



HAL
open science

GPS and travel diary: Two recordings of the same mobility

Thanh Tu Nguyen, Jimmy Armoogum, Jean Loup Madre, Thi Huong Thao
Pham

► **To cite this version:**

Thanh Tu Nguyen, Jimmy Armoogum, Jean Loup Madre, Thi Huong Thao Pham. GPS and travel diary: Two recordings of the same mobility. ISCTSC, 11th International Conference on Transport Survey Methods, Sep 2017, Esterel, Canada. 13p. hal-01588420

HAL Id: hal-01588420

<https://hal.science/hal-01588420v1>

Submitted on 15 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

11th International Conference on Transport Survey Methods - ISCTSC 11th 24th-29th September
2017 Esterel Quebec Canada

GPS and travel diary: Two recordings of the same mobility

Thanh Tu Nguyen^a, Jimmy Armoogum^b, Jean-Loup Madre^b, Thi Huong Thao Pham^b

^a*Faculty of Transport Economics, University of Transport and Communications, Hanoi, Vietnam.*

^b*University Paris Est - Department of Dynamic Economics Socials of Transport (DEST), French institute of science and technology for transport, spatial planning, development and networks (IFSTTAR).*

Abstract

GPS-based data collection methods have become particularly popular in travel behavior research, mainly because of the worldwide coverage and the accuracy of the GPS system. The main objective of this paper is compare the descriptions of mobility obtained by two methods survey reported and GPS recorded in the same days. This study shows that the GPS survey can be used successfully to complete the conventional transport surveys, but it is still too early to predict the complete substitution of conventional survey by the GPS mobility survey.

1. Introduction

Travel survey methods based on new technologies have evolved in the past few decades, shifting from limited experiments to large-scale travel surveys. GPS-based data collection methods have become particularly popular in travel behavior research, mainly because of the worldwide coverage and the accuracy of the GPS system.

A challenge in the GPS data post-processing is the development of methods to fill GPS missing data and to reconstitute automatically continuous sequence, both in space and time. The results obtained from the post processing software are calibrated by comparison with conventional methods, in particular a few questions when getting back the GPS device, providing for a few trips, taken at random, additional information on the reliability of the device and on more detailed characteristics (mode, purpose and accompanying persons).

The primary objective of this paper is to compare survey-reported trips and GPS-recorded trips for the same individuals to identify the reasons of differences between two sources of data. It maybe from the household, individual or travel reasons. Based on these results, we could assess the contributions and challenges of data collection using GPS devices in long-term travel survey. Besides, thanks to a better understanding of the factors influencing these differences, the travel survey data collected by GPS and also by interview could be improved.

To give full information, first this paper introduce some statistics about both data sources. In which, the percentage of unmatched trip by different characteristics are investigated. It means that all the trips recorded by GPS but not reported by respondents and the reported trips with no GPS trace are concerned in this step. Further, these unmatched trips are modeled using a logit model to estimate the variables that affect the difference between the survey-reported and GPS-recorded trips, without making any a priori assumptions about the accuracy of the data from each source. This finding could be useful in further researches for improving the quality of travel data collected

by GPS or by diaries. This paper analyses de GPS travel data from 2 sources: the French National Household Travel Survey (FNHTS) in 2007/08 and the travel survey test by GPS in Hanoi in 2010/11 using the same GPS data logger.

This study has two limitations. First, this paper accepts the self-selection bias and the representativeness of the sample caused by the limited acceptability by some group of respondents. Second, the GPS-recorded trips are being processed by algorithms that were developed by IFFSTTAR/DEST (Yuan, 2010) and we don't know whether similar studies using different algorithms would find similar results.

The remainder of this paper is structured in 4 parts. The next one provides a summary of using GPS in travel survey. Section 3 introduces an overview of the GPS pilot surveys in FNHTS and in Hanoi city. Section 4 presents the factors influencing the difference between two survey methods. The final section summarizes the important findings and recommends specific improvements for the capture of travel behavior by both diary and GPS survey methods.

2. GPS technology in household travel surveys

In recent years, GPS technology has been increasingly used in travel survey research to evaluate and improve the travel data reported by diary. The travel survey by GPS was conducted for the first time in 1997 in Austin - Texas (Casas and Arce, 1999) and since then there are several studies on the use of this technology in travel survey. GPS data show firstly the rate of missing trips obtained from conventional method (by CATI, CAPI, face-to-face or self-administrated). This is a significant shortfall considering that household travel surveys are used for transport modeling and planning. From data of six surveys using CATI and GPS in the US from 2001 to 2004, Wolf (2006) has shown that the rate of missing trips can be from 11% to 35%. However, the first experiment was only in car trips, which facilitated the performance of GPS at a time when the capacity of batteries was very limited. In the research of Forrest and Pearson (2005), they showed that in a CATI household travel survey in Laredo, Texas only 44 percent of trips recorded by a GPS were able to be successfully matched with a trip reported by CATI. By combining GPS and a prompted recall survey, Stopher et al. (2007) found that the rate of non-reported trips was only 7.4% but they ignored all the trips non-recorded in the GPS for some reasons (forgetting to wear, no satellite signal, etc.).

Besides the comparison between travel data by GPS and diaries, many studies tries to identify the key factors characterizing the under reported trips. Zmud and Wolf (2003) identified key demographic factors that may contribute to misreporting of trips. It is particularly the case for individuals younger than 25 years old with low income. Household size and household income are also significant factors of trip reporting accuracy. Bricka and Bhat (2006) have found that young adults (less than 30 years of age), men, individuals with less than high school education, unemployed individuals, individuals working in clerical and manufacturing professions, workers employed at residential, industrial, and medical land-uses, and individuals in nuclear families are all more likely to under-report trips than other respondents. Houston et al. (2014) have similar findings in their research. They have found that participants who were older, had a lower education, lower household income, or were employed, were associated with trip under reporting. In another research, Bricka et al. (2012) divided the trips into work and non-work for identifying the factors influencing difference in survey-reported and GPS recorded. Education levels, employment characteristics (full time or part time) and household income are the main significant variables of their model.

In addition to demographics, travel behavior characteristics have influenced the trip under reporting level, Wolf (2000) claims that short duration trips are often forgotten or omitted. The travel purposes and the number of trips undertaken on the survey day also have impacts on trip under-reporting, Forrest and Pearson (2005) found that home-based non-work and non-home-based trips are more significant for trip under-reporting in Laredo, Texas. Bricka and Bhat (2006) and Stopher et al. (2007) found that individuals with high mobility level tend to underreport their travel. Besides, trips with shorter time duration and trip distance for visits and picking up or dropping off a passenger were more likely to be under-reported.

Besides trip under reporting, the situation where the GPS unit recorded fewer trips than that reported by respondent, have been studied by some researchers. Wolf et al. (2003) looked for the reasons why the GPS trips were missed in comparison with the trip diaries. They found that the lack of power for GPS, cold-start, misrecording due to the loss of signal are the main reasons. Stopher and Shen (2011) found that two most probable reasons for the trip not being recorded in the GPS but being reported in the survey data, are the forgetfulness of respondents and the GPS device problems. The trips undertaken at early morning or late evening, are easier to be under-recording by GPS units. More details, (Bricka et al., 2012) have examined the differences in measures of intensity of travel

between survey and GPS trips for two trip purposes (work and non work purposes). They found that the GPS survey should be considered as the data collection method for the younger, more technology experiences and have high travel propensities but for the elderly and more leisurely travelers, the traditional survey method is recommended.

3. Data information and descriptive

3.1. France - GPS pilot survey 2007/08

Once per decade, the Ministry of Transport and the National Institute of Statistics use to conducting a National Household Travel Survey with the scientific support of INRETS (now IFSTTAR). It is an important source providing the most consistent data on mobility: daily mobility and long distance travel for households living in France. From a national random sample taken from the census, a sub-sample of approximately 750 accepted to participate to the GPS pilot survey. The FNHTS has been running on a full year (from May 1st 2007 to April 30 2008) and 450 interviewers were selected for a GPS pilot test with 170 GPS receivers.

The GPS travel survey protocol as follows: At the first visit, the interviewer asked the respondent if they accept for the GPS travel survey. If they accept and if a GPS device was available, it was given to the respondent. He/she has been asked to wear the GPS during 1 week. At the second visit, the interviewer retrieved the GPS and launched the application for downloading all data from GPS to laptop by bluetooth communication. The interviewer looked up the data and conducted a complementary interview (CAPI-GPS). If there are days without GPS recording, the first question is to understand the reasons why there are no GPS data (GPS receiver forgotten in the office/at home; battery problem, etc). The objective of this complementary questionnaire is to know if the GPS has recorded all trips made and to collect information on travel behavior that GPS cannot record (mode, purpose and accompanying persons).

3.2. Vietnam - Test GPS for travel data collection in Hanoi 2010/11

In the study of (Thanh Tu NGUYEN, 2013), the travel survey by GPS and trip diaries was conducted for these reasons:

- Comparison of travel behavior observed through different survey methods;
- Imputation of travel modes and travel purposes for trips recorded by GPS;

The travel surveys were conducted in 2010 and 2011 in Hanoi city using two main tools at the same time: a wearable GPS receivers and a travel diary self-reported. The methodology is mentioned as follows: all participants - who are author's relatives, friends, neighborhoods and colleagues - were asked to wear the GPS on all trips undertaken within a week, they were also asked to report their trips in the diary at the end of the day. The data collected consisted of the start time, end time, mode of transport, travel purposes and travel distance. The household information and personal socioeconomic characteristics were also self-reported. The socioeconomic information was: age, gender, educational level, occupation, income, study status, motorcycle/car driving license, the taste and frequency for driving. To remind the participants to always wear the GPS during their trips, text messages have been sent to them every early morning and late evening.

3.3. Trip identification process

The data recorded by the GPS consisted of a lot of raw data including GPS ID, latitude, longitude, date, time and instantaneous speed. Because of using the same GPS devices in France and Vietnam, an unique algorithm was used for trip identification process. Based on the raw data imported to the computer, at the first step, a rule-based algorithms developed by IFSTTAR/DEST was used for identifying trip (Yuan, 2010). The trip identification procedure as follows: a program was run to remove all burden points and create trip ends where the device was stationary for 120 seconds or more. The dwell time is 120s suggested by (Wolf, 2000) based on the traffic signal cycle at intersection is always less than 120s according to the US Highway Capacity Manual. The second program was run that created trip information, including origin, destination, travel distance, starting time, end time, and travel time. All identified trips were exported to one data table as .csv format.

- France:

957 volunteers participated in the pilot GPS survey within a dozen regions in France. This sample size was lower than expected. Finally, the description of trips was both CAPI reported and GPS recorded for 529 days with the participation of only 327 respondents. The travel dataset contains 2294 daily trips reported by CAPI and 2851 GPS trips.

Among 2294 daily trips reported by CAPI, there are 1288 daily trips which matched with GPS trips and 1006 unmatched. Among the 2851 trips recorded by GPS, there are 1563 unmatched trips with CAPI.

In the travel survey tests in Vietnam, we have a total of 95 participants and 1150 matching trips among 1872 diaries reported trips and 1680 GPS recorded trips. All participant had travel data in their GPS.

Table 1. Number of matched and un-matched trip in France and Vietnam

Number of trip	France				Vietnam			
	Matched	Unmatched	Total	% un-matched	Matched	Unmatched	Total	% un-matched
By Diary	1288	1006	2294	44%	1150	722	1872	39%
By GPS	1288	1563	2851	55%	1150	530	1680	32%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08 and travel survey test in Vietnam.

The unmatched percentage of survey reported and GPS recorded trips in France are respectively 44% and 55% while in Hanoi Vietnam are only 39% and 32%. This lower unmatched percentage in Vietnam is not only due to the higher accessibility rate by respondents (because they are relatives, friends, neighborhoods or colleagues of the author) but also to the improvement in survey methodology. In fact, when the GPS devices and travel diaries were given to the respondents, they always received text messages early in the morning and late in the day to remind them for wearing GPS and reporting trips.

Table 2 shows the distribution of unmatched trip by different group of respondents, in which, we have considered 2 main groups: the trip recorded by GPS without diaries reporting and the trip reported without GPS trace. In each main group, the comparison of unmatched percentage between France and Vietnam is shown.

In the descriptive analysis, the variables were divided into 2 groups:

- Demographic factors : sex, age, job characteristics (full time or part time), motorization (0-1-2 and 2+), health conditions (good, medium, bad);
- Travel characteristics: number of trip undertaken, trip distance, departure time, travel purposes, travel mode, travel day - the day of week that travel was recorded.

Table 2: Characteristics of respondents in 2 surveys

Variables	GPS trips (missing diaries reported)		Diaries trips (missing GPS recorded)	
	Fr	VN	Fr	VN
<i>Total trips</i>	2851	1680	2294	1872
Number of unmatched trip	1563	530	1006	722
% of unmatched	55%	32%	44%	39%
<i>Gender</i>				
Male	59%	33%	42%	33%
Female	51%	31%	46%	45%
<i>Age</i>				
<30 years	48%	27%	41%	36%
30-50 years	54%	35%	43%	48%
>50 years	61%	50%	49%	36%
<i>Employment status</i>				
Full time	54%	33%	42%	38%
Part time	53%	24%	47%	40%
Not employed	58%	33%	45%	40%
<i>Household Motorization</i>				
0	69%	21%	36%	46%

1	54%	33%	46%	40%
2	51%	21%	43%	22%
>2	56%	NA*	42%	NA
<i>Health conditions</i>				
Good	55%	28%	43%	40%
Medium	54%	32%	48%	30%
Bad	50%	71%	20%	47%
<i>Number of trip undertaken in the day</i>				
1-2	15%	14%	28%	39%
3-4	40%	24%	42%	43%
5-6	44%	34%	49%	39%
> 6	67%	43%	47%	26%
<i>Departure time</i>				
Before 9h	59%	18%	41%	41%
9h-11h30	52%	30%	49%	40%
11h30-14h00	56%	40%	41%	47%
14h -17h	56%	42%	46%	38%
17-21h	51%	33%	41%	33%
21-24h	65%	39%	48%	266%
<i>Trip purposes</i>				
	NA	NA		
Home	#	#	42%	40%
Shopping	#	#	56%	61%
Entertainment	#	#	38%	28%
Professional	#	#	32%	26%
Others	#	#	46%	53%
<i>Trip modes</i>				
	NA	NA		
Moto	#	#	41%	44%
Car	#	#	41%	8%
Bike	#	#	60%	60%
Walk	#	#	55%	42%
Bus	#	#	64%	23%
Others	#	#	25%	49%
Metro	#	#	42%	NA
Train	#	#	20%	NA
<i>Trip distance</i>				
< 2 km	72%	50%	54%	59%
2- 5 km	42%	21%	51%	36%
5-10 km	35%	26%	43%	33%
> 10 km	24%	17%	31%	24%
<i>Day of trips</i>				
Weekday from Monday to Friday	51%	30%	45%	38%
Weekend	58%	39%	44%	40%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08 and travel survey test in Vietnam.
NA: non available

In terms of demographics, there are not differences among employment status, motorization and health conditions of respondent, both in France and Vietnam. But the younger people tend to record more trip by GPS than the elderly. This finding is confirmed by (Bricka et al., 2012) in their research of factors influencing differences in survey-reported and GPS-recorded trips.

In term of mobility, the person with high travel frequency tend to miss more trips by GPS than those who travel less (e.g. 67% unmatched GPS trips for people making more than 7 trips per day, compared to 15% for those making 1-2 trip per day). The trip distance seems also to have an influence with 70% unmatched trips under 2km while this percentage is only 24% over 10km.

For the unmatched reported trip, there are similar difference in France and Vietnam (44% and 39%). The elders are more likely to have unmatched reported trip than the younger. When people are more mobile and make longer trips, rather for work or education than for shopping or personal purpose, the number of unmatched trip is lower.

4. Factors influencing of GPS recorded and survey reported trips

To identify the impacts on the unmatched trip, a logit binary model was used, in which the dependent variable is possibility of unmatched trip (i.e. Yes - No) from each of two types unmatched: (i) the GPS trips but no survey trip and (ii) the survey trips without GPS trip. The independent variables are demographic and mobility characteristics of the respondents. By using a logit binary regression (logistic regression), the explicative variables could be estimated. These independent variables could be one or more nominal, interval or ratio level independent variables.

By comparing the survey reported trips and survey recorded trips, 3 types of trip were identified:

- matched trips;
- GPS unmatched trips : trips recorded by GPS but not reported on diary
- survey unmatched trips: trips reported in diary but no recorded by GPS

4.1. Identification of the non-recorded GPS trips

The non- recorded GPS trips are trips reported in the travel diary but have not been detected by the GPS data loggers. We previously counted 2294 daily trips reported by CAPI of which 1006 trips are not matched with GPS traces in the French survey and 722/1872 unmatched trip reported in Vietnam. To better understand the factors affecting the difference between GPS and diary trips, we use a binary logistic regression with the dependent variable : matched (or unmatched) trip. Independent variables are characteristics of households, personal information and travel behavior.

Table 3 shows the factors that affect unmatched trips between GPS and diary surveys. These are trip distance, pattern, number of trips made the day before the interview and the characteristics of the respondent such as gender, occupation, state of health, type of household to which the respondent belongs, number of cars in the household, mode of travel. However, the respondent's age and zone of residence do not seem to have any influence.

Table 3. factors influencing the non-recorded GPS trips

Factors	France			Vietnam		
	DF	Wald Khi-2	Pr > Chi-2	DF	Wald Khi-2	Pr > Chi-2
Trip purposes	4	32.30	<.0001			NS
Trip distance	3	42.55	<.0001	3	107.6059	<.0001
Number of trips undertaken the day before	3	20.94	0.0001	3	11.2687	0.0104
Occupation of respondent	7	28.11	0.0002			NS
Household types	4	17.11	0.0018			NA
Gender	1	7.40	0.0065			NS
Healthy situation	3	10.69	0.0135			NS
Motorisation level	3	9.52	0.0231			NS
Means of transport	6	14.06	0.0290	6	39.62	<.0001
Starting time			NS	5	29.1928	<.0001

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08 and travel survey test in Vietnam.

NS: non significant

NA: non available

Trip distance has an important impact on unmatched reported trips: the percentage of unmatched for trips under 2km is 2.2 times higher than for trips over 10km in France and 5.2 times in Vietnam.

The next factors is trip frequency: the more trips are undertaken per day, higher is the unmatched percentage. The means of transport is also one of factors influencing the non-recorded trips (e.g. few trips by public transport).

Table 4 shows the odds ratio when matching survey trips and GPS trips, it is the percentage of non recorded trip of each groups comparing with the reference group.

Table 4: Odds ratios for matching survey trips and GPS trip (France)

Variable/Item	Odds ratios	Pr > Khi-2
<i>Trip purpose</i>		
Home	0.869	0.9268
Shopping	1.454	<.0001
Leisure	0.681	0.0367
Professional	0.595	0.0020
Other	Reference	
<i>Trip Distance</i>		
<2 km	2.191	0.0035
2 to 5 km	2.141	0.0025
5 to 10 km	1.643	0.8700
>= 10 km	Reference	
<i>Number of daily trips the day before</i>		
1--2	2.191	<.0001
3--4	0.955	0.2804
5--6	1.139	0.0009
>6	Reference	
<i>Profession</i>		
Farmer	2.191	0.4679
Craftman, tradesman	0.337	0.0416
Senior executive	1.460	0.0757
Intermediate Profession	1.885	0.0001
Employee	0.911	0.1314
Blue colar	1.351	0.2132
Retired	1.489	0.0387
Other non active	Reference	
<i>Type of household</i>		
Living alone	2.191	0.0041
Single parent family mono parental	0.372	0.1096
Couple without children	0.422	0.1202
Couple with children	0.530	0.5082
Other	Reference	
<i>Gender</i>		
Male	2.191	0.0065
Female	Reference	
<i>Health condition</i>		

Very good	2.191	0.8841
Good	2.961	0.3090
Medium	3.422	0.1449
Bad	Reference	
<i>Number of cars in the household</i>		
0	2.191	0.0114
1	1.141	0.0099
2	0.995	0.2903
>2	Reference	.
<i>Trip Mode</i>		
Other	2.191	0.2720
Bus	1.825	0.0559
Underground	0.777	0.6225
Train	0.304	0.2391
Motorcycle/Car	0.541	0.5177
Walking	0.832	0.2698
Bicycle	Reference	

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Table 5: Odds ratios for matching survey trips and GPS trip (Vietnam)

Variable/Item	Odds ratio	95% Confidence Interval of Wald	
<i>Departure time</i>			
Before 5:00	<0.001	<0.001	>999.999
From 5:00 to 9:59	2.633	1.402	4.943
From 9:00 to 11:59	1.792	0.941	3.412
From 12:00 to 16:59	2.425	1.295	4.541
From 17:00 to 20:59	1.401	0.754	2.602
After 21:00	Reference		
<i>Travel distance</i>			
< 2km	2.191	3.732	7.455
2- 5km	2.136	1.539	2.966
6-10km	1.622	1.173	2.243
>10km	Reference		
<i>Number of trip undertaken per day</i>			
1-2	1.490	1.096	2.026
3-4	1.553	1.183	2.038
5-6	1.252	0.930	1.685
>7	Reference		

Source : Travel survey test in Vietnam.

4.2. Reasons for the missing trips in GPS

For a better understanding about the reasons of non-recorded trips, we have asked the respondent in the FNHTS. So, the most important reason is not wearing the GPS (device forgotten at home or office, or shopping without GPS) with 60% of total answers. But in Vietnam this question has not been asked.

Table 6 : Reasons for the trips missed by GPS in France*

Reasons	Frequency	%
Forgotten GPS	441	44%
Cold start in the morning	110	11%
Shopping without GPS	86	9%
GPS left at the office	74	7%
Other reasons	295	29%
Total	1006	100%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

*: this information is not available in Vietnam GPS test.

Table 6 shows that a half of the missing trips are related to human failures, among them 44% correspond to the GPS device forgotten at home. There are 7% of people who forget the GPS at the offices when they go out for lunch or for professional purposes. 11% of the daily missing trips are due to a cold start or a restart requiring a period of time to capture the signals of the satellites, especially during the first trip in the morning. 9% of missing trips are related to shopping purpose, it means that in a shopping center, it is hard to receive the satellites signals.

4.3. Identification of non-reported survey trips

The non-reported survey trips are those detected by GPS but which were not reported in the travel diary. As shown in table 1, we have found 2851 GPS trips including 1563 GPS trips unmatched with the travel diary for France. For Vietnam, there are 530 trips omitted among 1680 GPS recorded trips. For a better understanding, we estimate a binary logistic regression on each survey.

Table 7 : Factors influencing the trips omitted in the diary

Factors	France			Vietnam		
	DF	Wald Khi-2	Pr > Chi-2	DF	Wald Khi-2	Pr > Chi-2
Number of GPS trips	3	166.7260	<.0001	3	41.697	<.0001
Speed of GPS trips	4	51.2141	<.0001	4		NS
Duration of GPS trips	4	102.5806	<.0001	1	7.838	0.0051
Occupation of respondent	7	20.2003	0.0052	7		NS
Distance of GPS trips	3	11.5147	0.0092	3	33.389	<.0001
Start time	3		NS	5	20.925	0.0008

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08 and travel survey test in Vietnam.

NS: non significant

Table 8 provides a detailed analysis for non reported trips in France. Among the influencing factors for omitting trips in the diary, the number of trip recorded by GPS is the most important factor. Individuals who make frequent trips tend to have more non-reported trips, the ratio of non-reported trips is 10 times lower for a day with 1 or 2 trips than for a day with 7 trips or more. Travel duration is also one of the factors influencing the non-reported trips: the rate of non reporting for trips under 5 minutes is 12 times higher than for trips over than 60 minutes. Personal occupation has only a light impact : employees and retired persons have more non reported trips.

For Vietnam, we have found similar results: short distance trips are under-reported (3 times more under 2 km than over 10 km), as well as short duration ones. Besides, the start time could have an influence : trips started early in the morning (from 5h00 to 9h00) are more usually reported.

Table 8 : Odds ratios for matching GPS recorded and diary reported trips in France

Variable/Item	Odds ratios	Pr > Khi-2
<i>Number of traces the day before</i>		
1--2	0.093	<.0001
3--4	0.420	0.2014
5--6	0.481	0.0106

> 6	Reference	
<i>Speed on the trace (in km/h)</i>		
< 7	1.885	<.0001
7--20	1.404	0.0054
20--50	0.689	<.0001
50--80	0.671	0.0037
> =80	Reference	
<i>Trace Duration</i>		
< 5 min.	1.885	<.0001
5 to 15 min.	6.221	<.0001
15 to 30 min.	3.733	0.5999
30 to 60 min.	1.911	0.0003
> = 60 min.	Reference	
<i>Linear distance</i>		
< 2 km	1.885	0.0012
2 to 5 km	1.026	0.2997
5 to 10 km	1.023	0.3330
>= 10 km	Reference	
<i>Profession</i>		
Farmer	1.885	0.6236
Craftsman, tradesman	1.181	0.2896
Senior executive	1.235	0.2228
profession	1.454	0.9641
Employee	2.462	0.0004
Blue collar	1.607	0.4775
Retired	2.351	0.0093
Other non active	Reference	
<i>Average number of daily trips from trip diaries</i>		
0	1.885	0.0040
1--2	2.926	<.0001
3--4	2.168	0.0892
5--6	1.335	0.0003
>=7	Reference	

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Table 9. Odds ratio for matching GPS trip and survey trips in Vietnam

Varibale/Item	Odds-ratio	Pr > Khi-2
<i>Departure time</i>		
Before 5:00	NS	0.2831
From 5:00 to 9:59	0.621	<.0001
From 9:00 to 11:59	1.274	0.3497
From 12:00 to 16:59	2.078	0.2656
From 17:00 to 20:59	1.334	0.9145
After 21:00	Reference	

Travel distance

< 2km	2.643	1.885
2- 5km	0.948	0.0031
6-10km	1.434	0.7295
>10km	Reference	
<i>Average number of daily trips</i>		
1-2	0.301	<.0001
3-4	0.52	0.3638
5-6	0.686	0.0933
>= 7	Reference	

Source : Travel survey test in Vietnam.

4.4. Comparison between GPS recorded trip and survey reported trips

Table 10. Comparison between the characteristics of GPS recorded trips and of survey reported trips

Absolute difference	France	Vietnam
<i>Departure time</i>		
From 0 to 2 minutes	17%	19%
From 2 to 5 minutes	16%	23%
From 5 to 10 minutes	17%	21%
More than 10 minutes	50%	37%
<i>Arrival time</i>		
From 0 to 2 minutes	21%	35%
From 2 to 5 minutes	21%	21%
From 5 to 10 minutes	14%	16%
More than 10 minutes	44%	28%
<i>Travel duration</i>		
From 0 to 2 minutes	21%	20%
From 2 to 5 minutes	24%	22%
From 5 to 10 minutes	23%	22%
From 10 to 15 minutes	11%	14%
More than 15 minutes	21%	20%
<i>Travel distance</i>		
From 0 to 0.5km	17%	28%
From 0.5 to 1.0km	12%	16%
From 1-2km	16%	20%
From 2-5km	26%	23%
More than 5km	30%	13%
Total	100%	100%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08 and travel survey test in Vietnam.

The results in table 10 for comparison between the trips collected by GPS and surveys show the following differences between the GPS recorded and survey reported trips:

- The departure time reported in the diary is earlier than that recorded by GPS for about 70% of trips. However in France, the difference exceeds 10 minutes for half of trips, while in Vietnam is only for 37%.

- The difference for ending time is lower than for departure time: In France, 44% of the difference in ending time between these survey is more than 10 minutes, while this percentage in Vietnam is only 28%. The smaller difference for ending time than for departure time may be due to the Time necessary to first fixe (TTFF) that is a measure of the time required for a GPS receiver to acquire satellite ...

- People tend to over-estimate the travel duration (Noble, 2001; Hubert, 2003). This research presents similar results : the reported travel duration is overestimated for about 80% of trips. However, the difference between GPS and diary is less than 15 minutes in France as well as in Vietnam for about 4/5 of trips.

The travel distance reported in the diary is usually longer than travel distance recorded by GPS but the difference is usually less than 5 km (for 30% of trips in France and only 13% in Vietnam).

5. Conclusion

In each of two contrasted contexts (i.e. the National Travel Survey 2007-08 in France and a GPS pilot survey in 2010-11 in Hanoi), two descriptions of daily mobility were collected for the same individuals during the same days: a travel diary and the GPS logs registered by a receiver worn by the respondent. Both surveys provide two descriptions of daily mobility (the travel diary and a post-processing of GPS logs). The first step consisted in matching diary reported and GPS recorded trips.

The paper aimed at two main objectives. First, it examined the influencing factors for the unmatched trips by using a logistic regression. Secondly, it analyzed the differences in the characteristics of travel for matched trips. Despite of very contrasted backgrounds in terms of socio-cultural context, of sampling methods (volunteers in a National travel survey vs. snow-ball approach, etc.) the results are convergent. The findings from this research can inform both survey-reported as well as GPS-recorded travel data collection approaches.

In terms of unmatched trips (non-recorded as well as non-reported), trip distance and the frequency of trip making are the main influencing factors: there are more unmatched for short trips and for days with a high number of trips. Male, those are employee or retired are more likely unmatched trip.

For matched trips, departure time, arrival time, trip duration and trip distance were compared between diary and GPS. Because of cold start, the departure time recorded by GPS is generally later than that reported in the diary, while there is less difference between the two measurements for arrival time. Trip duration and trip distance are over-estimated in the diary for about 80% of trips, but the difference is usually less than 15 minutes or 5km.

Overall, there are large difference in the measurement of daily mobility between GPS recorded and survey reported trips. these results suggest that improvements are needed in both methods to yield more accurate and unbiased data. For this reason, we recommend that GPS-only studies be undertaken with caution until methodological improvements resulting in more consistent data being obtained from both sources.

Acknowledgement

This paper results from a comparison between the results obtained by Thi Huong Thao Pham in chapter 6 of her PhD thesis (2016) and of the PhD thesis by Thanh Tu Nguyen (2013). It has been completed by Thanh Tu Nguyen within the inviting researcher program of the French institute of science and technology for transport, spatial planning, development and networks. The main authors would like to thank not only the support of the IFSTTAR but also the colleagues at the Laboratory "Economic and Social Dynamics of Transport" for their precious contributions and comments.

References

- Armoogum, J., 2000. Correction de la non-réponse et de certaines erreurs de mesure dans une enquête par sondage : Application à l'enquête transports et communications 1993-94 (Thèse de doctorat). Université Libre de Bruxelles.
- Bricka, S., Bhat, C., 2006. Comparative Analysis of Global Positioning System-Based and Travel Survey-Based Data. *Transp. Res. Rec. J. Transp. Res. Board* 1972, 9–20. doi:10.3141/1972-04
- Bricka, S.G., Sen, S., Paleti, R., Bhat, C.R., 2012. An analysis of the factors influencing differences in survey-reported and GPS-recorded trips. *Transp. Res. Part C Emerg. Technol.* 21, 67–88.
- Casas, J., Arce, C.H., 1999. Trip reporting in Household Travel Diaries: A Comparison to GPS-collected data. *Transportation Research Board, 78th Annual Meeting, CD-ROM, Washington, DC.*
- Forrest, T., Pearson, D., 2005. Comparison of Trip Determination Methods in Household Travel Surveys Enhanced by a Global Positioning System. *Transp. Res. Rec.* 1917, 63–71.
- Houston, D., Luong, T.T., Boarnet, M.G., 2014. Tracking daily travel; Assessing discrepancies between GPS-derived and self-reported travel patterns. *Transp. Res. Part C Emerg. Technol.* 48, 97–108. doi:10.1016/j.trc.2014.08.013

- Hubert, J.-P., 2003. Travel and time use surveys, a comparison of two belgian surveys. In: 25th IATUR Conference on Time Use Research, 17–19 September, Brussels.
- Noble, B., 2001. Using simple time use surveys to investigate travel. In: 6th International Conference on Survey Methods in Transport. Kruger Park, South Africa.
- Pham, T.-H.-T., 2016. Apports et difficultés d'une collecte de données à l'aide de récepteurs GPS pour réaliser une enquête sur la mobilité. Thèse de doctorat. Université Paris Est, Marne la Vallée.
- Roux, S., 2012. Transition de la motorisation en France au XXème siècle (Thèse de doctorat). Université de Paris 1 Panthéon Sorbonne.
- Stopher, P., FitzGerald, C., Xu, M., 2007. Assessing the accuracy of the Sydney Household Travel Survey with GPS. *Transportation* 34, 723–741. doi:10.1007/s11116-007-9126-8
- Stopher, P., Shen, L., 2011. An in-depth comparison of GPS and diary records. Presented at the Transportation Research Board Annual Meeting, 90th, 2011, Washington, DC, USA.
- Thanh Tu NGUYEN, 2013. Mise au point d'une méthode de collecte de données de mobilité en utilisant des récepteurs GPS qui soit comparable avec les enquêtes classiques et applicable dans les pays du Sud. Thèse de doctorat. Université Paris 1 Panthéon-Sorbonne, Paris, France.
- Wolf, J., 2006. Applications of New Technologies in Travel Surveys, in: *Travel Survey Methods*. Emerald Group Publishing Limited, pp. 531–544.
- Wolf, J., 2000. Using GPS data loggers to replace travel diaries in the collection of travel data. *Diss. Ga. Inst. Technol. Sch. Civ. Environ.* 58–65.
- Wolf, J., Loechl, M., Thompson, M., Arce, C., 2003. Trip Rate Analysis in GPS-Enhanced Personal Travel Surveys, in: *Transport Survey Quality and Innovation*. Elsevier, Amsterdam : Oxford :, pp. 483–498.
- Yuan, S., 2010. Méthodes d'analyse de données GPS dans les enquêtes sur la mobilité des personnes - les données manquantes et leur estimation. Université Paris 1 Panthéon-Sorbonne, Paris.
- Zmud, J., Wolf, J., 2003. Identifying the Correlates of Trip Misreporting –Results from the California Statewide Household Travel Survey GPS Study. Presented at the International Conference on Travel Behaviour Research, Lucerne.