum speed and the rate of acceleration observed during the trip. In a second method, they also used spatial data, such as public transport routes. Underground trips were identified by examining whether the previous trip ended in the vicinity of a metro station and the beginning of the following trip started near another metro station. Both methods were tested on the GPS logs of nine volunteers in Toronto and both achieved good results, although the method using GIS performed the best. See also Chung and Shalaby (2005).

For deriving destination types from GPS logs, the use of GIS is indispensable. GPS logs provide no information on the kind of location concerned. A few studies describe possibilities for deriving destinations by combining GPS logs with GIS maps, although current literature is limited to studies focusing on car travel (e.g. Wolf et al., 2001). Schönfelder et al. (2002) found on the basis of experience that an important precondition was the availability of detailed maps. Doherty et al. (2006) theoretically described a method that both combines GPS and GIS and enables us to determine travel modes and destination types. In their method, GPS logs are first split into trips, then activities and missing sections are determined via algorithms and subsequently start and end times, travel modes used, and activity locations and the use of space in the vicinity of the activity locations determined. The details derived from this process are then presented to the respondents in tables in an internet questionnaire. The respondents are asked to check the details and among other things, add information relating to the number of travelling companions. It is possible to view the trips on a map in the application should respondents wish to do so. Any adaptations to the data by the respondents are made in the tables.

In addition, we argue that a GPS-based method requires the validation of the data by its respondents by showing them the results derived and asking for validation, as well as corrections and the addition of trips and trip characteristics. Stopher and Collins (2005) have shown the value of an internet recall survey. The development of an interactive web-based system is complicated and as far as we know, no system has yet been designed that works almost perfectly. However, when a working system is developed, the processing of GPS data into trip characteristics should be relatively fast and inexpensive, as no manual data entry is needed. The Internet is a medium that enables the presentation of derived data as interactive maps and tables.

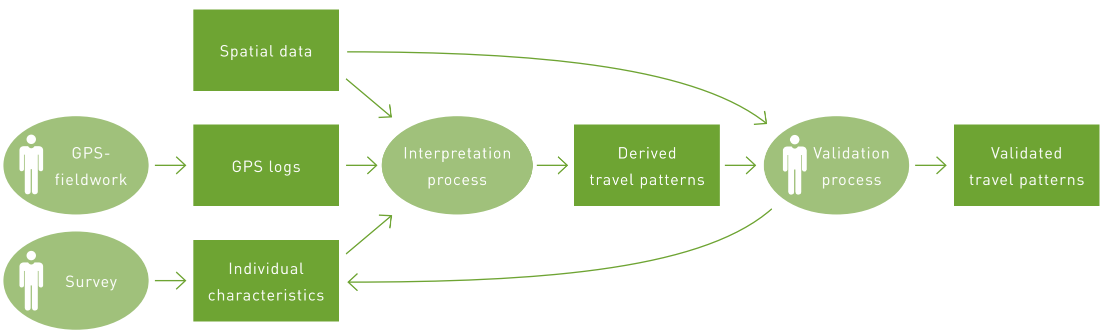
An option that does not seem to be discussed in literature is the use of individual characteristics (household composition, possession of travel modes, home and work addresses, etc.) as input for the algorithms that derive trip characteristics from the GPS logs. A small survey held in advance can yield valuable information and decrease the amount of information that has to be acquired at a later date (cf. chapter 13).

ARCHITECTURE OF THE GPS-BASED SYSTEM

This section describes the architecture of the GPS-based system that identifies modes, location of destinations, destination type and trip distances, times and duration as accurately as possible and with as little a burden on the respondents as possible. Attention is also focussed on validation of the results by the respondents.

The GPS-based system consists of two main processes: an interpretation and a validation process (see **Illustration 10.1**). Three different sources are placed in a spatial database (PostgreSQL/PostGIS): GPS logs created by providing respondents with GPS receivers for one or more days, individual characteristics of the respondents collected by a survey and spatial structure data. The spatial data consists of various sources, such as road networks, railways, public transport stops, shops and other services. Moreover, all addresses of the home and work locations are geocoded. Individual characteristics include gender, availability of a driving licence and a car, an annual season ticket for public transport, work status, and so on. During the interpretation process, a collection of scripts that include various algorithms runs on top of the database to combine and interpret the different data sources. When trip characteristics are reconstructed, as far as possible, they are forwarded to the validation process. The main part of the validation process consists of a web application (using Flash) that obtains its data from the spatial database. In the user interface of the web application, the derived data are presented to the respondents in maps and tables. The respondents are asked to use this validation application to correct and add to the derived trip characteristics. The link with the interpretation process is interactive. When the input of the respondents provides new individual information such as the addresses of friends and family or in the event that new trips are added, this information will be reused for further interpretations.

Illustration 10.1 - Architecture of the GPS-Based System



The interpretation process

In the interpretation process, characteristics of the trips made by the respondents are derived from the raw data in the GPS logs. In addition to the GPS logs, two other sources are used as input, namely spatial structure data and individual characteristics of the respondents. To derive trip information from the GPS logs, the following steps are used (cf. chapter 7): (1) filtering track points from the GPS log; (2) splitting the GPS log into trips; (3) deriving modal choice and destination types for the trips; and (4) post-processing. The parameters used in the algorithms are partly derived from other travel behaviour data, partly from validated data in this data set, and partly from logic thinking.

Filtering track points

Before the GPS log is analyzed, a number of cleaning operations are performed. Some of the filtering is carried out to reduce the amount of data and some to remove measurement errors from the GPS log. A GPS log consists of a series of points, so-called 'track points'. Each track point is described in the log on the basis of x and y coordinates and a time stamp. Data is also recorded on the quality of the measurement and the speed of the GPS. These data are used to remove track points with low quality measurements or an unrealistically high speed.

Special care needs to be taken to detect whether the GPS is in rest. Due to measurement errors, a GPS that is in rest does not return to the same position each time, but acts as moving slowly. In most cases, filtering out these points is straightforward, based on the typical movement behaviour and the known accuracy of the GPS device. However, when the receiver is indoors, it only sees a small fragment of sky and incorrectly produces tracks even when the receiver it is not moving.

Splitting tracks into trips

After the most unreliable track points have been removed, the log is divided into actual trips made. The parts of the log that must be considered for separate trips is deduced from the respondent's rest periods by means of the loss of satellite reception in buildings. Although dividing up a trip based on changes in the respondent's speed seems to be a good method, in practice, a single trip appears to include many speed changes. When the GPS logs indicate that someone remained at a certain location for at least three minutes, the location in question is classed as the destination of the previous trip. Wolf et al. (2001) compared the use of different thresholds and found that a three minute threshold resulted in the best prediction. As activity-based research increasingly analyses the complexity of multi-modal travel, multi-modal trips are divided into stages (a stage is a part of a trip, between mode changes), meaning that railway stations and bus stops are included as separate destinations. If a respondent rides his bicycle to the train station, takes the train and then walks to work from the station, this is classified as three trips.

When the log is split up in trips, this automatically shows the times and locations at which trips were started and completed. This is because if it is known which track point was the first to be placed at a location where someone remained for an extended period of time, it is also known what time the person arrived there and the x and y coordinates of the point in question. It is possible to deduce the departure time from the first track point placed at some distance from the location where the person remained for an extended period of time.

One of the shortcomings of the current generation of GPS receivers is the fact that after the receiver has been turned off or in the event of lost reception, it often takes a while (usually no more than 30 seconds) before a GPS receiver has found enough satellites and receives enough signals to be able to determine its location (this is also referred to as getting a 'fix'). This often means that the departure from a location is not logged. When the final track point of one trip and the start point of the next trip (these should always be the same) are far apart, the end point of the previous trip is chosen as the most reliable.

Determining travel modes

To determine the travel mode, different data are used. By deriving the average and nearly-maximum speed, in most cases, it can firstly be estimated whether the person walked, cycled or drove a car. Nearly-maximum speed is used to avoid using track points that were registered when satellite reception was not optimal, causing misplaced registration and the wrong estimation of speeds. When the average speed is below 10 km/h and the nearly-maximum speed is below 14 km/h, it is estimated that a person walked. An average speed between 10 and 25 km/h and a nearly-maximum speed between 14 and 45 km/h leads will lead to the estimate being made that a person cycled. When the average trip speed is between 25 and 200 km/h, it is estimated that a person travelled by car.

To determine whether a trip was completed by train instead of by car, a link to GIS data is required, as the speeds of these travel modes may be very similar. By checking whether at least one third of all track points of a trip lie within 50 meters of a railway and by comparing the coordinates of the starting and ending point of a trip with the locations of train stations via underlying GIS maps, it is determined whether a trip is likely to have been made by rail and that the respondent therefore probably took the train.

These algorithms will not always make the correct distinction between different travel modes. For example, when people are running, the estimated travel mode will probably be 'bicycle'. Furthermore, no algorithms have been developed to detect the use of scooters or other less common modes. When in future applications acceleration and speed changes within a trip are also analysed, it will be possible to increase the number of correctly predicted travel modes.

Determining category of the destination

GIS data are also used to estimate the category that the visited destinations fall within. All the potential destinations are classified in 13 categories, such as 'home', 'work', 'friends/family' and 'cultural'. Destinations of facilities are derived from GIS maps listing Points of Interest (POIs) such as the x and y coordinates of schools and other facilities. If a trip ends within a radius of approximately 50 metres from a known location, it is assumed that this is the location being visited. Because shopping centres and railway stations can be spread out over a large area, they are not represented by points, but their whole outline is drawn in GIS maps.

Due to the fact that the home addresses of respondents are already known (as this is where the GPS receivers were delivered to and collected from), these can be entered into the database. The postal codes of the work locations of the respondents are also translated into x and y coordinates. Because the work location will be frequently visited, the work address is asked for in the survey carried out in advance. If a respondent's trip ends within 100 meters from his home or work address, it will be assumed that he went home or to work.

If it is not possible to filter out a possible destination category on the basis of the underlying GIS maps and the known data on the respondent, the destination category will be listed as 'unknown' until the respondent has indicated in the internet application what category the visited destination is in.

The algorithms that are used to determine the destination types of the trips are relatively straightforward. In the near future, more complex algorithms will be constructed to enable us to derive increasingly accurate trip data from GPS logs. For example, when there are several known locations within a 50-metre radius of the end of a trip, it should be possible to not automatically select the nearest location, but to select the destination based on (individual) values allocated to the possible locations. This means that if a supermarket is located a little further away from the end of a log than a police station, it will, in the first instance, still be assumed that the respondent visited the supermarket. It is also possible to allocate individual values to the respondents for each type of destination included in the probability calculation. This means, for instance, that people with school-age children are more likely to be assumed to be visiting a school than people without children.

Post processing the data

After analysing individual trips, the generated trip diaries are post-processed and analyses are performed that consider more trips together. One example is merging small and short shopping trips within one shopping area, while another example is that if a trip ends near one railway station and reappears later near another, it is assumed that a train trip was made.

The validation process

After the data have been interpreted, they are presented to the respondents in the web-based user interface of the web application. The respondents are asked to check their travel behaviour data and to make corrections and/or additions where necessary.

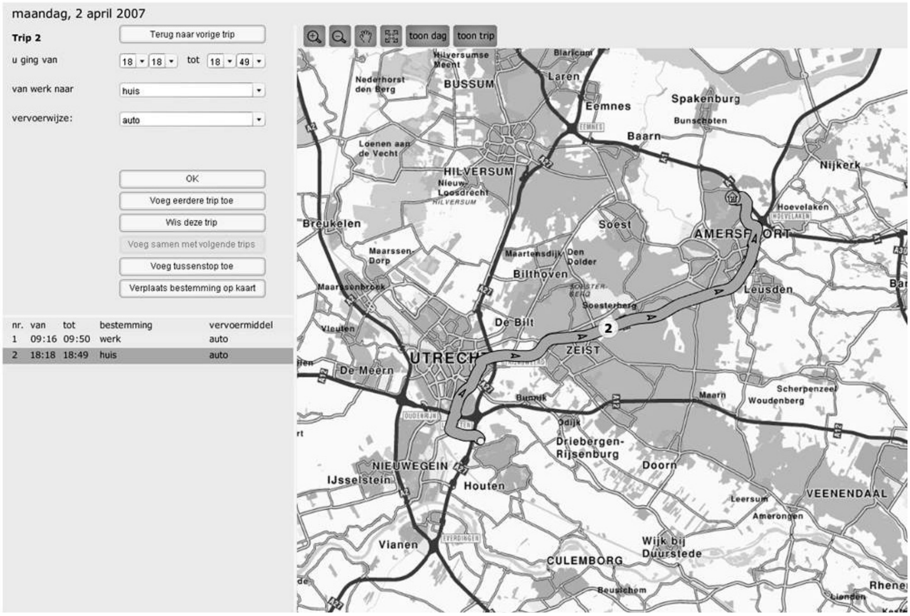
The decision to use the internet was taken for several reasons. Firstly, an important advantage of using the internet for a recall survey is that when entering information, the user interacts directly with the database and information provided can then immediately be used to better approximate the respondents' trips at a later date. In contrast to telephone surveys, respondents can answer the survey whenever they wish. Another important advantage is the possibility of easily showing respondents their travel behaviour depicted on a zoomable map and allowing them to move locations in their trips on the map. An experiment carried out by Stopher and Collins (2005) showed that respondents where able to indicate missed trips and destinations on maps depicting the routes they had taken. Finally, exchanging the data via the internet also eliminates the associated costs of delivery and data entry.

An important disadvantage could be the risk of generating a selective response. To be able to use the application, respondents have to have at least some experience with computer programmes and the computer and internet connection they use must not be too slow. These preconditions may lead to an under-representation of older and lower-educated people. However, in recent years, computer and internet accessibility have grown enormously and it can be expected that selectivity caused by lack of computer skills and respondents without access to a computer or with no or slow internet connections will decrease rapidly. Moreover, as surveys are generally over-represented by older people, this may be balanced by using computer-based surveys.

Interface

The interface of the validation application consists of a map and a table depicting all trips a respondent visited in a day as derived from the spatial database as well as a form that respondents can use to adjust the depicted trip characteristics (**see Illustration 10.2**). The map can be panned dynamically and zoomed in and out, with a changing level of detail while zooming. This is realized by using a map database with maps on six successive spatial scales. The maps, provided by Falk, include street names and points of interest on the lowest spatial level. The trips on the map are linked dynamically with a table that lists each trip with the originating location, the departure time, the travel mode used, the arrival time and the location of the destination.

Illustration 10.2
A screenshot of the user interface of the web application



Adjusting trips and trip characteristics

When the algorithms used in the interpretation process were not able to determine or wrongly determined the starting or ending times of a trip, or when the travel mode used or the category of the destination visited could not be derived correctly, respondents can adjust these trip characteristics by choosing an (other) option from the dropdown menus. When the respondent amends any data, the amendments concerned are passed on to the spatial database and are used for further interpretation, after which the new data are used to better estimate the trips of the respondent in the days after the adjustment was made (self-learning).

When the GPS log is not correctly divided into trips, respondents can merge or split trips. Respondents can also change the location of the destination of a trip if the GPS receiver stopped logging before the end of a trip or even add a whole trip if the receiver was forgotten, the battery was not charged or the receiver failed to acquire reception of satellites.

A CASE STUDY

The GPS-based data collection method we developed was applied for the collection of over one thousand one-week travel behaviour patterns. This data collection was part of a larger research project (Bohte & Maat, 2007), but the data can be used for a broad range of travel behaviour research (cf. chapter 11).

In this section, the results and evaluation of this case study are used to evaluate the method. The section starts with a description of the study area, the respondent sample and the fieldwork organisation. The interpretation process is then evaluated by discussing the percentages of modes and destination types that had to be changed or added by the respondents because they were not able to be correctly derived from the interpretation model. Certain results (trip frequencies and trip chains) are subsequently compared to data from the Dutch Travel Survey (AVV-MON, 2006; the government's national survey, a one-day paper recall survey). Finally, some of the results from the evaluation survey held among the respondents are presented.

Description of the case study

The fieldwork took place in the first half of 2007 among a sample of residents of Amersfoort (137,000 inhabitants), Veenendaal (61,000 inhabitants) and Zeewolde (19,000 inhabitants) – three municipalities in the centre of the Netherlands. Due to the facts that our research focuses on residential choice, most people in the Netherlands choosing to rent a house have limited options and availability is low and distribution (partly) regulated, we restricted our study to home owners.

The respondents were recruited from the respondents of an internet survey held at the end of 2005. In total, 1200 respondents participated in the GPS fieldwork, 1104 of whom completed the entire project. The respondents carried a hand-held receiver for one week. The receiver we used was an adjusted Amaryllo Trip Tracker (http://www.amaryllo.com, acc. April 7, 2007; see illustration 10.3) that was programmed to log a track point every 6 seconds. The battery lasts approximately 16 hours which was accomplished by disabling all unnecessary functions. One day after the receivers had been collected again, the GPS logs from the receivers were entered into the interpretation process to be combined with spatial structure data (location of shops,

Illustration 10.3
The Amaryllo tracker

railway stations, railways, schools and cultural services) and data on individual characteristics (car ownership, home and work addresses) that were collected beforehand. The same day, the respondents received an e-mail with a link to the user interface of the validation application, enabling them to validate and add to the registered and derived data.

Corrections by the Respondents

By determining what percentage of all travel modes and destination types had to be provided or changed by the respondents in the validation application, we can estimate the success of the interpretation process. In almost three-quarters of all cases, the travel mode proves to have been estimated correctly in the interpretation process. Car use was deduced correctly most often (75% of all trips), followed by cycling (72%) and walking (68%) respectively. Trips that ended at home were placed in the correct destination category most frequently due to the fact that the home location was already known (74% of all trips that ended at home). Visits to shops were estimated correctly in 35% of cases. Although the home addresses of friends and family were not known in advance, 11% of these visits were nonetheless recognised as such in the database. These latter estimates may contribute towards the 'learning' function, which remembers the category of the destinations that were visited earlier in the week.

Trip Characteristics

By comparing travel characteristics from the GPS dataset with data from the Travel Survey (AVV-MON, 2006; DTS) it can be estimated whether the GPS-based method produced reasonable results. Illustration 10.4 shows that the destination and mode shares in both data sets are quite similar, with the absolute numbers being higher in the GPS dataset. Nonetheless, the number of tours per day is almost equal in both datasets (GPS: 1.56 and DTS: 1.61). The main difference lies in the number of trips per tour. The respondents in the GPS sample have an average of 2.9 trips per tour while the respondents from the DTS survey demonstrate an average of 2.3 trips per tour. On one hand, it may be the case that the GPS method measures more trips than the traditional method. It can be expected that in a paper recall survey such as the DTS, small stops on a tour such as picking up a child or a short visit to a shop on the way back from work may be forgotten. On the other hand, we realize that the higher number of trips may be partly caused by trips that are incorrectly divided into more than one trip due to the fact that the reception of satellite signals failed for a while and respondents failed to merge these trip parts in the validation application later on. Further analyses of the GPS dataset are necessary to be able to determine more accurately the cause of this relatively high number of trips per tour.

Illustration 10.4
Average Number of Trips per Mode per Day and per Destination per Day

	GPS-based method (1,104 respondents, 7,395 days)		DTS recall survey (40,208 respondents/days)	
Destination	**Mean**	**Share**	**Mean***	**Share***
Work	0.84	18%	0.60	16%
Study	0.03	1%	0.02	1%
Shop	0.60	13%	0.42	11%
Social Visit	0.29	6%	0.26	7%
Recreation	0.47	10%	0.43	11%
Home	1.56	34%	1.61	42%
Other	0.77	17%	0.47	12%
Travel Mode	**Mean**	**Share**	**Mean***	**Share***
Car	2.44	54%	2.01	53%
Train	0.10	2%	0.09	2%
Bus/tram/metro	0.04	1%	0.09	2%
Bicycle	1.17	26%	0.81	21%
Foot	0.74	16%	0.75	20%
Other	0.06	1%	0.06	2%
Total Number Of Trips	**4.55**	**100%**	**3.80**	**100%**

* weighted to match age and education level of the GPS dataset

Results from the evaluation survey

After the respondents finished the validation of the week they participated in the fieldwork, they were asked to fill in an evaluation survey. Questions focused on their experience of the week they carried the GPS receiver and their evaluation of the user interface of the validation application.

The results of the evaluation survey showed that only 1% of the respondents found it a considerable nuisance to continuously carry the receiver, 11% found it somewhat a nuisance and 88% didn't mind at all. However, remembering to carry the receiver was found to be a problem by a large proportion of people (over 40%). A quarter of all respondents did in fact forget their receiver on one or more occasions. The vast majority of respondents (91%) did not consider the fact that the receiver had to be charged in a power point to be a problem.

Many respondents found checking and updating their trips fairly difficult. Nearly 25% said that they found following the programme quite difficult and 40% somewhat difficult. One of the reasons given was that there are still a great deal of gaps in the GPS logs. This was especially the case for respondents with a car with heat-resistant windscreens and respondents who frequently travel by train. People with little computer experience or an older computer experienced a relatively high degree of difficulty with the application. Some respondents found it difficult to remember the trips they had made in the initial days of the fieldwork more than a week later. Nonetheless, a third of all respondents were able to go through the entire validation process in fifteen minutes or less, while two-thirds needed half an hour. Although Doherty et al. (2006) assume that respondents will have difficulty interpreting the maps, this is not supported by the evaluation. Due to the frequent use of route planners on the internet and in-car navigation systems, it is not unreasonable to expect that many people already have some experience using digital maps.

CONCLUSION AND DISCUSSION

This chapter describes a GPS-based data collection method which we have developed for the collection of travel behaviour of individuals. We evaluated the system by applying it in large-scale data collection. The main aim of the development of this method was to build a system that can be used to collect travel times and distances, modal choice and destination types as accurately as possible, and which will also place a low burden on the respondents. The experience and evaluation of the use of other GPS-based methods was used as a starting point for the development of our method. Compared to other methods, this system has a stronger focus on validation by the respondents.

The GPS-based system consists of an interpretation process and a validation process. Three data sources are combined in the interpretation process: GPS logs, individual characteristics of the respondents and spatial structure data. When in the interpretation process trip characteristics are reconstructed, as far as possible they are passed on to the validation process. In the user interface of the validation application, the derived data are presented to the respondents in maps and tables. Here, the respondents can adjust the derived trip characteristics. The link with the interpretation process is interactive and new information delivered by the respondents is used for further interpretations.

The evaluation of the case study was satisfying. Due to the fact that GPS receivers can register exact location coordinates, the location of destinations can be determined with an accuracy that cannot be approached by traditional methods. Furthermore, because the method is able to derive substantial trip data before the respondents go through the validation application and also adds extra information while they are validating the data, the burden on the respondents is reasonably low. The data collected by means of the survey conducted beforehand proved to be very useful as input for the algorithms in the interpretation process. The participants did not experience carrying and charging the GPS receiver as being a nuisance and were enthusiastic about viewing their trips in the maps of the validation application. The majority of respondents were able to go through the validation application in a reasonably short period of time. This meant that the 7-day travel behaviour was able to be collected from more than a 1000 respondents. The comparison with data from the national travel survey showed that the GPS-based method was able to record a larger number of trips. This indicates that fewer trips were missed.

However, the evaluation also showed that the method has room for improvement. Firstly, the method is dependent on the quality of the GPS receivers used. The receiver that was used in the case study presented worked poorly inside trains, which meant that many train trips were missing. People had to remember to recharge their receivers every night. As a substantial part of the technical drawbacks such as the battery life and the size and weight of receivers may be largely solved or at least improved within the next few years, it is important that if possible, new research projects use state-of the-art GPS receivers.

Secondly, although the majority of the respondents required 30 minutes or less to go through the validation application, some did struggle a lot with it, especially people with very few computer skills, old computers and complicated travel behaviour. In addition, because of the poor reception of satellites inside trains, frequent train travellers had to add many trips, potentially meaning a lot of extra work. An extension of the algorithms used in the interpretation to derive more data automatically would mean a further decrease of the complexity of validating the results for the respondents. Progress can be expected to be made in a survey by asking the right questions. Moreover, algorithms could be constructed that compare trips of the same respondents on different days. Since a large part of people's travel behaviour is based on routines, gaps in one day can perhaps be repaired with information from other days, even in real-time when the respondents are using the validation application.

In summary, we can conclude that, at present, both GPS and GIS are starting to make a significant contribution to collecting data on travel behaviour of individuals. The system we proposed solves some of the previous shortcomings and has proved to perform well. However, if the current (technological) developments continue, and methods for collecting the necessary additional information such as the method described in this chapter are developed further, it is likely that data collection by means of paper diaries will disappear entirely in the near future, and will be replaced by methods that collect data with the aid of GPS and related technologies.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the financial assistance provided by the Delft University of Technology via the Delft Centre for Sustainable Urban Areas and by the Dutch government via the Habiforum Programme for Innovative Land Use. We furthermore owe much gratitude to Falk for providing outstanding digital maps of the Netherlands.

REFERENCES

- Arentze, T., Dijst, M., Dugundji, E., Chang-Hyeon, J., Kapoen, L., Krygsman, S., Maat, K. and Timmermans, H. (2001) New Activity Diary Format: Design and Limited Empirical Evidence, *Transportation Research Record: Journal of the Transportation Research Board*, No. 1768, TRB, National Research Council, Washington D.C., pp. 79-88.
- AVV-MON (2006) *Mobiliteitsonderzoek Nederland*, Rijkswaterstaat Adviesdienst Verkeer en Vervoer, Rotterdam.
- Boarnet, M.G. and Sarmiento, S. (1998) Can land use Policy Really Affect Travel Behaviour? A Study of the link between Non-Work Travel and Land Use Characteristics, *Urban Studies* 35(7): 1155-1169.
- Bohte, W., Maat, K. and van Wee, B. (2007) *Residential self-selection, the effect of travel-related attitudes and lifestyle orientation on residential location choice; evidence from the Netherlands.* Paper presented at the 11th World Conference on Transport Research, Berkeley.
- Chung, E-H. and Shalaby A. (2005) A Trip Reconstruction tool for GPS-based Personal Travel Surveys. *Transportation Planning and Technology*, 28(5): 381-401.
- Doherty, S., Papinski, D. and Lee-Gosselin, M. (2006) *An Internet-Based Prompted Recall Diary with GPS Activity Trip Detection: System Design*. Presented at 85th Annual Meeting of the Transportation Research Board, Washington, D.C..
- Forest, T.L. and Pearson, D.F. (2005) *A comparison of trip determination methods in GPS-enhanced household travel surveys*. Presented at 84th Annual Meeting of the Transportation Research Board, Washington, D.C..
- Golob, T.F. (2000) A Simultaneous Model of Household Activity Participation and Trip Chain Generation, *Transportation Research B,* 34, pp. 355-376.
- Kochan, B., Bellemans, T., Janssens, D. and G. Wets. (2006) *Dynamic Activity Travel Diary Data Collection, Using a GPS-Enabled Personal Digital Assistant.* Presented at the Innovations in Travel Modelling Conference, Austin, 2006.
- Krizek, K.J. (2003) Neighborhood Services, Trip Purpose, and Tour-Based Travel, *Transportation,* 30, pp. 387-410.
- Maat, K. and Timmermans, H. (2006) Influence of Land Use on Tour Complexity, a Dutch case, *Transportation Research Record: Journal of the Transportation Research Board*, No. 1977, TRB, National Research Council, Washington D.C., pp. 234–241.
- Ohmori, N., Nakazato, M. and Harata, N. (2005) GPS Mobile Phone-Based Activity Diary Survey. *Proceedings of the Eastern Asia Society for Transportation Studies*, 5, pp. 1104-1115.
- Steer Davies Gleave and Geostats (2003) *The Use of GPS to Improve Travel Data.* Study report. Prepared for the DTLR New Horizons Programme, submitted to the London Department for Transport, London.
- Schönfelder, S., Axhausen, K.W., Antille, N. and Bierlaire, M. (2002) Exploring the Potentials of Automatically Collected GPS Data for Travel Behaviour Analysis - a Swedish Data Source. In: J. Möltgen and A. Wytzisk, eds, *GI-Technologien für Verkehr und Logistik*, IfGIprints, 13, Institut für Geoinformatik, Universität Münster, Münster, pp.155-179.
- Stopher, P.R. and Collins, A. (2005) *Conducting a GPS Prompted Recall Survey over the Internet*. Presented at 84th Annual Meeting of the Transportation Research Board, Washington, D.C..
- Stopher, P.R. and Wilmot, C.G. (2000) *Some new approaches to designing household travel surveys – time use diaries and GPS*. Presented at the 79th Annual Meeting of the Transportation Research Board, Washington, D.C..
- Tsui, A. and Shalaby, A. (2006) Enhanced System for Link and Mode identification for Personal Travel Surveys Based on Global Positioning Systems, *Transportation Research Record: Journal of the Transportation Research Board*, No. 1972, TRB, National Research Council, Washington D.C..

- Wolf, J., Guensler, R. and Bachman, W. (2001) *Elimination of the Travel Diary: an Experiment to Derive Trip Purpose from GPS Travel Data*. Presented at 80th Annual Meeting of the Transportation Research Board, Washington, D.C..
- Wolf, J., Oliviera, M. and Thompson, M. (2003) Impact of Underreporting on Mileage and Travel Time Estimates: Results from Global Positioning System-enhanced Household Travel Survey, *Transportation Research Record: Journal of the Transportation Research Board*, No. 1854, TRB, National Research Council, Washington D.C., pp. 189-198.